International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

October 2023
Volume 34
Number 3
ISSN 1650-3414

Communications and Publications Division (CPD) of the IFCC
Editor-in-chief: Prof. János Kappelmayer, MD, PhD
Faculty of Medicine, University of Debrecen, Hungary
e-mail: ejifcc@ifcc.org

INFORMATION FOR CONTRIBUTING AUTHORS

In order to cope with the increase in submissions to the eJIFCC we decided to advance our manuscript handling process, and we subscribed for the Editorial Manager manuscript handling system. Each author has to register for the program. For the registration there are two possible ways: with an ORCID ID, or with an email address. To finalize the registration, a few questions will be asked that will be necessary for later communication and work-flow handling. The following questions can be expected during the end of the registration: workplace, contact details (email address, phone number), professional qualifications and availability lists. You will receive a system message about the successful registration. If you encounter any problems, or registration fails, log in to the site or if you would like to initiate the deletion of your account, you will be able to communicate it to us by using the email at the bottom of this page. After a plagiarism and a general technical check by the Editoral Office, manuscripts will be forwarded for evaluation to Reviewers. If your manuscript is not found acceptable you will receive detailed commnents that may help you to publish your work elsewhere. If your manuscript requires revision you will always be provided a deadline that should be strictly kept. Once your manuscript gets accepted you will receive the proof two weeks prior to publication.

Instructions for authors

Manuscript preparation: Microsoft Word is preferred, with the use of 12-point font, 2cm margins and double spacing.

The title page should list title and short title, authors (first name, last name), each authors affiliation, including e-mail addresses, corresponding authors contact information, key words.

All manuscript should be provided with structured Abstract not longer than 250 words. This information should proceed the text of the manuscript.

Tables should be created in a common word-processing or spreadsheet format and included within the text of the manuscript, close to the first location where they are referenced. The figures must be submitted as independent images or files, with high enough resolution for publishing and printing. All tables and figures must include a short title, with any further explanatory text to be included at the bottom of either.

The maximum number of references for a review article is 80, and 40 for original papers. Intext reference numbers should be included paranthetically before punctuation. The actual references shall be presented and punctuated consistently and are to be listed sequentially following the main text, with their numbers unlinked.
All references should be provided in the Vancouver style.

Manuscripts should provide any required Authors Disclosures. All studies involving human subjects must indicate that they are in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

For any further questions please contact:
ejifcc.editorialoffice@ifcc.org

Evaluation of four quality indicators of the Pre-Analytical Phase External Quality Assessment Subprogram of the Fundación Bioquímica Argentina
Gisela Unger, Silvia Fabiana Benozzi, Raúl Girardi, Graciela Laura Pennacchiotti 203

Disruption of laboratory activities during the COVID-19 pandemic: results of an EFLM Task Force Preparation of Labs for Emergencies (TF-PLE) survey
Giuseppe Lippi, Janne Cadamuro, Elisa Danese, Emmanuel J. Favaloro, Julien Favresse, Brandon M. Henry, Snezana Jovicic, Tomris Ozben, Mario Plebani, Jecko Thachil 213

Inter-laboratory method validation of CD34+ flow-cytometry assay: the experience of Turin Metropolitan Transplant Centre
Ivana Ferrero, Deborah Rustichelli, Sara Castiglia, Loretta Gammaaitoni, Alessandra Polo, Marisa Pautasso, Massimo Geuna, Franca Fagioli 220

Phenotype similarities in automatically grouped T2D patients by variation-based clustering of IL-1β gene expression
Lucio José Pantazis, Gustavo Daniel Frechtel, Gloria Edith Cerrone, Rafael García, Andrea Elena Iglesias Molli 228

Best Practices for Effective Management of Point of Care Testing
Anil K Chokkalla, Brandy D Recio, Sridevi Devaraj 245

Monoclonal Light Chains with alpha 2 mobility on Serum Protein Electrophoresis
S. Danalakshmi 250

Coffee Colored Serum, Adverse reaction of Eltrombopag
Carlos Rodríguez Rojas, Celia Juez Santamaría, Oscar David Pons Belda, Emilia Moreno Noguera 258

The parallel lives of prostate specific antigen in cardiac troponin assays
Eleftherios P. Diamandis 262

Evaluation of four quality indicators of the Pre-Analytical Phase External Quality Assessment Subprogram of the Fundación Bioquímica Argentina

Gisela Unger[1,2*], Silvia Fabiana Benozzi[1,2*], Raúl Girardi[3*], Graciela Laura Pennacchiotti[1,2*]

[1]*Cátedra de Bioquímica Clínica I, Departamento de Biología, Bioquímica y Farmacia; Universidad Nacional del Sur; San Juan 670, 8000 Bahía Blanca, Argentina*

[2]*Pre-Analytical Phase External Quality Assessment Subprogram, Fundación Bioquímica Argentina, Calle 148 Nº 584, B1900BVK La Plata, Argentina ;*

[3]*External Quality Assessment Program, Fundación Bioquímica Argentina, Calle 148 Nº 584, B1900BVK La Plata, Argentina*

Article Info

Author of correspondence:
Dr. Graciela Laura Pennacchiotti;
E-mail: grapen@uns.edu.ar;
Address:
Universidad Nacional del Sur, San Juan 670, 8000 Bahía Blanca, Argentina

Keywords

pre-analytical phase, pre-analytical phase quality, quality indicators, external quality assessment programs

Abstract

Pre-analytical phase external quality assessment programs contribute - through the interlaboratory comparison of quality indicators (QIs) - to the continuous improvement of the clinical laboratory total testing process. The purpose of the present work is to document the results derived from measuring four QIs within the framework of a pre-analytical phase external quality assessment subprogram in Argentina. The laboratories participating in this subprogram measured the following QIs: i) patients recalled for a new blood sample collection due to pre-analytical causes; ii) clotted samples from hemogram and coagulation tests; iii) clinical chemistry hemolyzed samples; and iv) requests with transcription errors entered into the laboratory information system. Results were expressed in percentage value and Sigma value. Databases were anonymized. A minimum acceptable quality level for the four QIs measured was recorded in the majority (75%) of the participating laboratories (Sigma > 3.0). It was nonetheless observed that the QIs of hemolyzed samples and requests with transcription errors entered into the laboratory information system deserve more attention. Through this pioneering experience in Argentina, the participating laboratories - some for the first time - could learn about their performance via interlaboratory comparison of results. This experience also proved to be motivating not only to improve the external assessment subprogram but also to continue working on the measurement of pre-analytical QIs for the continuous improvement of the clinical laboratory total testing process in Argentina.

Introduction

Over time, the concept of error in the clinical laboratory has evolved from a model focused primarily on the analytical phase to a model focused on errors that occur all throughout the clinical laboratory total testing process (TTP), including extra-analytical phases (1,2). The pre-analytical phase, in particular, is a stage that is not only key but also high-risk for patients as a result of its complexity and the variety of procedures and factors that it involves. Evidence has, in fact, been documented that errors in this phase represent up to 70% of the total errors in the TTP (3). For this reason, both internal and external quality control of the pre-analytical phase is of paramount importance to ensure clinical utility of the results issued by the laboratory (1,4). Risk analysis, systematic error detection and the implementation of pre-analytical quality indicators (QIs) are sine qua non requisites for an effective internal quality control (5). QIs have proven to be an effective tool to monitor processes as well as the efficacy of the corrective and/or preventive actions implemented, and therefore their measurement is required to comply with different clinical laboratory accreditation standard regulations (5-7). QIs are also objective tools with which it is possible to assess what happens in the laboratory during the TTP, thus allowing self- and inter-comparison among laboratories measuring the same QI (7). Still, the main recommendation is that, depending on their resources, laboratories should prioritize the most useful QIs for their processes to avoid unnecessary extra work that may limit the continuity and usefulness of their monitoring (5). By participating in pre-analytical phase external quality assessment programs, laboratories have the opportunity to access to documented and objective tools to achieve continuous process improvement through interlaboratory comparison (4,8). In the last 20 years, clinical laboratory societies worldwide have developed external quality assessment programs for the pre-analytical phase based on different types of strategies, namely: i) procedure recording (Type I-strategy applied in Spain, Norway, Germany, Finland); ii) distribution of simulating error samples (Type II-strategy applied in Denmark, United Kingdom, Switzerland, Sweden, Italy, Austria, Luxembourg, France, the Netherlands, Spain); and iii) QIs recording (Type III-strategy applied in the United States, Australia, Spain, Norway, United Kingdom, Brazil) (1,4,9). Among the programs that apply Type III-strategies is the project "Laboratory Errors and Patient Safety" (WG-LEPS) of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) (10). The WG-LEPS developed a harmonized model of QIs for the clinical laboratory TTP as well as a project to record them through an online platform within which laboratories from all over the world can record their measurements and evaluate their performance by interlaboratory comparison (8,10). In Argentina, the Pre-Analytical Phase External Quality Assessment Subprogram (preEQA Subprogram) was created in 2016 within the framework of the External Quality Assessment Program (EQA Program) of the Fundación Bioquímica Argentina (11) with the purpose of contributing to the continuous improvement of the clinical laboratory TTP. The specific aims of the preEQA Subprogram are to provide professional updating on the pre-analytical phase and to carry out interlaboratory comparison of procedures (Type I-strategy) and QIs (Type III-strategy) taking into account the context and characteristics of the participating laboratories (4,9). Based on the above, the purpose of the present work was to report the results collected from the first interlaboratory comparison of four QIs carried out by the preEQA Subprogram in Argentina during 2021 and 2022.

Materials and methods

Characteristics of the preEQA Subprogram

To be eligible for participation in the preEQA Subprogram, laboratories should be registered in the EQA Program and should apply for registration in the preEQA Subprogram. The latter consists of four annual surveys conducted through the Fundación Bioquímica Argentina web page (11). As both participation in the preEQA Subprogram and survey response submission are voluntary, the number of participants and responses received varies with time. After each survey, the participating laboratories receive a report of the results collected together with a commentary on the analysis and interpretation of the results, with recommendations and related literature.

QIs evaluated in the preEQA Subprogram

The preEQA Subprogram proposed to the participating laboratories to measure four QIs. Instructions on the registration of each QI were first delivered to the participating laboratories and the measurement of these QI was subsequently carried out following the survey schedule designed by the EQA Program for the years 2021 and 2022. The QIs evaluated were designed based on the IFCC QI model (10) with self-adaptations to facilitate their measurement taking into account previously observed characteristics of the laboratories participating in the preEQA Subprogram (12). This is the reason why the four QIs were measured only in the outpatient setting to standardize the interlaboratory measurement and with a bimonthly periodicity so that the small laboratories could obtain a significant number of records.

Table 1 lists the QIs evaluated chronologically.

Table 1: Quality indicators evaluated in the preEQA Subprogram

Quality Indicator (QI)	Data Collection Period	Formula %
NS-QI: percentage of patients recalled for a new blood sample collection due to pre-analytical causes	May-June/2021	= 100 x (number of patients recalled for a new blood sample collection due to pre-analytical causes / total number of patients of the laboratory)
CS-QI: percentage of clotted samples	October-November/2021	= 100 x (number of clotted samples from hemogram and coagulation tests / total number of samples from hemogram and coagulation tests)
HS-QI: percentage of hemolyzed samples	April-May/2022	= 100 x (number of hemolyzed clinical chemistry samples / total number of clinical chemistry samples)
TE-QI: percentage of requests with transcription errors entered into the LIS	October-November/2022	= 100 x (number of requests with transcription errors entered into the LIS / total number of requests entered into the LIS)

preEQA Subprogram: Pre-Analytical Phase External Quality Assessment Subprogram;
QI: quality indicator; LIS: laboratory information system

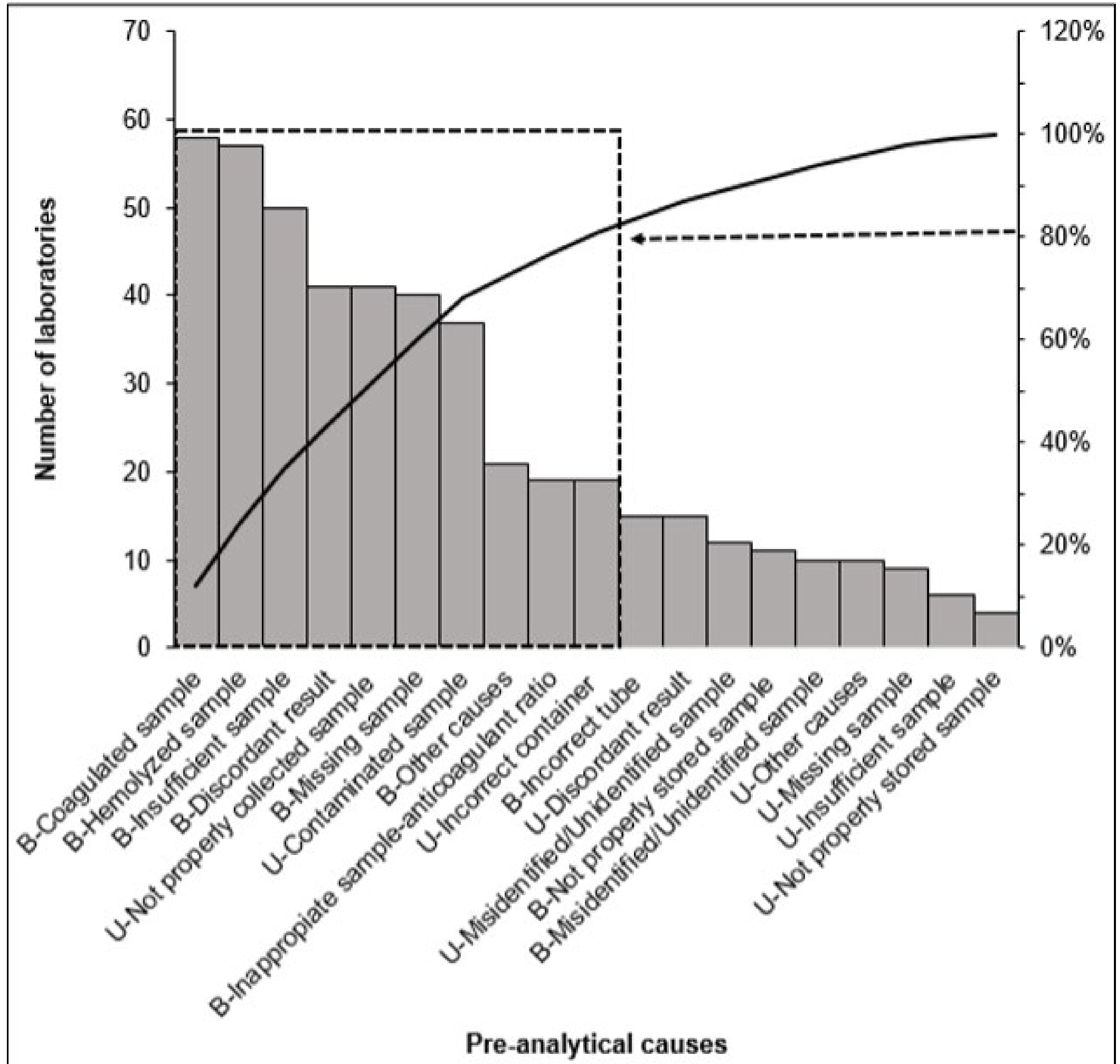


Figure 1: Pre-analytical causes for which patients were recalled for a new blood sample collection and number of laboratories that reported each of these causes for patient recalling during NS-QI measurement

Pareto Chart. Abscissa: pre-analytical causes predefined by the preEQA Subprogram for which patients were recalled for a new blood sample collection. Bars and left axis: number of laboratories that reported each pre-analytical cause for recalling patients for a new blood sample collection. Continuous line and right axis: accumulated percentage. Dotted lines: causes of patient recalling which, based on Pareto principle, have highest impact on the laboratories that responded to the survey. NS-QI: percentage of patients recalled for a new blood sample collection due to pre-analytical causes; B: blood sample; U: urine sample.

The percentage of patients recalled for a new blood sample collection due to pre-analytical causes (NS-QI) was designed for a first assessment of the pre-analytical phase among the laboratories participating in the preEQA Subprogram. The participating laboratories were asked to record the NS-QI as well as the pre-analytical causes due to which patients were recalled. A predefined list of pre-analytical causes, which are shown in Figure 1, was distributed to all the participating laboratories in order to standardize this registry. In addition, the laboratories were asked to group the pre-analytical causes that were not included in this predefined list under the category "other causes". The NS-QI was adapted for the present study from that proposed by the IFCC QI model. According to the latter, the number of patients recalled for a new blood sample collection as a result of any type of error should be measured as an indicator of the outcome of the clinical laboratory TTP (10,12). In the present study, the pre-analytical causes for which patients were recalled for a new blood sample collection were analyzed using Pareto´s principle in an Excel chart (Microsoft Office). This principle focuses efforts on either the causes or the factors that impact most on a given process by stratifying them by frequency and by considering that 80% of problems stem from 20% of causes (12,13). Based on this analysis, two QIs were selected in principle to evaluate the quality of the sample collection process, namely the percentage of clotted samples (CS-QI) and the percentage of hemolyzed samples (HS-QI) (5). As for the clotted samples, the IFCC QI model proposes to measure the CS-QI on either all the samples or tubes with anticoagulant that are checked for the presence of clots (10). In contrast, the preEQA Subprogram ordered to consider only all hemogram samples (EDTA tubes) and all coagulation test samples (citrate tubes) to facilitate the recording of this QI to those laboratories that are beginners in this process. In the case of the HS-QI, all laboratories, including those with an automated hemolysis index, were asked to use the same color scale (14) - which had been provided with the instructions by the preEQA Subprogram - to detect the presence of hemolysis by visual inspection and thus standardize its recording. In the present study, a sample was considered to be hemolyzed when its color was equal to or higher either than that of tube 2 of this scale or than 0.5 g/L of free hemoglobin (10,14). The IFCC QIs model determines that 0.5 g/L of free hemoglobin is the cut-off point for visual inspection and proposes to measure HS-QI on all samples that are checked for hemolysis (10). Therefore, in order to help the participating laboratories to record this QI, the preEQA Subprogram asked to use clinical chemistry plasma samples (tubes with heparin) and clinical chemistry serum samples. The percentage of requests with transcription errors entered into the laboratory information system (TE-QI) was subsequently evaluated in order to measure the quality of the administrative process (5). On account of the fact that in Argentina the majority of laboratories receive requests written in handwriting on paper, for this QI the participating laboratories were asked to consider requests entered into their information system with any type of transcription error (e.g., in patient or physician data, in omitted, wrong or added tests, etc.) detected outside the administrative entry procedure (i.e., detected at the moment of blood collection, sample processing, results validation, report delivery, in reply to the patient or physician´s claim, etc.). In contrast, in the IFCC QIs model, only test transcription errors are considered (10). Apart from the instructions, the participating laboratories were also provided with an Excel worksheet (Microsoft Office) which was designed to record each QI and included a formula for the percentage calculation of the QIs and the instructions for obtaining the short term Sigma metrics with the calculator available at www.westgard.com (15,16). Every laboratory uploaded its results through Fundación Bioquímica Argentina web page (11) within the established times. Those responses that were implausible (e.g., percentages higher than 100%) were eliminated from the data analysis (12). Percentiles 25 (p25) and 75 (p75) of the percentage value distribution were obtained using the software Statistical Package for Social Science 15.0 (Chicago, IL, USA). Laboratory performance was classified into three levels, namely i) high, i.e. with percentage values ≤ p25; ii) medium, i.e. with percentage values between p25 and p75; and iii) low, i.e. with percentage values ≥ p75) (4,8,17). The Sigma metrics corresponding to the percentage values of p25 and p75 of each QI measured was calculated using the above-mentioned calculator (15,16) in order to guarantee a universal and objective QI assessment. The Sigma metrics relates the defect error rate per million opportunities with the efficiency of the process. Based on this, the following quality levels were considered: i) minimum acceptable quality level equal to Sigma value of 3.0, which corresponds to 6.680% of error and 93.3% of yield; and ii) minimum desirable quality level equal to Sigma value of 4.0, which corresponds to 0.621% of error and 99.4% of yield (8,12,16,18,19). Where possible, the results collected in the present study were compared with those published by the IFCC QIs project in 2023. All the databases used were anonymized for the present study in compliance with the ethical requirements for data privacy and confidentiality (20).

Results

The results from the surveys carried out in the preEQA Subprogram show that 64% of the laboratories that responded to the surveys belonged to Buenos Aires province, 8% to the Autonomous City of Buenos Aires and 6% to Santa Fe province. They also show that the percentage corresponding to the laboratories from the remaining Argentine provinces was lower than those above-mentioned. Many of the participating laboratories (70%) belonged to the private outpatient setting. Approximately 5% of them attended less than 100 patients per month whereas another 5% attended more than 3800 patients per month. 80% of the laboratories reported that, prior to the preEQA Subprogram, they had not performed any external quality control of the pre-analytical phase. As for the pre-analytical procedures that are of interest to the present study, it was observed that i) 96% of the laboratories received medical requests written in handwriting on paper; ii) 80% used syringe and needle (open

system) for blood collection; and iii) 90% detected the presence of hemolysis by visual inspection. Only 40% of the laboratories did report that they systematically recorded pre-analytical errors, the main reasons for not recording them systematically being not knowing how to do it, not having time to do so and having the belief that the number of pre-analytical errors made is not enough to justify recording them.

Table 2 shows the number of laboratories participating in the preEQA Subprogram that measured four QIs and the number of valid responses received for each of them.

Table 2: Number of laboratories participating in the preEQA Subprogram for the measurement of four QIs and number of valid responses per QI

	Percentage of the laboratories that participated within the scheme of the preEQA Subprogram to measure QIs (received responses/total participants)	**Percentage of valid responses (valid responses/received responses)**
NS-QI	26% (113/406)	73% (83/113)
CS-QI	56% (226/407)	100% (226/226)
HS-QI	65% (262/405)	94% (246/262)
ET-QI	52% (240/427)	93% (222/240)

QI: quality indicator; preEQA Subprogram: Pre-Analytical Phase External Quality Assessment Subprogram; NS-QI: percentage of patients recalled for a new blood sample collection due to pre-analytical causes; CS-QI: percentage of clotted samples; HS-QI: percentage of hemolyzed samples; ET-QI: percentage of requests with transcription errors entered into the laboratory information system. A response was considered to be invalid when the value reported was higher than 100%.

Figure 1 shows the number of laboratories that detected the pre-analytical causes for which patients had been recalled for a new blood sample collection during the recording of NS-QI and Pareto principle-based analysis (12,13). Table 3 lists the results collected for each QI in the present study and those reported as quality specifications by IFCC QI project (8).

Table 3: Results obtained for four QIs evaluated in the preEQA Subprogram and their comparison with those reported as quality specifications by the IFCC QI project (8)

	preEQA Subprogram PERFORMANCE LEVEL		**IFCC PERFORMANCE LEVEL (8)***	
	HIGH	LOW	HIGH	LOW
	≤ p25% (p25 CI-95%)% [Sigma]	≥ p75% (p75 CI-95%)% [Sigma]	≤ p25%* (p25 CI-95%)%* [Sigma]	≥ p75%* (p75 CI-95%)%* [Sigma]
NS-QI	≤ 0.180% (0.110-0.340)% [4.5: desirable]	≥ 1.640% (1.210-2.220)% [3.7: acceptable]	Unpublished Data	Unpublished Data
CS-QI	≤ 0.000% (0.000-0.000)% [6.0: desirable]	≥ 0.245% (0.200-0.395)% [4.4: desirable]	≤ 0.126% (0.100-0.150)% [4.6: desirable]	≥ 0.527% (0.407-0.630)% [4.1: desirable]
HS-QI	≤ 0.330% (0.265-0.455)% [4.3: desirable]	≥ 2.578% (1.840-3.239)% [3.5: acceptable]	≤ 0.456% (0.000-0.739)% [4.2: desirable]	≥ 1.650% (1.590-1.820)% [3.7: acceptable]
TE-QI	≤ 0.558% (0.400-0.710)% [4.1: desirable]	≥ 5.702% (4.440-6.660)% [3.1: acceptable]	≤ 0.117% (0.078-1.105)% [4.6: desirable]	≥ 2.217% (1.705-2.518)% [3.6: acceptable]

The medium performance level is defined by the range p25-p75. *Performance level considered as quality specifications in 2023 (8). preEQA Subprogram: Pre-Analytical Phase External Quality Control Subprogram; IFCC: International Federation of Clinical Chemistry and Laboratory Medicine; NS-QI: percentage of patients recalled for a new blood sample collection due to pre-analytical causes; CS-QI: percentage of clotted samples; HS-QI: percentage of hemolyzed samples; TE-QI: percentage of requests with transcription errors entered into the laboratory information system; p: percentile; CI-95%: 95% confidence interval.

Discussion

Laboratories showed interest in participating in the interlaboratory comparison of the QIs proposed by the preEQA Subprogram. In this study, it was observed that the percentage of laboratories that measured the last 3 QIs (i.e., CS-QI, HS-QI and TE-QI) duplicated with respect to that of the first QI (NS-QI) measured. In the IFCC QI project, the number of laboratories that measured each pre-analytical QI during 2021 ranged from 25 to 289 (8); in the Spanish Preanalytical Quality Monitoring Program (SEQC), 72 laboratories participated during 2018-2019 (4); and in the Programa de Benchmarking e Indicadores Laboratoriais from Brazil, the number of responses obtained during 2016-2018 ranged from 34 to 1081 depending on the QI measured (12). Thus, the fact that between 83 and 262 laboratories responded to preEQA Subprogram surveys is indeed extremely encouraging particularly if one takes into account the data collection period, which corresponded to the SARS-CoV-2 pandemic context, mainly the year 2021. A plausible explanation for this - although there is still no published evidence - seems to be the fact that in Argentina the number of small and medium-sized laboratories is larger than that of large laboratories in contrast to other countries where centralized laboratories predominate and serve a large number of patients. The fact that it was nonetheless observed that a high percentage of the laboratories enrolled in the preEQA Subprogram did not measure any Qis, highlights the need to implement measures to determine the reasons for such non-participation. As regards the number of invalid responses received in the preEQA Subprogram, it was observed that the highest percentage of invalid responses corresponded to the NS-QI, which was the first QI measured by the participating laboratories. In contrast, the percentage of invalid responses for the remaining QIs was lower probably because the participating laboratories were already familiarized with the Excel chart and the web page to upload the results collected. Measuring the NS-QI proved to be useful to select future QIs in terms of their degree of priority, particularly those related to sample quality. The Pareto chart showed that, among the 10 pre-analytical causes for which patients were recalled for new blood sample collection and which impacted most on the group of laboratories that responded to the survey, clotted, hemolyzed and insufficient blood samples represented an accumulated frequency of 40%, followed by other causes that will be further taken into account for the planning of future QI measurements. It should be noted that not properly collected urine samples were included among these other causes. This is due to the fact that, although national recommendations discourage 24h urine sample collection tests in Argentina (21), they are very commonly performed (for example, for the measurement of albuminuria and proteinuria and creatinine clearance). In the case of the CS-QI, it was observed that the percentage values recorded in the present study were lower than those reported by the IFCC QI project (8), although in both cases the Sigma metrics indicated that such values achieved a desirable level of quality. In this respect, it seems likely that the higher percentage of clotted samples detected by the laboratories participating in the IFCC QI project is due to a higher availability or use of automated pre-analytical and analytical platforms, which are more sensitive than visual inspection, which is mostly used in preEQA laboratories to detect the presence of clots (4,22-24). In the case of the HS-QI, the p75% value collected in the present study was higher than that reported by the IFCC QI project (8) for this same QI obtained by visual inspection. The difference in the results observed could be attributed to the use of the open system (syringe and needle) for blood collection among the preEQA Subprogram laboratories, instead of the closed system (vacuum) which presents a lower risk of hemolysis and is also internationally recommended (25). In parallel, as visual inspection for hemolysis detection involves a certain degree of subjectivity in the results of the preEQA Subprogram, as well as in those of the IFCC QI project, it should not be considered as the main cause of the difference observed. Finally, although 75% of the laboratories responding to the survey achieved - as is the case internationally - a quality level ranging between acceptable and desirable, the low-performing labeled laboratories were suggested to concentrate their efforts on the implementation of actions aimed at improving the blood collection procedure The percentage values collected for the TE-QI were also found to be higher than those reported by the IFCC QI project (8) and the difference among the percentage values was - in the particular case of this QI - also higher, as evidenced by the Sigma metrics. This is a plausible outcome on account of the fact that the IFCC QI model takes into account only transcription errors of the tests requested. Therefore, although it is not possible to make a straightforward comparison of the results collected, such a comparison could only be orientative. Furthermore, although 75% of the participating laboratories achieved a level of quality ranging between acceptable and desirable, it is advisable that the low-performing labeled laboratories take extreme measures to improve the data entry and their control procedure. The percentage values obtained for the four QIs evaluated allowed the preEQA Subprogram and the laboratories that measured them to compare their performance with each other and with the international literature available. This, in turn, helped them. not only to determine the level of error in their pre-analytical processes but also to predict the high or low need to implement continuous improvement actions in the processes, particularly in the case of the low-performing labeled laboratories. All in all, the Sigma metrics proved to be useful to achieve an objective interpretation of the percentage values collected after measuring QIs in the preEQA Subprogram and for each individual laboratory that performed QI measurements. The decision to determine that a Sigma value of 3.0 is equivalent to a minimum acceptable quality level should not overshadow the established goal of a desirable quality level equivalent to a Sigma value either higher than or equal to 4.0. In this sense, that a laboratory achieved a medium performance level according to the distribution of percentage values of the QI in the group (p25-p75) but with an error rate > 0.819% (Sigma < 4.0 and < 99.4% of yield) means that the process in question is objectively

vulnerable and does require the implementation of improvement actions (16,19). Likewise, a low performance level but with an error rate > 8.076% (Sigma < 3.0 and < 93.3% of yield) should be interpreted as unsatisfactory low performance. Sigma metrics turns out to be a fundamental tool for a correct interpretation of the results derived from QI measurement (19).

Conclusion

The majority (75%) of the laboratories participating in the preEQA Subprogram to measure the four QIs evaluated in this study showed a minimum acceptable level of quality (Sigma > 3.0). It was also observed that both the HS-QIs and the TE-QI deserve more attention. In spite of its weaknesses, the importance of this pioneering pre-analytical phase external quality assessment conducted in Argentina lies in that the participating laboratories - some for the first time - could identify processes in need of improvement thanks to the interlaboratory comparison of their performance. This preEQA Subprogram experience has allowed us to document issues deserving improvement in this Subprogram and has, in parallel, been motivating to keep on working on the measurement of pre-analytical QIs for the continuous improvement of the clinical laboratory TTP in Argentina.

Conflicts of interest

None

Acknowledgements

We are grateful to the Fundación Bioquímica Argentina for letting us create and coordinate the preEQA Subprogram and to Dr. Cecilia Etchegoyen, sub-director of the EQA Program, for her support in this work. Thanks are also due to the laboratories participating in this Subprogram to measure quality indicators.

References

1. Aita A, Sciacovelli L, Plebani M. Extra-analytical quality indicators – where to now? Clin Chem Lab Med. 2019 Dec 18 [cited 2023 Jun 21];57(1):127–133. Available from: https://doi.org/10.1515/cclm-2017-0964

2. Plebani M, Laposata M, Lundberg GD. The brain-to-brain loop concept for laboratory testing 40 years after its introduction. Am J Clin Pathol. 2011 Jan 12 [cited 2023 Jun 20];136:829–33. Available from: https://pubmed.ncbi.nlm.nih.gov/22095366/

3. Mrazek C, Lippi G, Keppel MH, Felder TK, Oberkofler H, Haschke-Becher E, et al. Errors within the total laboratory testing process, from test selection to medical decision-making – A review of causes, consequences, surveillance and solutions. Biochem Med (Zagreb). 2020 Jun 15 [cited 2023 Jun 20];30(2):020502. Available from: https://doi.org/10.11613/BM.2020.020502

4. Caballero A, Gómez-Rioja R, Ventura M, Llopis MA, Bauça JM, Gómez-Gómez C, et al. Evaluation of 18 quality indicators from the external quality assurance preanalytical programme of the Spanish Society of Laboratory Medicine (SEQCML). Adv Lab Med. 2022 Jun 8 [cited 2023 Jun 20];3(2):188-200. Available from: https://doi.org/10.1515/almed-2021-0097

5. Marzana Sanz I, Ibarz Escuer M, Llopis Diaz MA, Barba Meseguer N Alsina Kirchner MJ, Martínez Espartosa D, et al. Recommendations for the design and implementation of a Pre-analytical Phase Quality Assurance Program [Recomendaciones para el diseño e implementación de un programa de aseguramiento de la calidad de la fase preanalítica]. Rev Lab Clin. 2019 Oct-Dec [cited 2023 Jun 20];12(4):e54-e65. Available from: https://doi.org/10.1016/j.labcli.2019.01.003

6. The International Organization for Standardization. ISO 15189:2022. Medical laboratories – requirements for quality and competence. Switzerland: International Organization for Standardization; 2012.

7. Fundación Bioquímica Argentina. LAM3:2012. Laboratory Accreditation Manual [MA3:2012. Manual de Acreditación de Laboratorios]. Argentina: Fundación Bioquímica Argentina; 2012 [cited 2023 Jun 20]. Available from: https://www.fba.org.ar/wp-content/uploads/2020/11/MA3-FBA.pdf

8. Sciacovelli L, Padoan A, Aita A, Basso D, Plebani M. Quality indicators in laboratory medicine: state-of-the-art, quality specifications and future strategies. Clinical Chemistry and Laboratory Medicine (CCLM). 2023 [cited 2023 Jul 11];497:35–40. Available from: https://doi.org/10.1515/cclm-2022-1143

9. Kristensen G, Aakre KM, Kristoffersen AH, Sandberg S. How to conduct external quality assessment schemes for the preanalytical phase? Biochem Med. 2014 Feb 15 [cited 2023 Jun 21];24:114–22. Available from: https://www.biochemia-medica.com/en/journal/24/1/10.11613/BM.2014.013

10. International Federation of Clinical Chemistry and Laboratory Medicine. Quality Indicators in Laboratory Medicine. 2017 [cited 2023 Jun 21]. Available from: https://ifcc.org/ifcc-education-division/working-groups-special-projects/wg-leps/quality-indicators-project/

11. Fundación Bioquímica Argentina. External Quality Assessment Program (EQA Program) [Programa de Evaluación Externa de la Calidad (PEEC)]. Pre-Analytical Phase External Quality Assessment Subprogram (preEQA Subprogram) [Subprograma Control de Calidad en la Etapa Preanalítica (PREAL)]. 2016 [cited 2023 Jun 21]. Available from: https://www.fba.org.ar/programas-de-la-fba/peec/subprogramas/

12. Shcolnik W, Berlitz F, Galoro CAO, Biasoli V, Lopes R, Jerônimo D, et al. Brazilian laboratory indicators benchmarking program: three-year experience on pre-analytical quality indicators. Diagnosis. 2020 Aug 31 [cited 2023 Jun 21];8(2):257-268. Available from: https://doi.org/10.1515/dx-2020-0043

13. New South Wales Health system. Clinical Excellence Commission Academy. Safety and quality improvement tools. Pareto Charts & 80-20 Rule. 1968 [cited 2023 Jun 21]. Available from:https://www.cec.health.nsw.gov.au/CEC-Academy/quality-improvement-tools/pareto-charts

14. Badrick T, Barden H, Callen S, Dimeski G, Gay S, Grahamet P, et al. Consensus Statement for the Management and Reporting of Haemolysed Specimens. Clin Biochem Rev. 2016 Dec [cited 2023 Jun 21];37(4):140-142. Available from: https://pubmed.ncbi.nlm.nih.gov/28167843/

15. Westgard OJ. Westgards QC. Six Sigma Calculators. DPM (Defects Per Million) Calculator. 2019 [cited 2023 Jun 21]. Available from: http://www.westgard.com/six-sigma-calculators.htm

16. Zorbozan N, Zorbozan O. Evaluation of preanalytical and postanalytical phases in clinical biochemistry laboratory according to IFCC laboratory errors and patient safety specifications. Biochem Med (Zagreb). 2022 Aug 5 [cited 2023 Jun 21];32(3):030701. Available from: https://doi.org/10.11613/BM.2022.030701

17. Plebani M, Task E, Speci P. Performance specifications for the extra-analytical phases of laboratory testing: why and how. Clin Biochem. 2017 Jul [cited 2023 Jun 21];50:550–4. Available from: https://doi.org/10.1016/j.clinbiochem.2017.02.002

18. Llopis MA, Trujillo G, Llovet MI, Tarrés E, Ibarz M, Biosca C, et al. Quality indicators and specifications for key analytical extranalytical processes in the clinical laboratory. Five years' experience using the Six Sigma concept. Clin Chem Lab Med. 2011 Jan 31 [cited 2023 Jun 21];49(3):463-70. Available from: https://doi.org/10.1515/cclm.2011.067

19. Carchio SM, Cappella AC, Goedelmann C, Pandolfo M, Bustos D. Application of Six Sigma in the Clinical Laboratory [Aplicación de Seis Sigma en el Laboratorio Clínico]. Acta Bioquím Clín Latinoam. 2019 [cited 2023 Jun 22];53(4):525-37. Available from: http://www.scielo.org.ar/pdf/abcl/v53n4/v53n4a13.pdf

20. Ministerio de Salud de la Nación Argentina. Guidelines for Human Health Research [Guía para investigaciones en salud humana]. 2011 Sep 21 [cited 2023 Jun 21]. Available from: https://www.argentina.gob.ar/normativa/nacionalresoluci%C3%B3n-1480-2011-187206/actualizacion

21. Inserra F, Torres ML, Alles A, Bonelli C, Ceci R, Corradino C, et al. Evaluation of the renal function for the detection and monitoring of chronic kidney disease. Multidisciplinary consensus document 2021. Special attention to specific clinical situations. [Evaluación de la función renal para la detección y seguimiento de la enfermedad renal crónica. Documento multidisciplinario de consenso 2021. Especial atención sobre situaciones clínicas específicas]. Acta Bioquím Clín Latinoam. 2022 [cited 2023 Aug 19];56(1):43-74. Available from: http://www.abcl.org.ar/edicionesanteriores.html

22. Sua LF, Amezquita MA, Hernández DE, Alcalá-Flores M, Leib CS, Aguirre Rojas M, et al. Automated preanalytical units in the hemostasis laboratory: an observational descriptive prospective study developed in a university hospital between April 15 to July 15, 2017. Iatreia. 2019 Jul-Sep [cited 2023 Jun 22];32(3):177-183. Available from: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-07932019000300177

23. Kristoffersen AH, Stavelin AV, Ajzner E, Kristensen GBB, Sandberg S, Van Blerk M, et al. Pre-analytical practices for routine coagulation tests in European laboratories. A collaborative study from the European Organisation for External Quality Assurance Providers in Laboratory Medicine (EQALM). Clin Chem Lab Med. 2019 Sep 25 [cited 2023 Jun 22];57(10):1511-1521. Available from: https://doi.org/10.1515/cclm-2019-0214

24. Fang, Kui, Dong, Zheqing, Chen, Xiling, Zhu, Ji, Zhang, Bing, You, Jinbiao, Xiao, Yingjun and Xia, Wenjin. Using machine learning to identify clotted specimens in coagulation testing. Clin Chem Lab Med. 2021 Mar 3 [cited 2023 Jun 22];59(7):1289-1297. Available from: https://doi.org/10.1515/cclm-2021-0081

25. Clinical and Laboratory Standards Institute. Collection of Diagnostic Venous Blood Specimens. 7th ed. CLSI standard GP41. Wayne, PA: Clinical and Laboratory Standards Institute; 2017.

Disruption of laboratory activities during the COVID-19 pandemic: results of an EFLM Task Force Preparation of Labs for Emergencies (TF-PLE) survey

Giuseppe Lippi[1*], Janne Cadamuro[2*], Elisa Danese[1*], Emmanuel J. Favaloro[3*], Julien Favresse[4*], Brandon M. Henry[5*], Snezana Jovicic[6*], Tomris Ozben[7*], Mario Plebani[8*], Jecko Thachil[9*]

[1]*Section of Clinical Biochemistry and School of Medicine, University of Verona, Verona, Italy*
[2]*Department of Laboratory Medicine, University Hospital Salzburg, Paracelsus Medical University, Salzburg, Austria*
[3]*Haematology, NSW Health Pathology, Sydney Centres for Thrombosis and Haemostasis, ICPMR, Westmead Hospital, Sydney, Australia*
[4]*Services of Clinical Biology, Clinic Saint-Luc, Bouge, Belgium*
[5]*Cardiac Intensive Care Unit, The Heart Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, USA*
[6]*Department for Medical Biochemistry, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia*
[7]*Department of Medical Biochemistry, Faculty of Medicine, Akdeniz University, Antalya, Turkey*
[8]*University of Padova, Padova, Italy*
[9]*Immune Thrombocytopenic Purpura (ITP) Clinic, Haematology Department, Manchester University NHS Foundation Trust - Manchester Royal Infirmary, Manchester, UK*

Article Info

Author of correspondence:
Prof. Giuseppe Lippi, MD
Section of Clinical Biochemistry;
E-mail: giuseppe.lippi@univr.it;
Tel.: +39-045-8124308;
Fax: +39-045-8122970;
Address:
University Hospital of Verona
Piazzale LA Scuro, 37134 Verona, Italy

Keywords

COVID-19; SARS-CoV-2; Laboratory Medicine; Survey

Abstract

Background: The EFLM Task Force Preparation of Labs for Emergencies (TF-PLE) created a survey that has been distributed to its members for gathering information on the key hazards experienced by European medical laboratories during the COVID-19 pandemic.

Methods: The survey was distributed to over 12,000 potential contacts (laboratory workers) via an EFLM newsletter, with responses collected between May 8 and June 8, 2023.

Results: Two hundred replies were collected and examined from European laboratories. 69.7% and 78.1% of all responders said they were short on non-COVID and COVID reagents, respectively. Exactly half of respondents (50.0%) said that they could not complete all laboratory tests required for a specific period, but this figure climbed to 61.2% for COVID tests. Finally, 72.3% of respondents expressed exhaustion during the pandemic, and 61.2% reported increasing patient hostility.

Conclusions: The COVID-19 pandemic had a significant impact on laboratory medicine in Europe. Cultural change, proactive planning, and even re-engineering in some parts of the laboratory industry may thus be necessary to prepare for future challenges.

Introduction

More than three years after the World Health Organization (WHO) declared coronavirus disease 2019 (COVID-19) a pandemic, the international emergency status has ended, but the organization continues to emphasize that COVID-19 remains a global health threat [1]. The fact that most health systems around the world were not prepared for this enormous challenge can certainly be considered one of the most important aspects that have contributed to increasing the morbidity, mortality, and resulting chronic disability caused by SARS-CoV-2 infections [2]. This refers specifically to the enormous burden placed on healthcare facilities by the pandemic, which exponentially increased shortages of beds, staff, and equipment [3], dramatically exacerbated by previous inadequacies in hospital funding. An interesting analysis conducted by Arsenault et al. in 10 different countries [4] shows that significant disruption occurred in almost all countries, characterized by a specific magnitude and duration, with no pattern related to income or pandemic burden. For example, treatment of chronic diseases was disrupted in all regions, while treatment of emergencies such as road traffic accidents was severely affected. The clear evidence that most health systems were woefully unprepared for the dramatic increase in the number of patients seeking diagnosis and treatment for SARS-CoV-2 infection during the initial phase of the pandemic goes hand in hand with evidence that even laboratory medicine was placed under unprecedented and perhaps unimaginable pressure [5,6]. A previous survey was conducted by the American Association for Clinical Chemistry (AACC) in four different periods during the early phase of the COVID-19 pandemic (May 1-24 2020: 100 responses; June 1-5 2020: 33 responses; June 24-July 6 2020: 53 responses; August 3-18 2020: 67 responses; September 17-29 2020: number of responses unavailable; and December 2 2020-January 4 2021: number of responses unavailable) [7]. The percentage of worldwide respondent labs reporting being unable to obtain supplies necessary to run routine laboratory testing ranged between 11-52% for non-COVID-19 tests and between 40-50% for COVID-19 tests, respectively. Contextually, the percentage of laboratories unable to process all requested COVID-19 tests due to supply issues and other challenges ranged between 14-22%. Importantly, up to nearly 80% laboratories responded that they were facing challenges to testing or increasing their testing capacity for COVID-19. Shortage of test kits and reagents affected as many as 60% of all respondent laboratories, whilst staff shortage was also commonplace, involving up to 80-90% of all respondent laboratories. Importantly, during the last surveyed period (December 2, 2020-January 4, 2021), some degree of burnout has also been reported by as many as 70% of all respondents. Nuñez-Argote reviewed 178 surveys completed during the early period of the pandemic by laboratory professionals, engaged in medical laboratories in the United States, and found evidence that the extent of overtime work increased almost every day from 3.4% to 13.5% from before to during the pandemic [8]. Jafri et al. interviewed 64 medical laboratory professionals in Pakistan between June 4th and 14th 2020, and evidenced several important aspects, including the fact that 42% and 78% of the respondents reported fear of employment termination and financial challenges, 96% answered that social life was strongly penalized and nearly 20% that they were largely unsatisfied about the measures taken by the hospital organization during the initial outbreak [9]. Another survey was conducted by the PeriAnalytic and Laboratory Medicine Society (PALMSoc) in Ireland, collecting 45 responses from 38 different medical laboratories [10]. Ac-cording to the results of this survey, nearly 60% of responders affirmed that maintenance of the quality management system was challenged and less than 20% reported as having 100% staffing level before the emergence of COVID-19. To determine the frequency of burnout and depression, along with their contributing factors and the impact of COVID-19, an electronic survey was distributed to a group of Canadian laboratory medicine residents [11]. The authors ultimately collected 79 responses, which revealed a prevalence of burnout and depression of 63% and 47%, respectively. The factors that contributed most to burnout were dissatisfaction with career, impairment of academic performance, lack of availability of sick leave, financial stress, and increased perception of fatigue. Regarding depression, the most important factors were poor availability of wellness resources, reduced free time, and experimentation with reduced sleep duration. In this challenging scenario, with a still unpredictable evolution regarding the risk of future natural and environmental disasters [12], the Task Force Preparation of Labs for Emergencies (TF-PLE) of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) has developed a specific survey to be sent to its members. The aim was to collect useful data and important information on the major threats faced by European medical laboratories during the COVID-19 pandemic for developing suggestions on how to avoid the next emergency (besides pandemics) with potentially similar unfavorable consequences.

Materials and methods

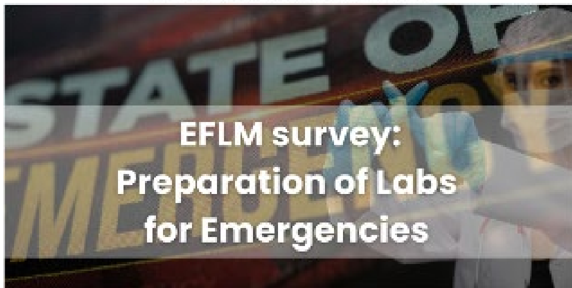
To obtain specific information about the extent and nature of the disruption of laboratory activities during the COVID-19 pandemic, a special questionnaire was developed by the EFLM TF-PLE that included general questions about the location and organization of laboratories, as well as a specific request for information about the most important organizational problem encountered, as follows: "Which types of troubles did you encounter (at least at some times) during the COVID-19 pandemic?", including six possible scenarios with four different answer options each "Yes, for long time"; "Yes, for short time"; "Never"; "N/A (not applicable)". The detailed questions included in the survey are summarized in table 1. The survey was then sent via an EFLM newsletter to the email addresses of over 12,000 potential contacts (laboratory professionals) from Europe and abroad, with responses collected between May 8 and June 8, 2023 (Figure 1). It had been requested that only one response can be accepted from each laboratory. The complete responses

were downloaded onto an Excel spreadsheet and graphically analyzed with Analyse-it for Microsoft Excel (Analyse-it Software Ltd, Leeds, UK). A single response for each medical laboratory was maintained. The statistics of the responses to the question about the threats encountered by the respondent laboratories during the pandemic was conducted by eliminating the results from participants who selected the "N/A (not applicable)" option (n=?). The survey was officially promoted and supported by the EFLM, and did not involve any medical treatment. Ethics Committee approval or patient's consent is not applicable to these types of studies.

Table 1: Questions and options presented in the questionnaire developed by the EFLM TF-PLE

Question	Formula %
Your continent	Asia; Africa; North America; South America; Europe; Oceania
Your country	Free text
Your lab (type of facility)	Private, Public; Other
Number of tests per year of your lab (both inpatients and outpatients)	1 million; 1-4 million; 5-8 million; >8 million
Do you perform stat (urgent) testing in your lab?	Yes; No
Which types of troubles did you encounter (at least at some times) during the COVID-19 pandemic? - Impossibility to run all lab tests requested - Impossibility to run all COVID tests requested - Shortage of some non-COVID reagents and/or supplies - Shortage of COVID reagents and/or supplies - Burnout - Patient aggressiveness	Yes, for long time; Yes, for short time; Never; N/A (not applicable)

The EFLM Task Force "Preparation of Labs for Emergencies", chaired by Prof. Giuseppe Lippi, asks for your attention on...



Dear Friends and Colleagues,

Please find below the link to an official survey developed by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Task Force: Preparation of Labs for Emergencies (TF-PLE).

The survey is aimed to collect data from Europe and abroad concerning the major environmental and biological threats and the natural disasters

We kindly invite you to fill this questionnaire (it will only take around 5 min of your time), since this represents a necessary premise for acknowledging lab threats and the state-of-the-art for their optimal management, providing insights as to whether the EFLM shall be engaged in providing universal guidance on this matter.

The survey will remain open **between May 8 and May 22, 2023.**

Best regards

Giuseppe Lippi
Chair
EFLM Task Force: Preparation of Labs for Emergencies

Click here to access the survey

Figure 1: The EFLM Newsletter about the TF-PLE survey that had been delivered to over 12,000 potential contacts among European and non-European laboratory professionals.

Results

During the one-month survey period, a total of 235 responses were received, 200 were from European laboratories (85.1%), which were used for the analysis. Most responses came from Italian laboratories (20.0%), followed by Serbian (9.0%), Turkish (6.5%), Spanish (5%), Croatian, Romanian, and Lithuanian (all 4.0%) institutions; laboratories from other countries accounted for less than 3% of all other responses. Most respondents were from public laboratories (78%), general (rather than specialized) laboratories (74.5%), of which 88.5% also perform urgent and/or emergency laboratory testing. In terms of size, most laboratories reported performing less than 1 million tests per year (42.0%), 32.0% reported performing between 1-4 million tests per year, 10.5% of all laboratories reported performing between 5-8 million tests per year, whilst 15.5% of responding laboratories reported performing more than 8 million tests per year (Figure 2). The responses to the specific question "which types of troubles did you encounter (at least at some times) during the COVID-19 pandemic?" are summarized in figure 3.

The percentage of "N/A (not applicable)" responses was 6.0% for "shortage of some non-COVID reagents and/or supplies", 15.5% for "shortage of COVID reagents and/or supplies", 5.0% for "impossibility to run all lab tests requested", 17.5% for "impossibility to run all COVID tests requested", 8.0% for "burnout" and 17.5% for "patient aggressiveness". Regarding reagents and/or supplies, other than for COVID diagnostics, 69.7% of all laboratories answered that they had some shortage, mostly (63.3%) for a short time, while they answered that some shortage of COVID reagents was more frequent (78.1%), but also mostly temporary (72.8%). Regarding the ability to perform tests, exactly half (50.0%) of all laboratories were unable to complete all tests requested over a certain period (mostly for a short-term, 44.2%), but this number increased to 61.2% for COVID tests ordered (57.0% on the short-term). Importantly, 72.3% of respondents emphasized that they had experienced some level of burnout during the COVID pandemic, 31.5% of them for an extended period. Finally, 61.2% of respondents reported that patient aggressiveness had increased during the COVID pandemic, 22.4% of whom had experienced this phenomenon over a longer period. A partial analysis of the data, stratifying all positive responses to the first four questions (i.e., the sum of "Yes, for long time" and "Yes, for short time" responses) by the amount of testing performed by laboratories, is shown in Figure 4. There are no major differences in the number of tests performed per year. In comparison to all other laboratory size categories, laboratories performing between 5-8 million tests experienced significantly more difficulties "to run all COVID tests requested" (chi-square statistic: 15.497; p=0.001). Apart thereof, no statistical significant difference between answers from different laboratory sizes could be found for "shortage of some non-COVID rea-gents and/or supplies" (chi-square statistic: 1.66; p=0.645), "shortage of COVID reagents and/or supplies " (chi-square statistic: 7.125; p=0.068) or "impossibility to run all lab tests requested" (chi-square statistic: 1.445; p=0.695).

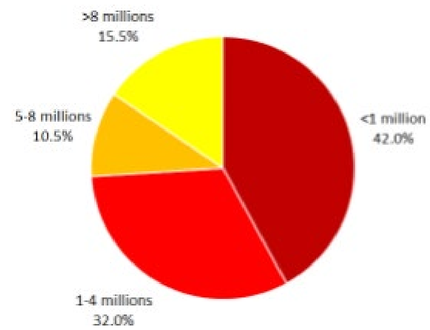


Figure 2: The percentages of the 200 EFLM European medical laboratories which responded to the EFLM survey stratified according to the volume of tests performed per year.

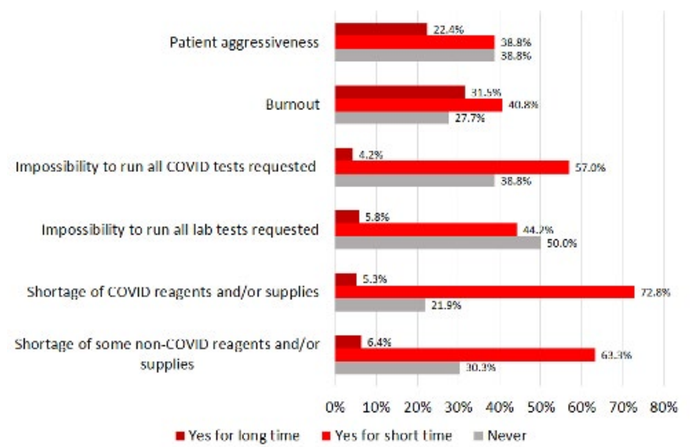


Figure 3: Responses to the question "Which types of troubles did you encounter (at least at some times) during the COVID-19 pandemic?" given by 200 European medical laboratories which responded to the EFLM survey.

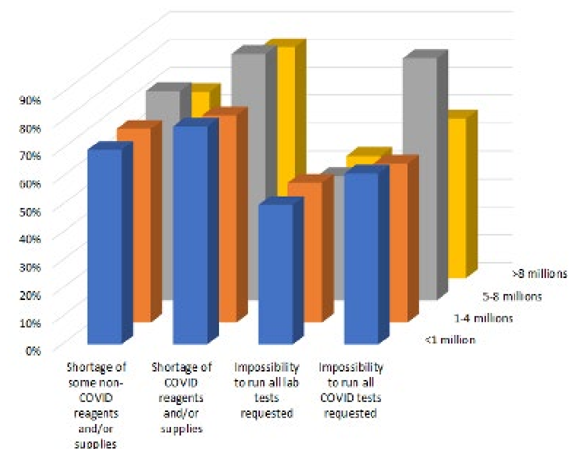


Figure 4: Positive responses to the question "which types of troubles did you encounter (at least at some times) during the COVID-19 pandemic?" given by 200 European medical laboratories which responded to the EFLM survey, stratified according to their testing volume (test per year).

Discussion

Although the Emergency Committee of the WHO made the predictable decision to end the global health emergency for COVID-19, many thousands of people affected by this disease are still in hospitals (some of whom require intensive care unit treatment), and several million others will suffer from the lingering after-effects of SARS-CoV-2 infection [13]. In addition, this (corona) virus is here to stay, and there is a risk that new and more severe lineages will emerge, causing new outbreaks with surging case numbers. The frequency of other natural and environmental disasters posing a public health challenge worldwide has increased significantly over the past few decades [14]. These may also act synergistically to amplify the harm to humans and animals [15]. These threats mainly include tornadoes, thunderstorms, hail, earthquakes and tsunamis, fires, floods, chemical and/or biological emergencies, mass casualties, terrorism and bioterrorism, wars, civil unrest, and so on. Recognizing that laboratory medicine plays a critical role in modern science and medicine [16] and that its contribution is indispensable for the management of frequently foreseeable emergencies (e.g., climate change), the EFLM recently established an ad hoc task force to improve the preparedness of medical laboratories to manage a variety of emergencies (EFLM TF-PLE). The first initiative, which provides an essential basis for planning future training activities, was the development and implementation of a special survey, aimed at collecting information on the extent of disruption of laboratory activities during the last three years of this COVID-19 pandemic. The results of which are presented and discussed in this article. In keeping with the evidence emerged from the previous AACC questionnaire, which was terminated at the beginning of 2021, several critical aspects could be identified from this EFLM initiative. From the responses obtained from over 200 European medical laboratories, the first fact that strongly emerges is that the capacity of both conventional and COVID-19 related diagnostics has been overwhelmed in over 50% of cases for at least some periods. This aspect underlines that around half of medical laboratories which responded to this EFLM survey were already running at their capacity limits, and this precarious stability was disrupted by an "exceptional" event like a pandemic. This inherently means that other similar (natural or environmental) disasters might generate a similar dramatic impact on laboratory medicine, causing important delays or even prolonged interruptions of the diagnostic activity, together with all ensuing patient safety risks. A second finding from our survey is that most European laboratories have suffered a temporary lack of reagents and supplies, which was not restricted to COVID-19 diagnostics but involved also many other testing areas. Intriguingly, nearly two-thirds of all European medical laboratories that responded to this EFLM survey stated that a lack of reagents or supplies for performing non-COVID tests was a tangible issue during the pandemic, thus emphasizing that the entire diagnostic industry was seemingly unprepared to face an exceptional event like this pandemic, and may remain so also in the unfortunate likelihood of future disasters. The third significant conclusion from the 200 replies collected in this EFLM survey is that more than 70% of European laboratory workers admitted to having experienced at least some degree of burnout during the pandemic. This statistic is consistent with prior research, which found that burnout was frequent among healthcare workers, especially in the early phases of the pandemic [17]. Macaron et al. conducted a systematic literature review and meta-analysis to determine the cumulative prevalence of burnout among physicians during the COVID-19 pandemic [18], reporting a peak burnout prevalence of up to 60% in the early stages of the pandemic, which is comparable to the prevalence found in our survey of European laboratory professionals (i.e., around 70%). Last but not least, almost 60% of respondents indicated that patient hostility increased during the pandemic, thus contributing to further aggravation of an already difficult working condition caused by environmental pressure and shortage of personnel. We acknowledge that only a minority of all potentially contacted laboratories responded to this survey, with a preponderance of Italians. This imbalance must be taken into account when interpreting our conclusions.

Conclusions

We live in a modern world where other major challenges such as the current COVID-19 pandemic are very likely to occur in the coming years. If we have learned anything from the recent pandemic, it is that proactivity and preparedness to respond in a much more expedient manner are critical. The findings of this first EFLM TF-PLE survey clearly reveal that the COVID-19 pandemic had a significant impact on laboratory medicine in Europe, both in terms of availability of material resources and professional well-being. Cultural change, proactive planning, and even reengineering in some parts of the laboratory industry may thus be required to prepare for future challenges.

Funding

This research received no external funding.

Acknowledgements

None.

Conflicts of Interest

The authors declare no conflict of interest

References

1. Wise J. Covid-19: WHO declares end of global health emergency. BMJ. 2023 May 9;381:1041. doi: 10.1136/bmj.p1041. PMID: 37160309.

2. Mattiuzzi C, Lippi G. Which lessons shall we learn from the 2019 novel coronavirus outbreak? Ann Transl Med. 2020 Feb;8(3):48. doi: 10.21037/atm.2020.02.06. PMID: 32154288; PMCID: PMC7036635.

3. Chehrehgosha M. The Unpreparedness of the Healthcare System for the Management of COVID-19 Pandemic Leading to the Mistreatment of the Elderly: A Newly Emerging Moral Dilemma. J Nutr Health Aging. 2020 Jun 6:1–2. doi: 10.1007/s12603-020-1410-8. Epub ahead of print. PMCID: PMC7275658.

4. Arsenault C, Gage A, Kim MK, Kapoor NR, Akweongo P, Amponsah F, Aryal A, Asai D, Awoon-or-Williams JK, Ayele W, Bedregal P, Doubova SV, Dulal M, Gadeka DD, Gordon-Strachan G, Mariam DH, Hensman D, Joseph JP, Kaewkamjornchai P, Eshetu MK, Gelaw SK, Kubota S, Leerapan B, Margoz-zini P, Mebratie AD, Mehata S, Moshabela M, Mthethwa L, Nega A, Oh J, Park S, Passi-Solar Á, Pé-rez-Cuevas R, Phengsavanh A, Reddy T, Rittiphairoj T, Sapag JC, Thermidor R, Tlou B, Valenzuela Guiñez F, Bauhoff S, Kruk ME. COVID-19 and resilience of healthcare systems in ten countries. Nat Med. 2022 Jun;28(6):1314-1324. doi: 10.1038/s41591-022-01750-1. Epub 2022 Mar 14. PMID: 35288697; PMCID: PMC9205770.

5. Plebani M. Laboratory medicine in the COVID-19 era: six lessons for the future. Clin Chem Lab Med. 2021 Apr 7. doi: 10.1515/cclm-2021-0367. Epub ahead of print. PMID: 33826810.

6. Tang JW, Dilcher M, Bird PW, Kok J, Lee CK, Nishimura H, Oon L, Alston G, Dwyer DE, Holmes CW, Jennings LC. Practical problems and responses for SARS-CoV-2 laboratory testing during the COVID-19 pandemic. Clin Microbiol Infect. 2023 May;29(5):560-562. doi: 10.1016/j.cmi.2023.02.005. Epub 2023 Feb 10. PMID: 36773772; PMCID: PMC9911976

7. AmericanAssociation for Clinical Chemistry. Coronavirus Testing Survey. Available at: https://www.aacc.org/science-and-research/covid-19-resources/aacc-covid-19-testing-survey. Last ac-cessed, Jun 10, 2023.

8. Nuñez-Argote L, Baker DP, Jones AP. Initial Clinical Laboratory Response to COVID-19: A Survey of Medical Laboratory Professionals. Lab Med. 2021 Jul 1;52(4):e115-e124. doi: 10.1093/labmed/lmab021. PMID: 33942859; PMCID: PMC8135790.

9. Jafri L, Ahmed S, Siddiqui I. Impact of COVID-19 on laboratory professionals-A descriptive cross sectional survey at a clinical chemistry laboratory in a developing country. Ann Med Surg 2020:57:70-7.

10. Leonard A, Murray B, Prior AR, Srinivasan R, Kane A, Boran G. Survey of laboratory medicine's national response to the Covid-19 pandemic in the Republic of Ireland. Ir J Med Sci. 2022 Feb;191(1):65-69. doi: 10.1007/s11845-021-02578-3. Epub 2021 Mar 4. PMID: 33665780; PMCID: PMC7932685.

11. Han R, Hahn E, Done SJ, Pun C, Shivji S, Lu FI. Resident Depression and Burnout During the COVID-19 Pandemic: A Survey of Canadian Laboratory Medicine Trainees. Arch Pathol Lab Med. 2023 Mar 1;147(3):368-375. doi: 10.5858/arpa.2021-0501-EP. PMID: 35802936.

12. Lippi G, Favaloro EJ, Plebani M. Laboratory medicine and natural disasters: are we ready for the challenge? Clin Chem Lab Med. 2010 May;48(5):573-5. doi: 10.1515/CCLM.2010.148. PMID: 20450330.

13. Kupferschmidt K, Wadman M. End of COVID-19 emergencies sparks debate. Science. 2023 May 12;380(6645):566-567. doi: 10.1126/science.adi6511. Epub 2023 May 11. PMID: 37167400.

14. Hidalgo J, Baez AA. Natural Disasters. Crit Care Clin. 2019 Oct;35(4):591-607. doi: 10.1016/j.ccc.2019.05.001. Epub 2019 Jul 15. PMID: 31445607.

15. Drake JM, Marty É, Gandhi KJK, Welch-Devine M, Bledsoe B, Shepherd M, Seymour L, Fortuin CC, Montes C. Disasters collide at the intersection of extreme weather and infectious diseases. Ecol Lett. 2023 Apr;26(4):485-489. doi: 10.1111/ele.14188. Epub 2023 Feb 27. PMID: 36849208.

16. Lippi G, Plebani M. A modern and pragmatic definition of Laboratory Medicine. Clin Chem Lab Med. 2020 Jul 28;58(8):1171. doi: 10.1515/cclm-2020-0114. PMID: 32069231.

17. Ulfa M, Azuma M, Steiner A. Burnout status of healthcare workers in the world during the peak period of the COVID-19 pandemic. Front Psychol. 2022 Sep 21;13:952783. doi: 10.3389/fpsyg.2022.952783. PMID: 36211838; PMCID: PMC9532965.

18. Macaron MM, Segun-Omosehin OA, Matar RH, Beran A, Nakanishi H, Than CA, Abulseoud OA. A systematic review and meta analysis on burnout in physicians during the COVID-19 pandemic: A hidden healthcare crisis. Front Psychiatry. 2023 Jan 12;13:1071397. doi: 10.3389/fpsyt.2022.1071397. PMID: 36713915; PMCID: PMC9877514.

Inter-laboratory method validation of CD34+ flow-cytometry assay: the experience of Turin Metropolitan Transplant Centre

Ivana Ferrero I.F.[1*], Deborah Rustichelli D.R.[1*], Sara Castiglia S.C.[1*], Loretta Gammaitoni L.G.[2*], Alessandra Polo A.P.[2*], Marisa Pautasso M.P.[3*], Massimo Geuna M.G.[4*], Franca Fagioli F.F.[1],[5*]

[1] *Stem Cell Transplantation Laboratory, Regina Margherita Children's Hospital, City of Health and Science of Turin, Italy; ivana.ferrero@unito.it, drustichelli@cittadellasalute.to.it, scastiglia@cittadellasalute.to.it*

[2] *Candiolo Cancer Institute, Fondazione del Piemonte per l'Oncologia (FPO) - IRCCS, Candiolo, Italy; loretta.gammaitoni@ircc.it, apolo@cittadellasalute.to.it*

[3] *Clinical Pathology Unit, University Hospital San Luigi Gonzaga, Orbassano, Italy; m61.pautasso@gmail.com*

[4] *Immuno-Pathology Unit, Mauriziano Umberto I Hospital, Italy; massimo.geuna@gmail.com*

[5] *Department of Science of Public Health and Paediatrics, University of Turin, Piazza Polonia 94, 10126 Torino, Italy; franca.fagioli@unito.it*

Article Info

Author of correspondence:
Ivana Ferrero
E-mail: ivana.ferrero@unito.it;
Tel.: +39-011-3135420;
Address:
Stem Cell Transplantation Laboratory, Regina Margherita Children's Hospital, City of Health and Science of Turin, Italy

Keywords

Validation, risk analysis, CD34+

Abstract

The Turin Metropolitan Transplant Centre (CIC 305) includes four flow-cytometry laboratories assessing quality control on hematopoietic stem cells (HSC) with different instruments and operators. Therefore, the CD34+ enumeration assay should be validated on a regular basis. We describe here the validation plan to test the inter-laboratory reproducibility of CD34+ enumeration assay, based on the risk analysis. Stabilized blood samples were analysed using Stem-Kit reagent according to manufacturer's instructions and acquired using the Beckman Coulter Navios at Regina Margherita Children's' Hospital (305-1), Beckman Coulter FC500 at Candiolo Cancer Institute FPO-IRCCS (305-2), BD Biosciences FACSLyric™ at S. Luigi Hospital (305-3), and Beckman Coulter Navios EX at Mauriziano Hospital (305-4). The ISHAGE guidelines were followed for estimating % and absolute number of CD34+ cells in single-platform method.

For each sample repeatability limit (r), reproducibility error, uncertainty of reproducibility error and coefficient of variation (CV) were reported. The repeated measurements from each laboratory or instrument have a variability, expressed as reproducibility error, lower than the repeatability limit for that single parameter. The corrected reproducibility error is always lower than the repeatability limit except for the percentage value of the "low" count. The analysis of inter-laboratory variance is within the maximum acceptable variance value, and the CV of all measurements for each parameter is less than 8%, indicating low measurement variability among laboratories. Evaluating the overall data, we can conclude that the four laboratories are perfectly aligned and the results are reproducible.

Introduction

According to Joint Accreditation Committee of ISCT and EBMT (JACIE) standards, relevant and standardized assay for quantifying hematopoietic stem cell (HSC) population needs to be established and periodically validated to keep the entire process under control. The efficiency of an autologous or allogeneic HSC graft is mainly determined by the number of CD34+ cells, and their measurement by flow cytometry is an important method to assess the graft quantity (1). When HSCs are used, manipulated or not, enumeration and viability are critical information that need to be considered. Therefore, a relevant and standardized assay for quantifying the initial and final cell population needs to be established and validated (2). For many years, multi-colour flow cytometry is the technology for cell surface marker detection, viability and enumeration (3). Measurement of viable absolute counts of cells can be performed using single-platform panels recommended by International Society for Hematotherapy and Graft Engineering (ISHAGE) and JACIE standards. This includes a cell viability dye and counting beads with a lyse/no-wash preparation using commercial kits for CD34+ cells and lymphocyte subpopulations, yet without a viability dye for the latter (4-5). The challenging aspect is validating laboratory-developed flow cytometry methods, as a successful validation makes the method readily available and adapted by quality control for laboratory processing, cryopreserving or manipulating products for cell production. (2). The Turin Metropolitan Transplant Centre (CIC 305) (TMTC) is a functional program established in 2012 from the partnership of four local programs, one paediatric and three adults, on the basis of clinical and laboratory collaborations. It promotes the exchange of scientific knowledge and professional experience for clinical and research activities, implementing shared quality policies for maintaining the performance of critical activities related to HSC transplantation (HSCT). The program shares policies and aims, technical procedures and discusses on a daily basis the key performance indicators. In the TMTC, there are four flow-cytometry laboratories assessing quality control on HSCs, with different instruments and operators. For backup and emergency, a continuity plan is in place to guarantee the prompt intervention and maintenance of many services, including CD34+ evaluation. The risk that an inaccurate CD34+ enumeration may affect the efficacy of the cell-based product administered to patients need to be taken into consideration. Therefore, the CD34+ evaluation assay must be aligned among the labs, standardized and validated on a regular basis. Any change of equipment, utilities, or process should be formally documented and the impact on the validation status or control strategy assessed. A process is validated by establishing objective evidence that it consistently produces an expected endpoint or result that meets predetermined acceptance criteria. A risk assessment should be performed for each validation study to assess how critical the process is and to define the level of risk (6). The risk assessment represents a basic step to go through in the validation process. There are several methods for the assessment of the risk, such as Failure Mode and Effects Analysis (FMEA) or Failure Modes, Effects and Criticality Analysis (FMECA), methods of assessing potential failure mechanisms and their impact in the system, identifying single failure points. FMEA method can be used to assess the risk of failures and identify criticisms at every step of the process, and is widely adopted in various health care settings, including transfusion medicine (7) and HSCT (8). In this work, we describe the validation plan with relevance to the risk analysis to test the inter-laboratory reproducibility of the CD34+ enumeration method.

Materials and methods

Risk analysis

All the steps from transport of samples to the final report were carefully examined adopting a failure mode and effect analysis scheme. For the FMEA analysis, the whole process was divided into several sub-phases, and for each step, one or more different failure modes were identified. For each failure mode, a risk priority number (RPN) was calculated considering three different scores expressing severity of effects, frequency of occurrence and identifying the detection, on a scale from 1 to 4, as described in the tables 1 and 2.

Severity was multiplied for occurrence to obtain the failure risk (range 1-16). The failure risk was then multiplied for detection to obtain the RPN, the quantitative expression of each failure, ranged from 1 (1x1x1), as the "best" score, to 64 (4x4x4) as the "worst" one. RPNs rated between 1 and 9 do not need intervention; RPN between 12 and 18 need to be monitored; RPN between 24 and 32 need corrective actions; RPN between 36 and 64 require urgent corrective measures and validation plans.

To rate the occurrence of specific hazards we considered the incidence of previous failures, the presence of trained operators, the instrument maintenance process and the availability of specific standard operating procedures (SOPs).

According to the performed analysis, the most hazardous step was the analysis made by different operators with different instruments which needed a validation plan to prevent or limit errors in CD34+ detection and, as a consequence, errors in assessing quality control on HSCs.

Validation plan

To develop our work, a validation plan was prepared. This document describes the procedure and the aspects to be evaluated during the validation of the method and includes the aim of validation, the description of the method, roles and responsibilities, the acceptance criteria, and conclusion remarks.

Table 1: Failure Mode and Effects Analysis (FMEA) risk assessment. A) Calculation of failure risk; B) Calculation of risk priority number (RPN). C) Remediation plan. RPNs rated between 1 and 9 do not need intervention; RPN between 12 to 18 need to be monitored; RPN between 24 to 32 need corrective actions; RPN between 36 to 64 require urgent corrective measures and validation plan.

A

				OCCURENCE			
				LOW	MODERATE	PROBABLE	VERY PROBABLE
				FEW FAILURES	OCCASIONAL FAILURES	FAILURES ALMOST CERTAIN	VERY HIGH OCCURENCE
				1	2	3	4
SEVERITY	LOW	NO EFFECT, NO DANGER	1	1	2	3	4
	MODERATE	MINOR PART AFFECTED	2	2	4	6	8
	HIGH	LOSS OF PRIMARY FUNCTION	3	3	6	9	12
	VERY HIGH	SAFETY HAZARD	4	4	8	12	16

SEVERITY X OCCURENCE = FAILURE RISK

B

				FAILURE DETECTION			
				VERY HIGH	HIGH	MODERATE	LOW
				1	2	3	4
FAILURE RISK	LOW	NO EFFECT, NO DANGER	1	1	2	3	4
			2	2	4	6	8
			3	3	6	9	12
	MODERATE	MINOR PART AFFECTED	4	4	8	12	16
			6	6	12	18	24
	HIGH	LOSS OF PRIMARY FUNCTION	8	8	16	24	32
			9	9	18	27	36
	VERY HIGH	SAFETY HAZARD	12	12	24	36	48
			16	16	32	48	64

FAILURE RISK X FAILURE DETECTION = RPN

C

RPN			
1-9	12-18	24-32	36-64
No intervention required	Need to be monitored	Corrective actions	Corrective actions and validation plan

Table 2: Definition and interpretation of Failure Mode and Effects Analysis (FMEA) failure scores.

		Description	Examples
Severity			
Low	1	No effect, no danger	The error has no effect on results
Moderate	2	Minor part affected	Temporary instrument failure
High	3	Loss of primary function	Wrong dose of the collected graft
Very high	4	Safety hazard	Wrong dose of the infused graft
Occurrence			
Low	1	Few failures	No registered cases
Moderate	2	Some occasional failures	1-6 cases/year
Probable	3	Failures almost certain	7-12 cases/year
Very probable	4	Very high occurrence	Daily occurrence
Detection			
Very high	1	Always detected	Expiration date of reagents
High	2	Probably detected	Calibration of pipettes
Moderate	3	Middle probability of detection	Reverse pipetting error
Low	4	Probably undetected	Operator error

Samples

Stabilized blood (BD TM Stem Cell Control cod. 340991, Becton Dickinson) samples were stained with Stem-Kit reagent (cod. IM3630, Beckman Coulter) according to manufacturer's instructions. As used in external quality assessment (EQA), we chose to use the BD™ Stem Cell Control, a stable control with assigned values, routinely used to monitor the immunophenotyping process for CD34+ cells. The test was repeated on LOW (lot BC0622L: CD34+/ul=11.3 [5.9-16.7], % CD34+= 0.187 % [0.099-0.275]) and HIGH (lot BC0622H CD34+/ul =29.4 [19.5-39.3], % CD34+= 0.472 % [0.314-0.630]) samples.

Antibodies

All antibodies belonged to Stem-kit reagents: CD45 FITC Isoclonic Control-PE, CD45 FITC-CD34 PE, and 7-AAD Viability Dye were used according to manufacturer's instruction.

Lyse/no-wash single-platform viable cell enumeration and instruments

The ISHAGE guidelines were followed for estimating % and absolute number of CD34+ cells in single-platform method, using the Beckman Coulter Navios at Regina Margherita Children's Hospital (CIC 305-1), Beckman Coulter Navios at Candiolo Cancer Institute FPO-IRCCS (CIC 305-2), BD Biosciences FACSLyric™ at S. Luigi Hospital (CIC 305-3), and Beckman Coulter Navios EX at Mauriziano Hospital (CIC 305-4).

The count was performed by all four laboratories once on the LOW control, and once on the HIGH control for two consecutive days, for two weeks.

Statistical analysis

For each analysis repeatability limit (r), reproducibility error, uncertainty of reproducibility error and coefficient of variation (CV) were reported. **The reproducibility error** is the difference between the highest and lowest value measured for each parameter by each laboratory and describes the maximum variability of precision using different instruments and different operators. **Repeatability limit** (r) is the maximum acceptable deviation from the mean and it is lower than standard deviation. **The uncertainty of reproducibility** of error is the sum of the errors deriving from all the components of the test (pre-analytical phase, analytical phase, analysis phase), considered intrinsic to the measurement of that parameter. The **repeatability limit (r)** and the reproducibility error were used to compare test results within and between laboratories (9). **The coefficient of variation (CV)** is the ratio between the standard deviation and the mean and shows the extent of variability in relation to the mean. The higher the CV, the greater the dispersion. The **variance** was calculated for all the values of each laboratory and was used to calculate the **inter-laboratory variance.** The **maximum inter-laboratory variance** was the squared difference between the maximum and minimum values obtained by all laboratories.

The acceptance criteria were:

1) Reproducibility error lower than *r;*
2) CV%≤10%,
3) Inter-laboratory variance < maximum inter-laboratory variance.

Results

Risk analysis

In the validation plan, three main phases were identified: pre-analytical phase (transport of samples), processing and analysis phases, as reported in table 3.

For each step, checkpoints, criticisms and failures have been identified, and severity, occurrence and detection were assigned to each step to calculate the RPN. We found RPN ranged from 1 to 16 in most steps, where corrective actions are not requested, even if monitoring of the failure events is always recommended; while the data analysis and reporting phases reached a RPN of 36 needing specific action plans as a measure to prevent possible failures. These steps are the most difficult steps to keep under control, influenced by the interpretation of the data and by inter-laboratory reproducibility, which has been validated.

Method validation

The absolute CD34+ count was performed by the ISHAGE method on the same starting material with different instruments, operators and reagents. Counting has been performed on stabilized peripheral blood samples of known titres of LOW level (low CD34+ content) and HIGH level (high CD34+ content) by all four laboratories for two consecutive days and repeated in the same way the following week. The results obtained by four laboratories for low and high control was in the range indicated by BD TM Stem Cell Control. The repeated measurements (in percentage and absolute value) from each laboratory (or for each instrument) have a variability, expressed as reproducibility error, systematically lower than the (r) for that single parameter (low: 0.0421 for CD34+%, 2.3121 for CD34+/ul; high: 0.1026 for CD34+%, 3.6249 for CD34+/ul). The corrected reproducibility error, which represents the reproducibility error subtracted from the uncertainty of reproducibility, is always lower than (r) (low: 0.0427 for CD34+%, 2.2549 for CD34+/ul; high: 0.0854, for CD34+%, 3.1654 for CD34+/ul) except for the percentage value of the "low" count where it is slightly higher (0.0427 with r = 0.0421). The analysis of inter-laboratory variance (low: 0.0002 for CD34+%, 0.3343 for CD34+/ul; high: 0.0011 for CD34+%, 0.6025 for CD34+/ul) is within the maximum acceptable variance value (low: 0.0009 for CD34+%, 2.5122 for CD34+/ul; high 0.0036 for CD34+%, 4.9506 for CD34+/ul). In addition, the CV of all measurements for each parameter analysed is less than 8%, indicating low measurement variability among laboratories. Overall results were summarized in table 4.

Table 3: Validation plan

PROCESS STEPS	MATERIALS	INSTRUMENTS	CHECK POINTS	CRITICISMS	FAILURES	SEVERITY	POTENTIAL CAUSES	OCCURENCE	FAILURE RISK	CONTROLS	DETECTION	RPN
1. Pre-analytical phase												
Transport of samples	PB, PBMC, BM	Courier Transport box	Check on integrity of sample	Temperature during transport	Deterioration of sample	3	Courier not compliant Temperature not compliant	1	3	Training of courier	1	3
				Delay in delivery	Delay in processing		Communication errors Problems during transport	2	6	Training of courier	1	6
2. Processing												
Evaluation of [WBC]	PB/BM	Haematology analyser	Cell concentration	Cell concentration not compliant for staining	Results not accurate	2	Instrument failure	2	4	EQA	1	4
	PBMC											
Labelling	PBS, Monoclonal antibodies	Pipette, tips	Control sample	Errors in staining	Results not accurate	3	Pipette failure	2	6	Annual calibration	2	12
				Reverse pipetting			Operator not trained/ distraction			Competence evaluation, training and retraining		
Acquisition	Lysis solution	Pipette	Stem Count Fluorospheres	Lysis solution not compliant	Results not accurate	3	Pipette failure	2	6	Annual calibration	1	6
	Fluorosphere			Reverse pipetting			Operator not trained/ distraction			Competence evaluation, training and retraining		
3. Analysis												
Data analysis and reporting	Software	Cytometer	Daily QC	OOS	Errors in results	4	Cytometer failure	2	8	Periodic OQ/PQ	2	16
			Isotype control	Gating strategy		4	QC reagent expired	2	8	Reagent control	1	8
							Operator not trained/ distraction			Competence evaluation, training and retraining		
			Results	Inter-laboratory reproducibility	Results not accurate	4	Operators/ Instruments/ reagents	3	12		3	36

Table 4: Summary of results

	BD TM Stem Cell Control LOW		BD TM Stem Cell Control HIGH	
	CD34+ %	**CD34/ul**	**CD34+ %**	**CD34/ul**
Sample Range	0.099-0.275	5.9-16.7	0.314-0.630	19.5-39.3
Mean	0.1950	12.2690	0.4705	29.6510
SD	0.0150	0.8258	0.0366	1.2946
CV	7.7148	6.7305	7.7856	4.3662
Repeatability limit (r)	0.0421	2.3121	0.1026	3.6249
Reproducibility error	0.0600	3.1700	0.1200	4.4500
Uncertainty of reproducibility error	0.0173	0.9151	0.0346	1.2846
Corrected reproducibility error	0.0427	2.2549	0.0854	3.1654
Inter-laboratory variance	0.0002	0.3343	0.0011	0.6025
Maximum acceptable variance	0.0009	2.5122	0.0036	4.9506

Discussion

FACT–JACIE standards have evolved over time, with scheduled review and revision based on the rapidly changing fields of HSCT and cellular therapy. Quality management vision also changed. The systems were primarily based on a "safety first approach.", while nowadays it is moving towards ensuring that "as many things as possible go right" (10). When HSCs are used, CD34 flow cytometry is the gold standard for stem cell enumeration both in peripheral blood (PB) and in the stem cell product (apheresis, bone marrow or cord blood) prior to transplantation. It requires well-trained and experienced operators and is a technique that can be technically challenging (11). The number of infused viable CD34+ cells strongly correlates with the time to hematologic recovery of the patient (12). Therefore, in case of autologous transplantation, the number of CD34+ cells in PB is a guide to the start of collection, the blood volume to be processed, and the number of apheresis procedures to be performed. Successful mobilization of PB stem cells and adequate stem cell collection are of critical importance. Doses of 2×10^6 to 5×10^6 CD34+ cells/kg body weight are associated with more rapid engraftment and a lower probability of graft failure (13). Low CD34+ cell doses are associated with increased cost and worse outcome after tandem autologous stem cell transplantation in patients with relapsed or refractory germ cell tumours. Moreover, in the autologous setting, enumeration of viable CD34+ cells at the time of infusion becomes particularly relevant in patients in whom stem cell mobilization has been problematic and/or in which a total amount of stem cells was collected which was borderline to ensure a safe transplant procedure (i.e. 1.0-2.5 × 106 CD34+/kg) (14). Stem cell laboratories should have policies and procedures that address interruption in routine activity due to equipment failure or other emergency that may occur, so that such interruptions do not adversely affect cellular therapy products, critical supplies, and processes (1). In a program that involves the sharing of processes and procedures, the availability of different services and laboratories is a strength. However, the risk that an inaccurate CD34+ enumeration may affect the efficacy of the cell-based product administered to patients need to be taken into consideration. The inter-laboratory variability and reproducibility of the method is even more important when talking about cryopreserved cells. CD34+ cell counting techniques is well standardized on fresh samples, whereas the cytometry analysis of thawed samples is still controversial, and no validated techniques are yet available (15,16). In this context, we set up a risk analysis to test the inter-laboratory reproducibility of the CD34+ enumeration method, among the four TMTC laboratories, analysing the repeatability limit (r), the reproducibility error, the uncertainty of reproducibility error and the CV (CV). As reported since many years, participation in external EQA proficiency testing improve the accuracy of the method. EQA participation coupled with effective laboratory monitoring and remedial action is strongly associated with improved laboratory accuracy, and therefore with more appropriate patient treatment decisions (17). All the TMTC laboratories participate in external quality assessment that is an integral part of laboratory work and mandatory when the results have a clinical application. As used in EQA, we chose to use the the BD™ Stem Cell Control, a stable control, with assigned values routinely used to monitor the immunophenotyping process for CD34+ cells, and to ensure that the lab processes and operations run efficiently and guarantees the production of accurate and reproducible results. BD™ Stem Cell Control has been adopted by several Transfusional Centers around Europe. In the last 12 months this specific Quality Control is in use in 310 Laboratories in Europe (65 of these Laboratories are located in Italy), Middle East and Africa. The low values of the reproducibility errors indicated high reproducibility between laboratories. Moreover, the low inter-laboratory variability was demonstrated also by CV%< 10%. Evaluating the overall data, we can assume that the four laboratories are perfectly aligned and the results are reproducible. The statistical data obtained in this validation work, led us to demonstrate the crucial importance an accurate risk analysis shared by all the laboratories involved. The standardization of the method in use among the different laboratories of the TMTC allows the optimization of the processes, and guarantees the continuity of the services even in situations of emergency and disasters. As already reported, the use of a common standardized protocol, targeted training and external quality assessment significantly reduces intra- and inter-laboratory CD34+ cell count variation (18).

Acknowledgements

We acknowledge Camilla Francesca Proto for English revision.

Author Contributions:

IF supervised the work and was responsible of writing and revision of the paper; DR, SC were responsible for designing and writing the validation protocol, analysing the samples and contributed to write the paper; LG, AP, MP were responsible for analysing the samples and contributed to write the paper; MG was responsible for extracting and analysing data and interpreting results; FF was responsible of critical reading of the paper.

Competing Interests

The authors declare that they have no competing interests in relation to the work described.

Data Availability Statement

Raw data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Wuchter P. Processing, Cryopreserving and Controlling the Quality of HSCs. In: The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies. Cham (CH), Springer. 2019;. DOI: 10.1007/978-3-030-02278-5 17

2. Mfarrej B, Gaude J, Couquiaud J, Calmels B, Chabannon C, Lemarie C. Validation of a flow cytometry-based method to quantify viable lymphocyte subtypes in fresh and cryopreserved hematopoietic cellular products. Cytotherapy 2021;23:77-87. https://doi.org/10.1016/j.jcyt.2020.06.005.

3. Jaye DL, Bray RA, Gebel HM, Harris WAC, Waller EK. Translational applications of flow cytometry in clinical practice. J Immunol 2012;188:4715-9. https://doi.org/10.4049/jimmunol.1290017.

4. Sutherland DR, Anderson L, Keeney M, Nayar R, Chin Yee I. The ISHAGE guidelines for CD34+ cell determination by flow cytometry. International society of hematotherapy and graft engineering. J Hematother 1996;5:213 26. https://doi.org/10.1089/scd.1.1996.5.213

5. Keeney M, Chin-Yee I, Weir K, Popma J, Nayar R, Sutherland DR. Single platform flow cytometric absolute CD34+ cell counts based on the ISHAGE guidelines. Cytometry 1998;34:61-70.

6. Monteleone R and Klarmann D. Qualification and Validation. In Quality Management and Accreditation in Hematopoietic Stem Cell Transplantation and Cellular Therapy: The JACIE Guide. Chapter 5, Springer 2021. ISBN 978-3-030-64491-8 ISBN 978-3-030-64492-5 (eBook).

7. Najafpour Z, Hasoumi M, Behzadi F, Mohamadi E, Jafary M, Saeedi M. Preventing blood transfusion failures: FMEA, an effective assessment method. BMC Health Serv Res 2017;17:453. https://doi.org/10.1186/s12913-017-2380-3

8. Teofili L, Bianchi M, Valentini CG, Bartolo M, Orlando N, Sica S. Validation plan of bone marrow collection, processing and distribution using the failure mode and effect analysis methodology: a technical report. Cytotherapy 2022;24:356-364. https://doi.org/10.1016/j.jcyt.2021.10.005

9. Cool P, Ockendon M. Stats Book. http://countcool.com. ISBN 978-0-9560215-7-1

10. Snowden JA, McGrath E, Orchard K, Kröger N, Sureda A, Gratwohl A. Visions for a JACIE Quality Management System 4.0. Bone Marrow Transplant 2021;56:2876-2881. https://doi.org/10.1038/s41409-021-01467-8

11. Whitby A, Whitby L, Fletcher M, Reilly JT, Sutherland DR, Keeney M, Barnett D. ISHAGE protocol: are we doing it correctly? Cytometry B Clin Cytom 2012;82:9-17. https://doi.org/10.1002/cyto.b.20612

12. Schmitz N, Barrett J. Optimizing engraftment source and dose of stem cells. Semin Hematol 2002;39:3-14. https://doi.org/10.1053/shem.2002.29245.

13. Hyder MA, Goebel WS, Ervin KD, Schwartz JE, Robertson MJ, Thakrar TC, Albany C, Farag SS. Low CD34+ Cell Doses Are Associated with Increased Cost and Worse Outcome after Tandem Autologous Stem Cell Transplantation in Patients with Relapsed or Refractory Germ Cell Tumors. Biol Blood Marrow Transplant. 2018;24:1497-1504. https://doi.org/10.1016/j.bbmt.2018.01.032

14. Lanza F, Campioni D, Hellmann A, Milone G, Wahlin G, Walewski J, et al. Editorial Transfusion and Apheresis Science 59 (2020) 102785 2. Individual quality assessment of autografting by probability estimation for clinical endpoints: a prospective validation study from the European group for blood and marrow transplantation. Biol Blood Marrow Transplant 2013;19:1670-6. https://doi.org/10.1016/j.bbmt.2013.08.005

15. Fournier D, Lewin A, Simard C, Trépanier P, Néron S, Ballerini L, Codinach M. Multi-laboratory assay for harmonization of enumeration of viable CD34+ and CD45+ cells in frozen cord blood units. Cytotherapy 2020;22:44–51. https://doi.org/10.1016/j.jcyt.2019.10.009

16. Lanza F, Saccardi R, Seghatchian J. New horizons on stem cell cryopreservation through the artificial eyes of CD34+, using modern flow cytometry tools. Transfus Apher Sci 2020;59:102785. https://doi.org/10.1016/j.transci.2020.102785

17. Bainbridge J, Rountree W, Louzao R, Wong J, Whitby L, Denny TN, Barnett D. Laboratory Accuracy Improvement in the UK NEQAS Leucocyte Immunophenotyping Immune Monitoring Program: An Eleven-Year Review via Longitudinal Mixed Effects Modeling. Cytometry B Clin Cytom 2018 Mar;94(2):250-256. https://doi.org/10.1002/cyto.b.21531

18. Barnett D, Granger V, Kraan J, Whitby L, Reilly JT, Papa S, Gratama JW. Reduction of intra- and interlaboratory variation in CD34+ stem cell enumeration using stable test material, standard protocols and targeted training. DK34 Task Force of the European Working Group of Clinical Cell Analysis (EWGCCA). Br J Haematol. 2000;108:784-92. doi: 10.1046/j.1365-2141.2000.01932.

Phenotype similarities in automatically grouped T2D patients by variation-based clustering of IL-1β gene expression

Lucio José Pantazis[1*], Gustavo Daniel Frechtel[2,3*], Gloria Edith Cerrone[2,4*], Rafael García[1], Andrea Elena Iglesias Molli[2*]

[1]Centro de Sistemas y Control, Instituto Tecnológico de Buenos Aires (ITBA), Lavardén 315 1437, Ciudad Autónoma de Buenos Aires, Argentina.
[2]CONICET-Universidad de Buenos Aires. Instituto de Inmunología, Genética y Metabolismo (INIGEM). Laboratorio de Diabetes y Metabolismo. Avenida Córdoba 2351, Ciudad Autónoma de Buenos Aires, Argentina.
[3]Universidad de Buenos Aires. Facultad de Medicina. Departamento de Medicina. Cátedra de Nutrición. Avenida Córdoba 2351, Ciudad Autónoma de Buenos Aires, Argentina.
[4]Universidad de Buenos Aires. Facultad de Farmacia y Bioquímica. Departamento de Microbiología, Inmunología, Biotecnología y Genética. Cátedra de Genética. Avenida Córdoba 2351, Ciudad Autónoma de Buenos Aires, Argentina.
**Corresponding author: lpantazis@itba.edu.ar*

Article Info

Author of correspondence:
Lucio José Pantazis;
E-mail: lpantazis@itba.edu.ar;
Address:
Centro de Sistemas y Control, Instituto Tecnológico de Buenos Aires (ITBA), Lavardén 315 1437, Ciudad Autónoma de Buenos Aires, Argentina

Keywords
Shape-based clustering, Longitudinal data, Gene expression, Type 2 Diabetes, Knowledge Discovery in Databases.

Abstract

Background: Analyzing longitudinal gene expression data is extremely challenging due to limited prior information, high dimensionality, and heterogeneity. Similar difficulties arise in research of multifactorial diseases such as Type 2 Diabetes. Clustering methods can be applied to automatically group similar observations. Common clinical values within the resulting groups suggest potential associations. However, applying traditional clustering methods to gene expression over time fails to capture variations in the response. Therefore, shape-based clustering could be applied to identify patient groups by gene expression variation in a large time metabolic compensatory intervention.

Objectives: To search for clinical grouping patterns between subjects that showed similar structure in the variation of IL-1β gene expression over time.

Methods: A new approach for shape-based clustering by IL-1β expression behavior was applied to a real longitudinal database of Type 2 Diabetes patients. In order to capture correctly variations in the response, we applied traditional clustering methods to slopes between measurements.

Results: In this setting, the application of K-Medoids using the Manhattan distance yielded the best results for the corresponding database. Among the resulting groups, one of the clusters presented significant differences in many key clinical values regarding the metabolic syndrome in comparison to the rest of the data.

Conclusions: The proposed method can be used to group patients according to variation patterns in gene expression (or other applications) and thus, provide clinical insights even when there is no previous knowledge on the subject clinical profile and few timepoints for each individual.

Introduction

Type 2 Diabetes (T2D) is one of the most complex, prevalent and heterogeneous diseases whose etiology involves multiple interactions between genetic predisposing factors and environmental triggers [1]. Inflammation is a relevant component of the pathophysiological alterations that define the progression from metabolic syndrome to T2D [2]. Interleukin-1 beta (IL-1β) is a proinflammatory cytokine related to this clinical inflammation in T2D individuals, and is a well-known immune system modulator secreted by activated macrophages that can affect β-cell function and reduce insulin secretion [3]. Currently, one of the most important lines of research in diabetes is precision medicine, with the principal aim of grouping T2D individuals in different clinical subtypes defined by biomarkers. This group identification could be translated into an emerging approach to disease treatment and prevention that considers individual variability in genes, environment and lifestyle [4]. In this way, Ahlqvist et al. could recently break down T2D subjects into five distinct subgroups, with an improvement prediction of disease progression and outcome by including six variables (age at diagnosis, body mass index [BMI], glycated haemoglobin [HbA1c] , Glutamic Acid Decarboxylase Autoantibodies [GADA], estimation of insulin secretion [HOMA-B] and estimation of insulin-sensitivity [HOMA-IS]) [5]. The measurement of GADA by Ahlqvist et al. assessed the possible diagnosis of LADA (Latent Autoimmune Diabetes in Adults). The role of precision medicine in diabetes management was recognized by the American Diabetes Association (ADA) in collaboration with the European Association for the Study of Diabetes (EASD), which launched the Precision Medicine Initiative in 2018 in Diabetes [6]. The ultimate goal of precision medicine is the personalized provision of medical care, with better recognition of people at high risk for the development of T2D and its complications, and the implementation of personalized treatments at the individual level. In this sense, artificial intelligence could be used to detect clinical subtypes by matching individuals to their combinations of different biomarkers, with techniques such as large-scale prediction models. In the last decades, there have been great developments in methods for gene expression analysis, giving rise to an abundant quantity of data [7]. Since the amount of data grows faster than the ability to understand their implications, methods that allow drawing conclusions from gene expression data can be very useful to narrow this gap. The analysis of gene expression data can be very challenging due to limited prior knowledge on the observed phenomenon, heterogeneity, noise in the data and missing observations in the subject data [8]. Therefore, data mining tools that can provide potential relationships among framework. Longitudinal studies include repeated measures of a variable of interest -usually called a response- in the same subject over time, yielding multiple responses per individual noted as a response trajectory. In this work, the response variable relates to gene expression at a certain time point and the response trajectory describes the evolution of gene expression for a certain individual over time. It must be pointed out that when there are few time points and mistimed measurements, the mathematical tools that can be used are limited. For example, Fourier transformations, the standard procedure for time series, are no longer valid for few measurements. In this setting, the increase or decrease of gene expression between different measurement occasions can be studied [9]. Variables can be very useful for a clinical comprehension Clustering algorithms aim to group observations according to some measure of similarity, or conversely, to separate observations according to dissimilarity. When quantitative variables are involved, the dissimilarity can be based on distance measures. The selection of these features is closely related to the application area and the research objective. Regarding clustering algorithms, K-Means is the most popular method due to the low computational complexity of the algorithm and performance in big data. A variation of this method is the Kernel Based K-means algorithm [10]. The major disadvantage of these algorithms is the susceptibility to outliers and to the random initial group assignment. Another alternative is the K-Medoids algorithm [11]. This algorithm is more robust to outliers and initialization than K-means. Some works proposed clustering subjects according to the corresponding variation of gene expression, suggesting associations between a certain behavior in the gene expression over time with other variables [9], [12]. Many publications used this approach assuming simultaneous measurements to cluster different genes according to the increase or decrease in their expression, defining groups of co-expressed genes, or activating and repressing genes [13-18]. On most occasions, data corresponding to different subjects are not simultaneously collected, and other strategies must be used. Möller et al., applied a clustering algorithm to the transcriptional program of budding yeast, allowing mistimed measurements [19-20]. In a similar way, similarities in the variation of gene expression, can suggest associations with observable clinical features, which can be a starting point for further investigation.

Objectives

The main objective of this work was to cluster subjects in order to find relationships between patterns in Interleukin 1-beta (IL-1β) variation and clinical metabolic variables from a database of T2D patients. Also, to focus potential associations with obesity and metabolic syndrome as central clinical phenotypes. In this article, we perform a new analysis of data from a cohort of patients previously studied by our research group [21].

Materials and methods

1. Prospective controlled study database

The database used for the development of the clustering algorithm included the results of a prospective controlled study conducted in patients with newly diagnosed T2D and hyperglycemia (HbA1c > 8%), and after 6 and 12 months of treatment to achieve metabolic remission (HbA1c < 7%). The treatment was personalized: each participant received the first-line pharmacological treatment, and in all cases lifestyle changes were included through diet and physical exercise. Detailed information on this population can be found in our previously published manuscript [21]. It was the first follow-up study that evaluated IL-1β mRNA expression in hyperglycemic people with T2D after glycemic normalization treatment.

The study was conducted in a group of 30 adults (23.33% were female subjects and 76.67% male subjects) with a median age of 46 years (IQR 18.75 years) recruited from the Diabetes Care Unit. All procedures performed in the study were in accordance with the ethical standards of the institutional research committee, the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Ethics Committees of the Hospital de Clínicas "José de San Martín" from Ciudad Autónoma de Buenos Aires and all the participants gave their written informed consent. An anonymized database for pre and post intervention (6 and 12 months) instances was constructed for the data mining study. All individuals informed their age and gender and anthropometric measurements (height, weight, and waist circumference), BMI and systolic and diastolic blood pressure (SBP and DBP, respectively) were determined by standardized protocols. Venous blood samples were drawn of every individual, high-density lipoprotein cholesterol (HDL-c), triglycerides (TG), fasting blood glucose (FBG) and HbA1c were measured in serum using standardized procedures [21]. Low-density lipoprotein cholesterol (LDL-C) was calculated by the Friedewald equation. Blood anticoagulated with EDTA K2 was used for mRNA extraction and IL-1β mRNA expression analysis.

2. Notation

The subjects considered in the study required the same number of repeated measures over time (noted as variable t at 0, 6 and 12 months from intervention). The different subjects were grouped according to the variation over time of the IL-1β gene expression (noted as variable r). With this notation, for example, t(i,j) and r(i,j) represent the time point and the gene expression at measurement number j for subject i, respectively.

3. Clustering methods

We used hard partitioning methods for quantitative features. Clustering applications was performed after three sequential definitions:

A) set of variables to be considered;
B) distance between different variable observations defined in item A;
C) clustering algorithm that groups observations defined in item A according to the distance function defined in item B.

A) The clustering objective was set on grouping subjects according to the increase or decrease of gene expression, therefore the algorithm considered two subjects as similar if the corresponding slopes between time points are similar. For each subject i, the variation of gene expression r between time points j−1 and j is given by the following slope value:

$$m(i,j)=\frac{(r(i,j)-r(i,j-1))}{(t(i,j)-t(i,j-1))}$$

Therefore, if each subject i has a set of J repeated measures noted r(i), the same subject has a corresponding set of slopes m(i) with J-1 values. These sets of slopes will be noted as slope vectors. Thus, the slope vectors m(i) were used instead of using the response vectors r(i) for each subject i. Hence, the automatic grouping relied on the distance between slopes.

B) Whenever it was applicable, two distance functions were considered:

• the Euclidean distance [22], that adds the squared differences of slopes and applies a square root to the results. For example, for two subjects i and k the distance is computed as:

$$dE\ (m(i),m(k))=\sqrt{(m(i,1)-m(k,1))^2+\ldots+(m(i,J-1)-m(k,J-1))^2}$$

• the Manhattan distance [23], that adds the absolute values of the slope differences. For example, for two subjects i and k the distance is calculated as follows:

$$d_M\ (m(i),m(k))=|m(i,1)-m(k,1)|+\ldots+|m(i,J-1)-m(k,J-1)|$$

C) Regarding the clustering methods, three alternatives were applied

• K-Means (based on the Euclidean distance)
• Gaussian Kernel based K-Means (based on the Euclidean distance)
• K-Medoids (based on the Manhattan distance)

These clustering methods are used to group individuals according to their corresponding set of slopes m(i). Regarding the Kernel function required for Kernel based K-Means, a Gaussian kernel function was applied [10]. It must be pointed out that the K-Means based algorithms apply exclusively the Euclidean distance, whereas the K-Medoids algorithm allows the use of other distance functions, such as the Manhattan distance. The algorithms were applied using standard commands of R software. Details on the clustering algorithms are available in Hastie et al. [11]. In the sequel we will refer to clustering algorithms applied to the individuals' set of slopes as shape-based. For example, K-Medoids applied to the slope vectors m(i) can be referred as Shape-Based K-Medoids. More details on these procedures can be found in Appendix C.

4. Statistical Inference

Once the data was grouped in clusters, statistical tests were applied to the anthropometric and metabolic variables of the

database, searching for group differences in BMI, HDL-c, TG, LDL-c, FBG, HbA1c, Waist circumference, Age and number of Metabolic Syndrome components [ncMS], according to the Adult Treatment Panel III (ATPIII) guidelines [24]. To assess the statistical significance of differences between and within groups we performed non-parametric tests due to the small sample size and unverifiable assumptions. Kruskal-Wallis test was performed to assess the differences between groups [25], and paired Wilcoxon test to assess the differences within groups [26].

Results

The database was analyzed and subjects with at least one missing response value were excluded, due to the impossibility to attain a slope set comparable with other subjects. After removal, a total of 26 individuals remained for further research. The responses were scaled prior to partitioning, thus, the mean gene expression was subtracted and the result was divided by the corresponding standard deviation [27]. Figure 1 shows the different groups resulting from the applied clustering algorithms. The partitions of the algorithms involving K-means (upper [a] and central [b] panel of Figure 1), result in groups that are likely to mix stable and highly variable gene expression trajectories. This effect can be explained by the lack of robustness of the K-means algorithm. Inspecting the results, K-Medoids clustering (lower panel [c] of Figure 1) is preferred in this application based on the following observations: subjects in Cluster 1 had an initial decrease and a posterior increase, subjects in Cluster 2 showed an initial increase and a posterior decrease, whereas subjects in Cluster 3 had a stable level of IL-1β expression throughout the study, with small increases or decreases over time. Therefore, in the following, the results of the K-Medoids will be shown since the requirements of variation similarities are met. Furthermore, although all clustering methods are subject to randomness, the K-Medoids algorithm showed such robustness that running several times the procedure yielded the same partition. For the K-Medoids algorithm, it is worth mentioning that there was a subject in Cluster 2 whose gene expression increased in both time intervals and has been classified in this group due to the initial increase, which is not present in other clusters, and therefore, the algorithm located the subject in the most similar group. This subject could be morphologically seen as an outlier, and perhaps should have been classified in a separate group. However, a single subject cluster does not allow a correct between-group comparison. Given this clustering, a subsequent analysis was performed in the remaining variables of the database. The main results are given in Table 1. We found significant differences across groups in waist circumference, BMI, HDL-c and TG; and a tendency for LDL-c; but we did not find significant differences in FBG and HbA1c (Table 1). This similarity across groups is explained by the main objective of the original design of the study in order to follow up on the T2D individuals: to attain a decrease in HbA1c levels for all the participants. Also, since the Kruskal-Wallis detects differences between groups, further inspection of the values of most variables suggest that this difference is mainly observed in subjects from Cluster 1. Table 1 shows that subjects in Cluster 1 presented a decrease in LDL-c, TG and increase in HDL-c over time, whereas these values were stable for other clusters. Also, BMI and waist circumference values for subjects in Cluster 1 were smaller compared to those of the other clusters, also suggesting healthier features for Cluster 1. In addition, the Wilcoxon paired test was applied to all variables comparing the values at the start and the end of the study. The Wilcoxon test was not performed in ncMS and Age since the values do not vary over time. The lowest p-values corresponded to Cluster 1, suggesting greater differences in key variables for subjects in this group. Even if statistical significance was not achieved, the p-value is close to 10%, which represents considerable differences in the variables, given the small number of subjects and that non-parametric tests generally provide less statistical power. In addition, the p-values for Cluster 1 are considerably lower than the values corresponding to other clusters, reinforcing the observable difference between the evolutions of people from different clusters. Although we found differences in age, none of the variables analyzed showed a significant association with age (data not shown).

5. Discussion

In the current application, the K-Medoids clustering method using the Manhattan distance applied to the slopes attained the best results concerning the main objective, which was grouping subjects according to the variation in the response of IL-1β expression and showing differential behaviour in clinical variables. The other clustering algorithms considered in our work ([11]), when applied to the slopes yielded heterogeneous groups and therefore, did not meet the desired qualities for such clustering. Similar results are shown when applied to another controlled database in Appendix B. The use of the slopes as the key features of the grouping, allows to generalize previous proposals [20]. In this new framework, any traditional clustering method can be applied to group subjects according to variations in the response. Unlike the application of clustering algorithms in the original data r(i), small distances between the slope vectors m(i) provided similar characteristics in the variation of gene expression. Therefore, the use of the slopes expands the already vast world of clustering methods since these algorithms can be applied in both settings, but yielding different results. More details in Appendix A. The clustering yielded three distinct groups, evidently differentiable when clinically and biochemically compared in Table I. There were significant differences in waist circumference and BMI between the different clusters, so it would also be necessary to analyze the contribution of obesity in the expression of IL-1β that allowed these groups to be separated. Intra-cluster analysis showed that in Cluster 1, although the proposed metabolic compensation goal was reached, the decrease in FPG and HbA1c did not reach

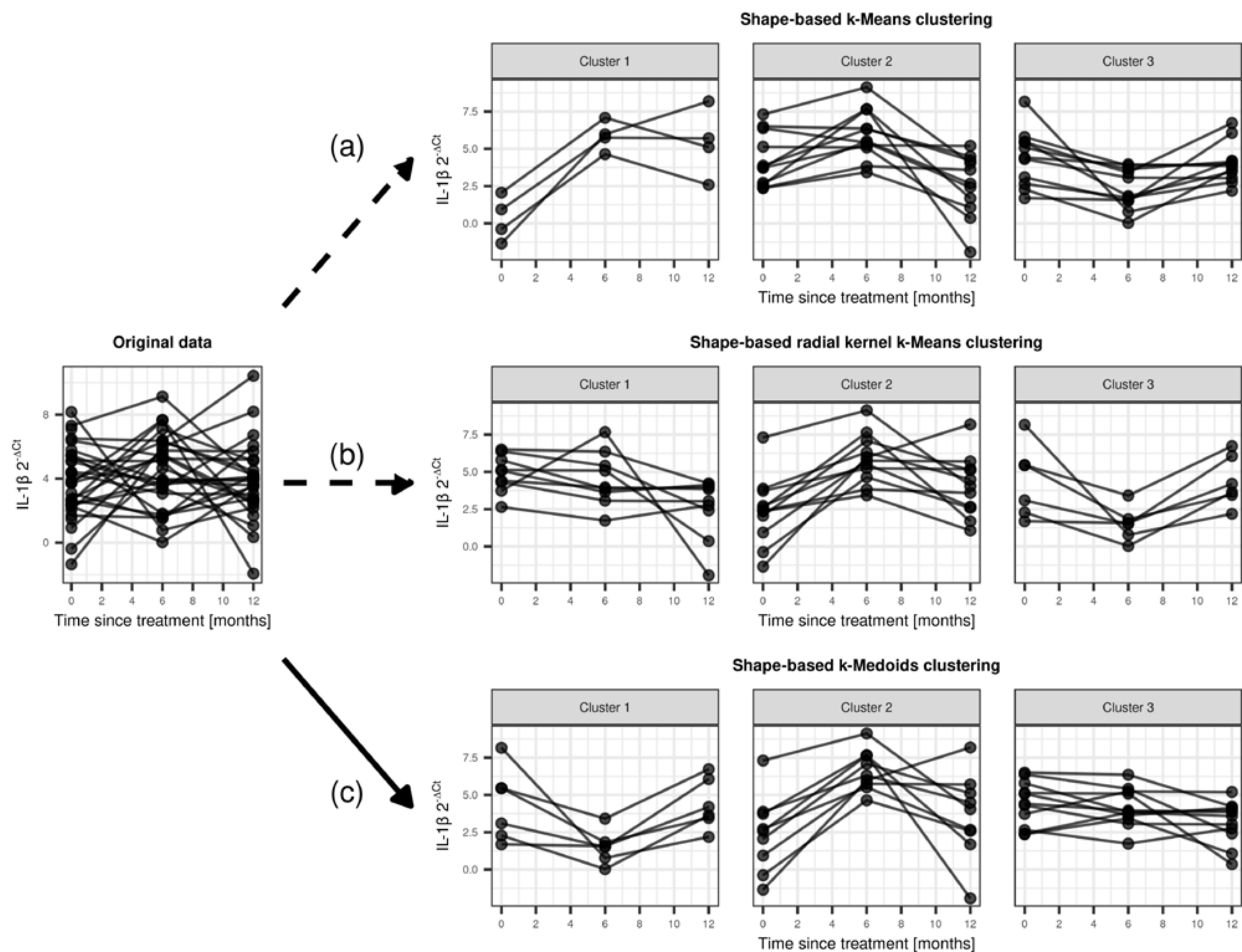


Figure 1: IL-1β ($2^{-\Delta Ct}$) expression over time, grouped according to the slopes between
time points using the three clustering algorithms described in Section 3.3: (a) K-Means (Upper panel), (b) Kernel K-Means (Center panel) and (c) K-Medoids (Lower panel).

Table 1: Observed differences in quantitative variables of the dataset, separated by time measurement (at 0, 6 and 12 months). The waist circumference results at 6 months were omitted due to a low proportion of observed data. m: median; IQR: interquartile range; BMI: body mass index; HDL-c: high-density lipoprotein cholesterol; LDL-c:low-density lipoprotein cholesterol; TG: triglycerides; HbA1c: glycated haemoglobin; FBG: fasting blood glucose; ncMS: number of Metabolic Syndrome components.

Variable	Time	Cluster 1 m (IQR)	Cluster 2 m (IQR)	Cluster 3 m (IQR)	P-value (Kruskal-Wallis)
Waist circumference (cm)	**0 mo**	100 (92-102)	104 (100-109)	112 (100-117)	*0.0149*
	12 mo	100 (97-103)	106 (101-114)	111 (98-119)	
p-value (Wilcoxon)	*0/12 mo*	*0.8922*	*1.0000*	*0.9056*	
BMI (kg/m²)	**0 mo**	31.11 (29.22-31.42)	32.91 (31.23-38.02)	34.02 (31.60-37.92)	*0.0106*
	6 mo	29.68 (29.18-30.11)	33.57 (29.56-38.22)	32.50 (31.65-35.75)	
	12 mo	30.48 (29.91-30.85)	33.28 (30.35-37.90)	32.80 (28.19-33.85)	
p-value (Wilcoxon)	*0/12 mo*	*0.0544*	*0.0852*	*0.1358*	
HDL-c (mmol/L)	**0 mo**	1.10 (1.01-1.31)	1.09 (0.83-1.16)	1.01 (0.91-1.03)	*0.0470*
	6 mo	1.22 (1.00-1.47)	1.06 (0.92-1.11)	1.05 (0.94-1.14)	
	12 mo	1.32 (1.34-1.45)	1.11 (0.98-1.40)	1.09 (0.98-1.11)	
p-value (Wilcoxon)	*0/12 mo*	*0.0544*	*0.0852*	*0.1358*	
LDL-c (mmol/L)	**0 mo**	3.22 (2.97-3.30)	3.00 (2.22-3.08)	3.29 (2.90-3.75)	*0.0718*
	6 mo	3.19 (2.87-3.44)	2.56 (2.37-2.97)	3.11 (2.82-4.40)	
	12 mo	2.28 (2.22-2.38)	2.72 (2.57-3.09)	3.13 (2.81-3.60)	
p-value (Wilcoxon)	*0/12 mo*	*0.1250*	*0.9453*	*0.8125*	
TG (mmol/L)	**0 mo**	1.46 (1.27-1.69)	2.06 (1.45-2.74)	1.51 (1.32-2.01)	*0.0047*
	6 mo	1.64 (0.97-1.88)	2.42 (2.09-3.20)	2.27 (1.86-2.94)	
	12 mo	0.93 (0.90-0.99)	2.19 (1.67-2.72)	1.47 (1.32-2.75)	
p-value (Wilcoxon)	*0/12 mo*	*0.1250*	*1.0000*	*0.7597*	
HbA_{1c} (%)	**0 mo**	8.6 (8.0-10.1)	9.5 (9.0-10.8)	8.1 (7.9-11.2)	*0.6652*
	6 mo	6.2 (6.1-6.4)	6.4 (5.9-6.9)	6.7 (5.8-7.2)	
	12 mo	5.9 (5.7-6.1)	6.1 (5.6-6.8)	6.2 (5.9-7.0)	
p-value (Wilcoxon)	*0/12 mo*	*0.0625*	*0.0039*	*0.0029*	
FBG (mmol/L)	**0 mo**	8.69 (7.41-15.17)	8.16 (7.38-15.01)	8.77 (7.33-12.10)	*0.8086*
	6 mo	5.91 (5.76-6.52)	5.94 (5.27-6.97)	6.33 (5.83-7.89)	
	12 mo	5.99 (5.83-6.22)	6.33 (5.61-6.66)	6.49 (6.27-7.44)	
p-value (Wilcoxon)	*0/12 mo*	*0.0625*	*0.0078*	*0.0322*	
ncMS		3 (2-4)	4 (3-4)	4 (3-5)	*0.05907*
Age (Years)		60 (57-62)	42 (39-52)	46 (40-58)	*0.00423*

statistical significance. Also, a decrease trend in BMI and metabolic improvements in HDL-c values were observed. In Clusters 2 and 3, the compensation goal was reached as shown by a significant decrease in HbA1c and FBG. In Cluster 2 we also found a downward trend in BMI and HDL-c; but there were no anthropometric or lipid variations in Cluster 3. These results demonstrated that Cluster 3 showed the worst metabolic profile. In subsequent studies, it would be interesting to evaluate variables related to cardiovascular risk. Usually, non-parametric tests are less powerful (prone to discard real differences as non-significant) and the p-values can also be affected by the small sample size [25]. Consequently, the standard significance level of 5% can be too restrictive for this particular application of the statistical tests and p-values which are higher but close to 5% were considered for analysis. However, the strength of the obtained results is enforced by the large time changes considering nutrition and physical individual habits, and also by the time-varying nature of the system under study. As future work, it would be necessary to analyze a larger number of individuals to improve the individualized model and to reinforce our conclusions. Most clinical applications of gene clustering algorithms, which can be phenotype-based or gene-based, do not consider the longitudinal evolutions of gene expression. To the best of our knowledge, this approach has not yet been addressed as a clinical application in the literature. In the work of Pearson et al. [28], the consideration of longitudinal evolution was focused on phenotype follow-up, rather than gene expression and our work considered both gene expression and phenotype over time. Further investigation could profit from the use of all these perspectives to improve algorithm performance. Furthermore, works of clinical application that considered the longitudinal evolution of gene expression used supervised learning algorithms, in which the outcome variable was known and used for further predictions [29-33]. The methodology presented in this work involves unsupervised learning and can be applied when this prior knowledge is absent or limited, and new associations are required. Also, since most available gene expression data comes from countries with strong European ancestry, further research could provide data from other countries that can enrich precision medicine, based on more diverse data sources [34]. Our work used hard partitioning to automatically group individuals from the study. Many other works focused on the use of soft clustering, which is preferred for big data [35]. However, in small studies like ours, with patients undergoing large time treatments, the groups should be well-defined in order to achieve an adequate between-group comparison. A possible extension of this work could be the application of soft clustering to the slopes in longitudinal studies with a great number of subjects, allowing the determination of larger groups according to a strong association. One of the advantages of the proposed data mining procedure is that it does not require time measurements to be equal among all individuals, which is a frequent imposition for similar algorithms. However, in this study, the measurements were taken with the same protocol for every subject and do not differ with great impact in the calculations, and the algorithm easily adapts to these situations. Furthermore, the algorithm is not restricted to gene expression and performs well in other applications, or in cases in which other methods are not recommended, with few time points in which there is no prior knowledge regarding the observed phenomenon, which is a frequent issue in case studies observed in clinical investigation. Also, it is important to remark that this lack of prior knowledge allows us to search for associations between variables that are not previously thought to be linked. However, it must be pointed out that any prior knowledge regarding the application can be used to improve the algorithm, allowing the selection of specific distances between slopes. Since the presented database does not have a massive number of observations, the computational cost of K-Medoids was a drawback without major consequences. However, in other types of databases as massive databases, K-Means or Kernel-based K-means can be a better option. Another issue worth mentioning is that the proposed method is analytical and should not be used as a statistical inference tool. Any result obtained with the method should be further tested in a controlled experiment with a bigger sample size in order to attain satisfactory and pertinent inferences.

Conclusions

Our study showed that clustering individuals according to the variation in gene expression enabled us to find important clinical features that could allow the identification of differentially grouped metabolic behaviors not attained by other data analysis. With further studies, this could be translated into clinical improvement management of each individual considering the group assignment. The achieved results show that the proposed approach can significantly improve predictive performance and is effective when other established methods are not recommended due to the nature of the data, such as small sample sizes, few timepoints, heterogeneity and abrupt changes in gene expression for different timepoints. T2D is a complex and heterogeneous disease. Therefore, identifying clusters with similar clinical phenotype, will allow health professionals to evaluate increased risk, assess clinical evolutions and apply specific and personalized treatment to these groups of individuals. Precision medicine can improve the quality of life of people with T2D and help them improve glycemic control, prevent complications and provide a better quality of life.

Construction of the slope space

Selecting a simmilarity measure for a specific problem is a challenging task, due to the vast number of choices. Since this work is focused on grouping subjects according to the increase or decrease of a certain response variable, the euclidean distance between vectors may fail to recognize differences in the variation. A clear example of this problem is shown in Figure 1:

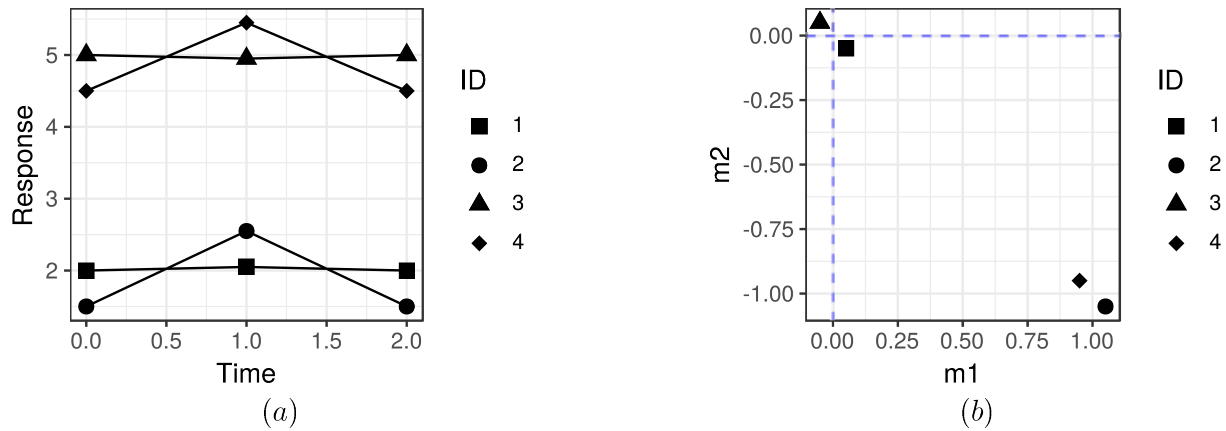


Figure 1. (a) Simulated response trajectories of 4 subjects at 3 time points. (b) Representation of these trajectories en the slope space.

In Figure 1 (a), the responses for subjects 1 and 2 are close to each other at different times, yielding a small euclidean distance between both vectors r_1 and r_2. The same feature can be observed between subjects 3 and 4. Conversely, the euclidean distance, does not consider subjects 1 and 3 as similar, as well as subjects 2 and 4. However, based on the variation of the response, the proposed distance should qualify subjects 1 and 3 as similar, the same should occur between subjects 2 and 4.

To address this issue, a natural distance is considering that to subjects are similar if the slopes between timepoints are similar. For example, according to the previous notations, for subject i, the variation of the response variable r between timepoint j and $j+1$ is given by the following slope value:

$$m_{i,j} = \frac{r_{i,j} - r_{i,j-1}}{t_{i,j} - t_{i,j-1}} \quad (1)$$

Under this construction, negative and positive slopes will correspond to decreases and increases in the response variable, respectively. On this basis, it may seem natural to group individuals according to the slope sign in the same instance. However, this classification omits the magnitude of the slope. If two subjects have slopes of different signs, but similar absolute value, these response trajectories reflect certain stability, and thus, are not qualitatively different.

Figure 1 (b) shows clearly that the slope space meets the goal of reducing the distance between subjects 1 and 3, as well as subjects 2 and 4. Furthermore, the figure shows how dividing observations by quadrants fails to assess the qualitative features of a trajectory, since subjects 1 and 3 belong to different quadrants but both correspond to stable trajectories in time and should belong to the same group. Grouping observations according to the distance in the slope space achieve this feature.

The resulting slope space yields a new set of quantitative values per individual, and the clustering methods detailed in the manuscript can be applied in this setting. For example, response trajectories $r_i = (r_{i,0}, r_{i,1}, \cdots, r_{i,J})$ and individual timepoints $t_i = (t_{i,0}, t_{i,1}, \cdots, t_{i,J})$ have $J+1$ real valued coordinates, whereas $m_i = (m_{i,1}, m_{i,2}, \cdots, m_{i,J}) \in \mathbb{R}^J$. Therefore, traditional distances for numeric vectors can be applied to the vector of slopes. For example, the distance between subject i and i' (with responses $r_{i,j}$ and

1

$r_{i',j}$ at times $t_{i,j}$ and $t_{i',j}$, respectively) can be expressed by the euclidean distance between the corresponding slopes $m_{i,j}$ and $m_{i',j}$:

$$(2) \qquad ||m_i - m_{i'}||_2 = \sqrt{\sum_{j=1}^{J} (m_{i,j} - m_{i',j})^2} = \sqrt{\sum_{j=1}^{J} \left(\frac{r_{i,j} - r_{i,j-1}}{t_{i,j} - t_{i,j-1}} - \frac{r_{i',j} - r_{i',j-1}}{t_{i',j} - t_{i',j-1}} \right)^2}$$

The K-means algorithm is based on the euclidean distance, and thus is frequently used in practice. However, other distances for quantitative vectors can be used, for example, the manhattan distance, based on the absolute value of the differences:

$$(3) \qquad ||m_i - m_{i'}||_1 = \sum_{j=1}^{J} \left| m_{i,j} - m_{i',j} \right| = \sum_{j=1}^{J} \left| \frac{r_{i,j} - r_{i,j-1}}{t_{i,j} - t_{i,j-1}} - \frac{r_{i',j} - r_{i',j-1}}{t_{i',j} - t_{i',j-1}} \right|$$

This distance is less susceptible to outliers, and can be used in both Hierarchal and K-Medoids algorithms, increasing the number of optional metrics beyond the Euclidean distance.

Unlike the application of these distances in the original data r_i, small distances in the slope space vectors m_i ensure similar characteristics in the response variation. Furthermore, any clustering method applied to the slope space succeeds to capture these features in a greater measure than the same clustering method applied to the trajectories. Therefore, the use of the slope space expands the already vast world of clustering methods since it can be applied in both settings, but yielding different results.

2

Figure 1: Appendix A

1. Growth Mixture Models

A different approach is to propose a model for the temporal evolution of the responses and a specified number K of unknown classes (the procedure is also called Latent Class Mixed Models [2],[1]). The model is usually polynomial with mixed effects for each group:

$$\begin{aligned} r_{i,j}^{(k)} &= \sum_{h=0}^{H} \alpha_h^{(k)} \cdot t_{i,j}^h + \varepsilon_{i,j} \\ \alpha_h^{(k)} &= \beta_h^{(k)} + \gamma_{h,i}^{(k)} \end{aligned} \quad (1)$$

where each β_h is a fixed effect per group and each $\gamma_{h,i}$ is an individual random effect.

Once the model is specified, the parameters are fitted by Maximum Likelihood and a posterior group classification of each observation is performed based on the estimated parameters. This classification yields K groups that can also be seen as a clustering or partition.

This approach has been proven useful for slowly changing trends. However, the model needs to be previously specified (requiring prior information which is not always available) and the model is not always clear, specially for data with small sample sizes. Furthermore, whenever the changes in temporal trends are abrupt, the estimated coefficients are greatly affected by these variations, influencing the entire fitted trajectory.

To avoid this drawback, the trends can be modeled with polynomial splines per group, diminishing the influence of poorly estimated coefficients. However, the timepoints used as knots are not always explicit and must be determined adding a new difficulty.

2. Clustering quality measures

Given an automatic clustering, it is not always clear how to evaluate the quality of the partition. The aforementioned diversity in clustering problems and objectives lead to several different indices, but these measures can be qualified in two groups: internal and external criteria.

Internal criteria are used to evaluate desirable qualities of a clustering, such as high between-cluster variability (or separability) and low within-cluster variability (or homogeneity), without a reference grouping. On the other hand, external criteria require two partitions and do not focus on the properties of the clusterings, they assess the similarity between two different groupings. In this work, only external criteria are addressed since the aim of the experiments is set on identifying a partition given as reference.

2.0.1. *External criteria.* Measuring agreement between clustering partitions is not as simple as matching the number of objects belonging to a certain clusters, mainly because most clustering algorithms have an initial random assignment. For example, running K-means twice on the same data, can lead to the same clustering, but with different labels given as an output.

To overcome this drawback, given a database of n individuals, most external criteria focus on the $\frac{n(n-1)}{2}$ different pairings of the data observations. If two observations x_i and $x_{i'}$ are grouped in the same cluster in one partition, the other partition will agree with this result if x_i and $x_{i'}$ are also in the same cluster, regardless of the cluster labels. Also, if two observations belong to different clusters in one partition, the other partition should also be assigned to different groups.

According to the notation used in the literature [4], given two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ for a dataset of n observations, and its corresponding $n_P = \frac{n(n-1)}{2}$ pairings, the following numbers are computed:

- yy pairs of observations grouped in the same cluster in both partitions.
- yn pairs of observations grouped in the same cluster in $\mathcal{P}_1$, but not in $\mathcal{P}_2$.
- ny pairs of observations grouped in different clusters in $\mathcal{P}_1$, but not in $\mathcal{P}_2$.
- nn pairs of observations grouped in different clusters in both partitions.

We remark that these notations do not require clusters to have the same label. Furthermore, the number of clusters of each partition can be different. In all cases, the sum of these four numbers add to the number of pairings n_P.

Based on these definitions, there are several criteria that can be applied to compare different partitions. This work relies on the following criteria:

- **Precision (P):** $C_P = \dfrac{yy}{yy + ny}$
- **Recall (RC):** $C_{Rc} = \dfrac{yy}{yy + yn}$
- **Rand (RN):** $C_{Rn} = \dfrac{yy + nn}{n_P}$
- **Jaccard (J):** $C_J = \dfrac{yy}{yy + yn + ny}$
- **Czekanowski-Dice (CD):** $C_{CD} = \dfrac{2yy}{2yy + yn + ny}$
- **Folkes-Mallows (FM):** $C_{FM} = \dfrac{yy}{\sqrt{(yy + yn) \cdot (yy + ny)}}$
- **Kulczynski (K):** $C_K = \dfrac{1}{2} \cdot \left(\dfrac{yy}{yy + ny} + \dfrac{yy}{yy + yn} \right)$
- **Rogers-Tanimoto (RGT):** $C_{RGT} = \dfrac{yy + nn}{yy + nn + 2(yn + ny)}$

It is worth mentioning that all these criteria correspond a higher index with a greater agreement between partitions. Furthermore, note that in the best case scenario, both yn and ny are zero, and all these indices have a maximum value of 1.

These criteria are very useful whenever there is a reference partition and the goal is to assess the agreement of an automatic partition to the reference grouping.

3. Benchmark database

The variation-based clustering algorithms are tested in a longitudinal benchmark database, in which the subjects are naturally grouped. The main goal is to cluster automatically the trajectories according only to the variation in the response, and compare the results to the reference grouping via external criteria.

TLC Data. The Treatment of Lead-Exposed Children (TLC) trial ([3]) is a randomized study that analyses the effects of a drug named succimer in children with similar blood lead levels. These data consist of four repeated measurements of blood lead levels obtained at baseline (or week 0), week 1, week 4, and week 6 on 100 children, randomly assigned to treatment with succimer or placebo.

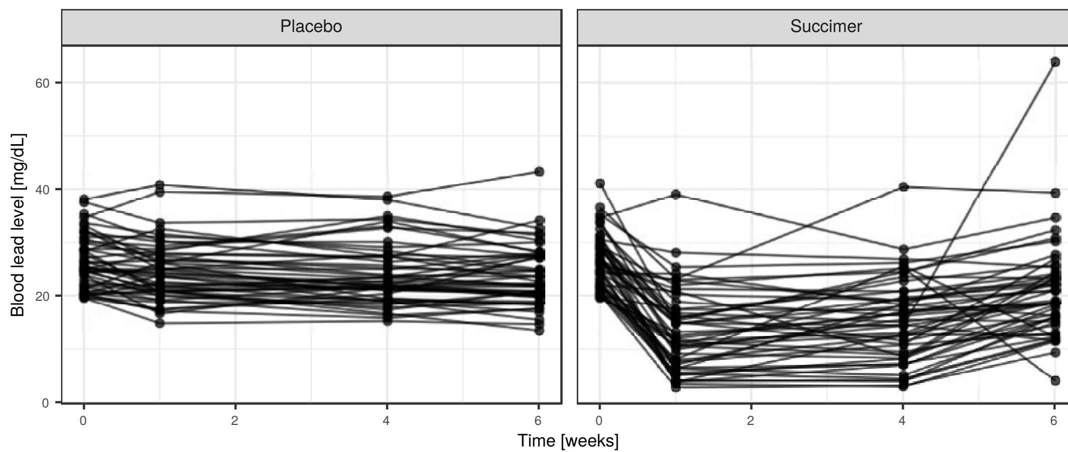


Figure 1. Response trajectories for subjects in the TLC study,

As the response trajectories in Figure 1 show, the blood lead levels are stable in the placebo group and there is a strong decrease in blood lead levels in the succimer group during the first week, but an increase in the remain of the study. Therefore, the slopes in the placebo group are expected to be close to zero.

2

On the other hand, the slopes in the succimer group are expected to be negative in the first instance, and relatively stable after the first week.

3.1. **Results .** Figure 2 reflects the quality indices for the different clustering methods described in the manuscript, and adding Latent Class Mixture Models applied to the TLC data. Following Figure 1, a linear spline model is considered with a knot in the timepoint corresponding to the first week. Also, Kernel-based K-Means is included with an automatic selection of parameter σ.

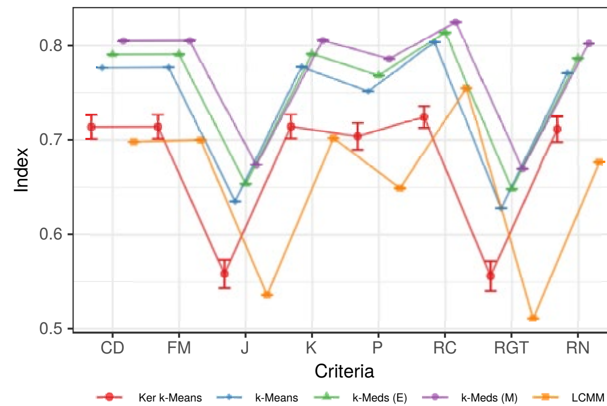


Figure 2. Mean indices for the TLC data

The k-Medoids algorithm using the Manhattan distance yields the best performance. Another detail worth mentioning is that even though for each method $M = 100$ repetitions were conducted, the indices remain unaffected and converge to the same partition, except for the Kernel-based k-Means. Therefore, there is a null standard error in almost all cases. The standard error for Kernel-based k-Means is considerable and can be explained by the sensitivity of this method to initialization and outliers.

References

[1] NGP Den Teuling, SC Pauws, and ER van den Heuvel. A comparison of methods for clustering longitudinal data with slowly changing trends. *Communications in Statistics-Simulation and Computation*, pages 1–28, 2020.

[2] Daniel P Martin and Timo von Oertzen. Growth mixture models outperform simpler clustering algorithms when detecting longitudinal heterogeneity, even with small sample sizes. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(2):264–275, 2015.

[3] WJ Rogan, RL Bornschein, Jr Chisolm, AI Damokosh, DW Dockery, ME Fay, RL Jones, GG Rhoads, NB Ragan, M Salganik, et al. Safety and efficacy of succimer in toddlers with blood lead levels of 20-44 μg/dl. *Pediatric research*, 48(5):593–599, 2000.

[4] Sławomir T Wierzchoń and Mieczysław A Kłopotek. *Modern algorithms of cluster analysis*. Springer, 2018.

Figure 2: Appendix B

1. Clustering methods

The objective of clustering methods is to group observations (in this case, the response trajectories r_i) according to some measure of similarity, or conversely, separating observations according to dissimilarity. The concept of similarity (or dissimilarity) is vague and can be applied to categorical or quantitative features (and transformed features), for different similarity measures, kernels, and even stablishing degrees of association between an individual and a group, leading to an immense number of options. The selection of these features is deeply related to the application area and the research objective.

This work focuses on crisp partitioning methods for quantitative features. The advantage of grouping quantitative values methods is that they can be based on distance measures for vectors, such as the Euclidean distance.

In the sequel and unless otherwise specified, the observations are noted as $x_i \in \mathbb{R}^m$, with $1 \leq i \leq n$. The clusters corresponding to a certain partition are noted $\mathcal{C}_k$ $(1 \leq k \leq K)$, where $K \leq n$, each with n_k elements, and an average value noted $\mu_k = \frac{1}{n_k} \cdot \sum_{x_i \in \mathcal{C}_k} x_i$. The number n_k of observations in each cluster satisfy $n = \sum_{k=1}^{K} n_k$.

In this work we focus on five clustering methods with corresponding variations: K-Means, Hierarchal clustering (Single, Average and Complete methods) and K-Medoids (using Manhattan and Euclidean distance), Kernel-based k-Means (Radial Kernel) and Latent Class Mixed Models.

Details regarding the used clustering methods are available in Hastie, Tibshirani & Friedman [18] and Wierzchoń & Kłopotek [26].

1.1. K-Means

K-Means is the most popular clustering method, mostly due to the low computational complexity of the algorithm and its performance in big data.

The main goal of the algorithm is to assign objects to clusters in order to minimize the within-group variance of the partition, and thus, based on the euclidean distance between vectors. However, this optimization problem is $\mathcal{NP}$-hard. Therefore, a heuristic approach is implemented iteritavely searching for a local minimum.

The algorithm requires a number of clusters K. In the first step, observations are randomly assigned to K groups. Within each group $1 \leq k \leq K$, the group mean $\mu_k \in \mathbb{R}^m$ is calculated. After this computation, each observation x_i is assinged to the group k, where μ_k is the closest group mean to x_i. The mean computation and further assignment is iterated until the resulting groups remain unchanged.

The major disadvantage to this algorithm is the lack of robustness to outliers of the calculated group mean, and to the initial group assignment. Also, the algorithm requires the number K of clusters and in practice, it can be very difficult to know in advance the number of groups.

1.2. K-Medoids

An alternative to K-Means is the K-Medoids algorithm. This algorithm is more robust to outliers and initialization than K-means, since it relies on observations of the database as group centers, instead of group averages.

However, the robustness comes with a cost of increased computation complexity. Therefore, this algorithm can preform very well when the number of observations are not massive.

The algorithm starts selecting K random observations as cluster representatives, and assigns the remaining $n - K$ observations to the closest center. Once the clusters are assigned, each cluster center is updated to the observation that minimizes the within-group distance, iterating until cluster assignation does not change.

1.3. Kernel-based K-Means

Kernel-based K-Means is a usual alternative when the observations are not linearly separable. In order to increase separability, a non-linear transformation Φ is applied to the data and the same K-Means algorithm described in Section 1.1 to the transformed data.

1

One of the most popular kernels is the Gaussian Kernel or Radial Basis Function:

$$\mathcal{K}(x_i, x_{i'}) = e^{-\frac{||x_i - x_{i'}||}{2\sigma^2}} \qquad (1)$$

where σ is a parameter defined by the user.

The use of kernels flexibilizes the use of K-Means. However, this extension does not modify the previous drawbacks: the algorithm is very sensitive to outliers and initialization. Furthermore, the inclusion of parameter σ adds a new issue: the results can be qualitatively different when the parameter values are modified.

1.4. **Algorithm.** Algorithm 1 describes the pseudocode for the structure of the algorithm. Only subjects with complete responses and time variables necessary to calculate the slope are included in the clustering algorithm. For these remaining subjects, the vector of slopes is attained and a clustering algorithm is performed. It must be pointed out that any distance function or clustering method can be applied at this point in the algorithm, yielding a great versatility for this approach.

Algorithm 1 Shape-based Clustering of Longitudinal data

```
procedure SHAPE-BASED CLUSTERING(Data, Resp, Time, J, ID)
    NetData ←FilterMissingData(Data, Resp, Time, ID)
    IDs ← UniqueID(Data, ID)
    NumIDs ← Length(IDs)
    i ← 1
    IDsComp ← ∅
    m ← ∅
    loop1:
    while i ≤ NumIDs do
        IndData ← SelectID(Datos, IDs(i))
        IndResp ← SelectResponse(IndData, Resp)
        IndTime ← SelectTime(IndData, Time)
        if Length(IndResp) = J + 1 and Length(IndTime) = J + 1 then
            IDsComp ← Append(IDs(i), IDsComp)
            m_Aux ← ∅
            j ← 1
            loop2:
            while j ≤ J do
                m_Aux(j) ← (IndResp(j + 1) − IndResp(j)) / (IndTime(j + 1) − IndTime(j))
                j ← j + 1
                go to loop2.
            m ← RowBind(m, m_Aux)
        i ← i + 1.
        go to loop1.
    AssignClusts ← ClusteringAlgorithm(m)
    Results ← RowBind(IDsComp, AsignClusts)
     return Results
```

L. Pantazis, R. García, G. Frechtel, G. Cerrone, A. Molli, A. Millán

References

[1] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
[2] Sławomir T Wierzchoń and Mieczysław A Kłopotek. *Modern algorithms of cluster analysis*. Springer, 2018.

3

Figure 3: Appendix C

References

1. Bowman P, Flanagan SE, Hattersley AT. Future Roadmaps for Precision Medicine Applied to Diabetes: Rising to the Challenge of Heterogeneity. J Diabetes Res. 2018 Nov 27;2018:3061620. doi: 10.1155/2018/3061620. PMID: 30599002; PMCID: PMC6288579.

2. Cruz NG, Sousa LP, Sousa MO, Pietrani NT, Fernandes AP, Gomes KB. The linkage between inflammation and Type 2 diabetes mellitus. Diabetes Res Clin Pract. 2013 Feb;99(2):85-92. doi: 10.1016/j.diabres.2012.09.003. Epub 2012 Dec 14. PMID: 23245808.

3. Nackiewicz D, Dan M, He W, Kim R, Salmi A, Rütti S, Westwell-Roper C, Cunningham A, Speck M, Schuster-Klein C, Guardiola B, Maedler K, Ehses JA. TLR2/6 and TLR4-activated macrophages contribute to islet inflammation and impair beta cell insulin gene expression via IL-1 and IL-6. Diabetologia. 2014 Aug;57(8):1645-54. doi: 10.1007/s00125-014-3249-1. Epub 2014 May 12. PMID: 24816367.

4. Florez JC. Precision Medicine in Diabetes: Is It Time? Diabetes Care. 2016 Jul;39(7):1085-8. doi: 10.2337/dc16-0586. Epub 2016 Jun 11. PMID: 27289125.

5. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, Wessman Y, Shaat N, Spégel P, Mulder H, Lindholm E, Melander O, Hansson O, Malmqvist U, Lernmark Å, Lahti K, Forsén T, Tuomi T, Rosengren AH, Groop L. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018 May;6(5):361-369. doi: 10.1016/S2213-8587(18)30051-2. Epub 2018 Mar 5. PMID: 29503172.

6. Nolan JJ, Kahkoska AR, Semnani-Azad Z, Hivert MF, Ji L, Mohan V, Eckel RH, Philipson LH, Rich SS, Gruber C, Franks PW. ADA/EASD Precision Medicine in Diabetes Initiative: An International Perspective and Future Vision for Precision Medicine in Diabetes. Diabetes Care. 2022 Feb 1;45(2):261-266. doi: 10.2337/dc21-2216. PMID: 35050364; PMCID: PMC8914425.

7. Baldi P, Hatfield GW. DNA Microarrays and Gene Expression: From Experiments to Data Analysisand Modeling. Cambridge University Press. 2002.

8.Altiparmak F, Ferhatosmanoglu H, Erdal S, Trost DC. Information mining over heterogeneous and high-dimensional time-series data in clinical trials databases. IEEE Trans Inf Technol Biomed. 2006 Apr;10(2):254-63. doi: 10.1109/titb.2005.859885. PMID: 16617614.

9. Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. Nat Rev Genet. 2012 Jul 18;13(8):552-64. doi: 10.1038/nrg3244. PMID: 22805708.

10. Marin D, Tang M, Ayed IB, Boykov Y. Kernel Clustering: Density Biases and Solutions. IEEE Trans Pattern Anal Mach Intell. 2019 Jan;41(1):136-147. doi: 10.1109/TPAMI.2017.2780166. Epub 2017 Dec 6. PMID: 29990278.

11. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media. 2009.

12. Ernst J, Nau GJ, Bar-Joseph Z. Clustering short time series gene expression data. Bioinformatics. 2005 Jun;21 Suppl 1:i159-68. doi: 10.1093/bioinformatics/bti1022. PMID: 15961453.

13. Chira C, Sedano J, Camara M, Prieto C, Villar JR, Corchado E. A cluster merging method for time series microarray with production values. Int J Neural Syst. 2014 Sep;24(6):1450018. doi: 10.1142/S012906571450018X. Epub 2014 Jul 24. PMID: 25081426.

14. Cinar O, Ilk O, Iyigun C. Clustering of short time-course gene expression data with dissimilar replicates. Annals of Operations Research. 2018;263(1-2):405-428.

15. Coffey N, Hinde J, Holian E. Clustering longitudinal profiles using p-splines and mixed effects models applied to time-course gene expression data. Computational Statistics & Data Analysis. 2014;71:14-29.

16. Futschik ME, Carlisle B. Noise-robust soft clustering of gene expression time-course data. J Bioinform Comput Biol. 2005 Aug;3(4):965-88. doi: 10.1142/s0219720005001375. PMID: 16078370.

17. Hestilow TJ, Huang Y. Clustering of gene expression data based on shape similarity. EURASIP J Bioinform Syst Biol. 2009;2009(1):195712. doi: 10.1155/2009/195712. Epub 2009 Apr 23. PMID: 19404484; PMCID: PMC3171421.

18. Son YS, Baek J. A modified correlation coefficient based similarity measure for clustering time-course gene expression data. Pattern Recognition Letters. 2008;29(3):232–242. doi:10.1016/j.patrec.2007.09.015.

19. Chechik G, Oh E, Rando O, Weissman J, Regev A, Koller D. Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network. Nat Biotechnol. 2008 Nov;26(11):1251-9. doi: 10.1038/nbt.1499. PMID: 18953355; PMCID: PMC2651818.

20. Möller-Levet CS, Klawonn F, Cho KH, Yin H, Wolkenhauer O. Clustering of unevenly sampled gene expression. Fuzzy sets and Systems. 2005;152(1):49-66.

21. Iglesias Molli AE, Bergonzi MF, Spalvieri MP, Linari MA, Frechtel GD, Cerrone GE. Relationship between the IL-1β serum concentration, mRNA levels and rs16944 genotype in the hyperglycemic normalization of T2D patients. Sci Rep. 2020 Jun 19;10(1):9985. doi: 10.1038/s41598-020-66751-x. PMID: 32561825; PMCID: PMC7305205.

22. Wierzchoń ST,Kłopotek MA. Modern algorithms of cluster analysis. Springer, 2018.

23. Yu J, Tian Q, Amores J, Sebe N. Toward Robust Distance Metric Analysis for Similarity Estimation. Computer Vision and Pattern Recognition. In IEEE Computer Society Conference. 2006;1:316-322.

24. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, Gordon DJ, Krauss RM, Savage PJ, Smith SC Jr, Spertus JA, Costa F; American Heart Association; National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation. 2005 Oct 25;112(17):2735-52. doi: 10.1161/CIRCULATIONAHA.105.169404. Epub 2005 Sep 12. Erratum in: Circulation. 2005 Oct 25;112(17):e297. Erratum in: Circulation. 2005 Oct 25;112(17):e298. PMID: 16157765.

25. Ostertagova E, Ostertag O, Kováč J. Methodology and application of the Kruskal-Wallis test. Applied mechanics and materials. 2014;611115-120.

26. Wilcoxon F. Probability tables for individual comparisons by ranking methods. Biometrics. 1947 Sep;3(3):119-22. PMID: 18903631.

27. Milligan G, Cooper M. A study of standardization of variables in cluster analysis. Journal of Classification. 1988;5(2):181-204.

28. Wesolowska-Andersen A, Brorsson CA, Bizzotto R, Mari A, Tura A, Koivula R, Mahajan A, Vinuela A, Tajes JF, Sharma S, Haid M, Prehn C, Artati A, Hong MG, Musholt PB, Kurbasic A, De Masi F, Tsirigos K, Pedersen HK, Gudmundsdottir V, Thomas CE, Banasik K, Jennison C, Jones A, Kennedy G, Bell J, Thomas L, Frost G, Thomsen H, Allin K, Hansen TH, Vestergaard H, Hansen T, Rutters F, Elders P, t'Hart L, Bonnefond A, Canouil M, Brage S, Kokkola T, Heggie A, McEvoy D, Hattersley A, McDonald T, Teare H, Ridderstrale M, Walker M, Forgie I, Giordano GN, Froguel P, Pavo I, Ruetten H, Pedersen O, Dermitzakis E, Franks PW, Schwenk JM, Adamski J, Pearson E, McCarthy MI, Brunak S; IMI DIRECT Consortium. Four groups of type 2 diabetes contribute to the etiological and clinical heterogeneity in newly diagnosed individuals: An IMI DIRECT study. Cell Rep Med. 2022 Jan 4;3(1):100477. doi: 10.1016/j.xcrm.2021.100477. PMID: 35106505; PMCID: PMC8784706.

29. Baranzini SE, Mousavi P, Rio J, Caillier SJ, Stillman A, Villoslada P, Wyatt MM, Comabella M, Greller LD, Somogyi R, Montalban X, Oksenberg JR. Transcription-based prediction of response to IFNbeta using supervised computational methods. PLoS Biol. 2005 Jan;3(1):e2. doi: 10.1371/journal.pbio.0030002. Epub 2004 Dec 28. PMID: 15630474; PMCID: PMC539058.

30. Borgwardt KM, Vishwanathan SV, Kriegel HP. Class prediction from time series gene expression profiles using dynamical systems kernels. Pac Symp Biocomput. 2006:547-58. PMID: 17094268.

31. Costa IG, Schönhuth A, Hafemeister C, Schliep A. Constrained mixture estimation for analysis and robust classification of clinical time series. Bioinformatics. 2009 Jun 15;25(12):i6-14. doi: 10.1093/bioinformatics/btp222. PMID: 19478017; PMCID: PMC2687976.

32. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, Miller-Graziano C, Moldawer LL, Mindrinos MN, Davis RW, Tompkins RG, Lowry SF; Inflamm and Host Response to Injury Large Scale Collab. Res. Program. A network-based analysis of systemic inflammation in humans. Nature. 2005 Oct 13;437(7061):1032-7. doi: 10.1038/nature03985. Epub 2005 Aug 31. Erratum in: Nature. 2005 Dec 1;438(7068):696. PMID: 16136080.

33. den Teuling NGP, Pauws SC, van den Heuvel ER. A comparison of methods for clustering longitudinal data with slowly changing trends. Communications in Statistics: Simulation and Computation. 2023;52(3):621-648.

34. Aiming for equitable precision medicine in diabetes. Nat Med. 2022 Nov;28(11):2223. doi: 10.1038/s41591-022-02105-6. PMID: 36333401.

35. Ben Ayed A, Ben Halima M, Alimi A. Survey on clustering methods: Towards fuzzy clustering for big data. Conference: International Conference on Computational Intelligence in Security for Information Systems. 2015;9. DOI:10.1007/978-3-319-47364-2_55

Best Practices for Effective Management of Point of Care Testing

Anil K Chokkalla[1,2*] Brandy D Recio[2*] and Sridevi Devaraj[1,2*]

[1] *Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX, USA*
[2] *Department of Pathology, Texas Children's Hospital, Houston, TX, USA*

Article Info

Author of correspondence:
Sridevi Devaraj, PhD, DABCC, FRSC, FAACC
Professor of Pathology & Immunology, Baylor College of Medicine;
E-mail: sxdevara@texaschildrens.org;
Address:
Medical Director, Clinical Chemistry and POCT, Texas Children's Hospital
6621 Fannin Street, West Tower, Pathology BB110.06.
Houston TX 77030, USA.

Keywords

Point-of-care tesing; laboratory management; quality management; clinical collaboration

Abstract

With the recent COVID-19 pandemic, point-of-care testing has gained tremendous attention, particularly in acute care settings. The point-of-care testing landscape is rapidly expanding and being contemplated for any crucial test with a central laboratory turnaround time >25% of the clinical decision time. A typical point-of-care testing program within a large hospital system encompasses a multitude of operators utilizing a wide range of devices across multiple testing sites. Thus, managing a large point-of-care testing network remains a daunting task with challenges related to staffing, standardization, quality management, training and competency assessment, and data management. This review will focus on understanding the general organization as well as the roles and responsibilities of various point-of-care testing stakeholders in addressing these challenges. More importantly, it will discuss the strategies and best practices for effective point-of-care testing management based on consensus recommendations from professional societies as well as our experience at Texas Childrens Hospital.

1. Introduction

Point-of-care testing (POCT) refers to testing conducted at or near the patient's bedside, providing rapid results to aid in immediate patient care management. A typical POCT formulary for a large network may include glucose, hemoglobin A1c, ketones, occult blood, blood gases, electrolytes, creatinine, urea, drugs of abuse, pH, human chorionic gonadotropin (hCG), coagulation and infectious disease testing for influenza, streptococcus, HIV, and mononucleosis (1). POCT devices are incorporated with biosensors, microfluidics and lateral flow immunoassays and are capable of providing both qualitative (e.g., urine drug screen) and quantitative results (e.g., glucometer) (1). A majority of POCT devices fall under the waived complexity category under the US Clinical and Laboratory Improvement Amendments of 1988 (CLIA), and do not require performance verification before patient testing (2). However, certain POCTs, such as blood gas instruments, are classified as moderately complex and require method verification for precision, accuracy, reference range and analytical measurement range before implementation (2).

POCT is performed outside the central laboratory, in settings like emergency rooms, intensive care units, physician's offices, pharmacies, ambulances, and outpatient clinics, by medical staff with minimal training in laboratory testing. Moreover, medical staff are constantly distracted by other tasks like drug administration, patient monitoring and data recording. Hence, major challenges in POCT are related to quality assurance and regulatory compliance. Despite these limitations, the POCT menu continues to expand, particularly post-pandemic, and it is projected that the global market value will exceed $44.6 billion by 2025, with an estimated annual compound growth rate of 9% (2). Undoubtedly, the advantages of swift traversal within the hospital system and circumventing central laboratory testing far outweigh the limitations posed by POCT in clinical decision-making. Managing the POCT program presents unique challenges, as it involves numerous operators utilizing a wide range of devices across multiple sites within the hospital system. Since operators may lack in-depth knowledge regarding troubleshooting laboratory variations and identifying potential sources of errors during the pre-analytical, analytical, and post-analytical phases of testing, it becomes vital to establish a robust organizational support system spearheaded by laboratory medicine professionals (3). This ensures that expertise and guidance are available to address complex issues and optimize the quality and reliability of POCT results.

2. POCT management:

A large POCT network refers to a system of interconnected testing sites within a healthcare organization managed centrally. For example, in our large pediatric institution, organized POCT services are delivered across multiple units spanning various hospitals and specialty care centers situated in nine geographical locations within the Houston area. With the continuous expansion of the POCT catalog, especially with the introduction of novel molecular assays for infectious diseases, it becomes imperative to continually monitor and evaluate POCT policies and procedures. The following sections will explore the challenges associated with managing a multi-center POCT network. Key considerations in areas such as staffing, standardization, quality management, training and competency assessment, data management, and continuing education will be reviewed (Figure 1).

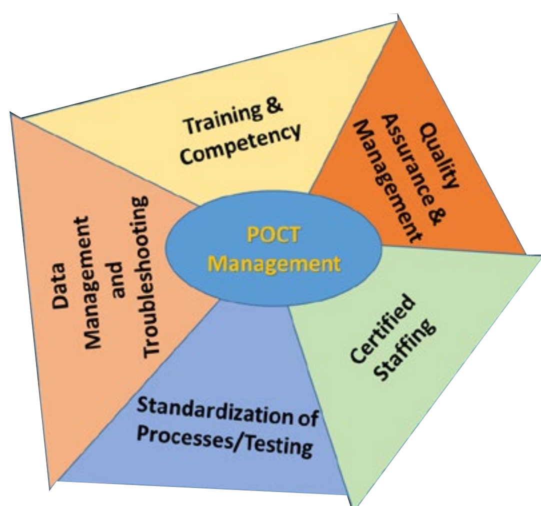


Figure 1: Pinwheel shows key components related to point-of-care testing management.

2.1. Staffing:

An interdisciplinary team with well-defined roles and responsibilities is the key to successful POCT management, particularly in larger networks (4). The key stakeholders of POCT management include the director, manager, co-ordinators, and POCT committee comprising laboratory, administrative and clinical representatives. These levels are dependent on the volume and complexity of testing and the number of testing sites within the healthcare organization. While the POCT director provides expert advice on the overall operations and implementation, the POCT manager and co-ordinators lead the day-to-day activities. The key duties of POCT co-ordinators (POCC) may include training medical staff, installing and validating POCT assays, troubleshooting spurious results and instrument failures, writing procedures, preparing worksheets and logbooks, maintaining and disseminating reagents and test strips to end-users, and conducting audits and competency reviews (5). Hence, POCCs require a wealth of knowledge on clinical workflows, interfaces, test environments, and regulatory requirements to act as technical consultants for medical staff. There are no mandatory federal requirements for certification or a higher degree other than a high school diploma for performing waived testing. However, if the POCC is acting as a technical consultant for non-waived tests the College of American Pathologists (CAP) requirement is to have a Bachelor of Science (BS) in chemical, physical, biological, clinical laboratory science or medical technology and a 2 years experience with POCT. A survey involving 98 POCCs across the US revealed that ~87% are certified in medical laboratory science (5). Medical laboratory scientists with over 3 years of experience in the central laboratory are ideal for the POCC role. Importantly, the number of POCCs in a program should be proportional to the number of testing centers and menus. This is quite challenging due to ongoing staffing shortages, high staff turnover and contract-based staffing, particularly in post-pandemic era healthcare. A recent study has highlighted several strategies, such as improving visibility, promoting diversity and inclusion, and early recruitment and retention of laboratory medicine professionals to overcome this problem (6). Furthermore, clinical collaboration and communication are critical to efficiently managing POCT at multiple sites with limited staff.

2.2. Standardization:

Currently, many POCT platforms are available for any given analyte of interest. For instance, there are 137 FDA-waived POCT assays specifically designed for pregnancy testing using urinary hCG, as reported by the CLIA test complexity database. Each of these assays comes with manufacturer-specific instructions regarding quality control and specimen handling procedures. Consequently, it becomes essential to standardize POCT by utilizing the same assays and instrumentation across all testing sites within a hospital system. The advantages of POCT standardization are manifold (4). First, it simplifies the management process by allowing a single policy to be shared among operators for training and competency purposes. Additionally, it facilitates quick troubleshooting via running parallels for instrument comparisons across different sites. Moreover, it enables monetary savings by cutting down on the vendor fees associated with instrument maintenance and reagent expenses due to increased test volume. Standardization also allows POCCs to maintain loaner devices to be deployed across the hospital system as needed. POCT system selection is another major consideration that requires the evaluation of operational and economic parameters. As per the Association for Diagnostics and Laboratory Medicine guidance document, the key characteristics for review should include device footprint, connectivity, electrical, waste, training and maintenance requirements, implementation and operational cost comparison (4).

2.3. Quality management:

For a large POCT network, having a homogenous quality management system (QMS) is desired. Waived tests are performed according to the manufacturer's instructions provided in the package insert. Moderately complex tests require two levels of quality control each day before patient testing or developing an individualized quality control plan (IQCP) after conducting a risk analysis, particularly for infectious disease and coagulation testing (4). A QMS plan should provide guidelines to conduct: 1. monthly site audits, 2. reporting questionable test results, 3. corrective/preventive actions, 5. proficiency testing, and 6. POCT operator training and competency (7). While conducting monthly site audits, POCCs should assess the internal quality control and patient testing logs, instrument maintenance logs, temperature and humidity logs, proper storage and expiry of the regents and test strips, and the testing area. Deficiencies from the site audit should be documented and communicated to the testing unit leadership for corrective or preventive measures, which may include retraining, direct observation, and removal of testing privileges (7). Moreover, it is crucial to review risk management strategies such as error reporting systems, root cause analysis, and corrective actions in a center or location-specific manner, considering the distinctive challenges associated with each particular center. Well-established tools such as Ishikawa diagrams and failure modes and effect analysis (FMEA) can be adapted in this context (8). Quality Indicators (QI) are a set of benchmarks that form an integral part of QMS for continuous quality improvement. Currently, 57 QIs to evaluate the key processes within pre-analytical, analytical and post-analytical phases are recommended for central laboratory testing by the International Federation of Clinical Chemistry and Laboratory Medicine (9). High variability in the testing environment is a major barrier to developing POCT-specific QIs. Based on a national survey of quality practices across 12 large POCT networks within 8 Canadian provinces, the following QIs were evaluated: patient identification, instrument lockouts, sample collection errors, failed quality control, transcription errors, result amendment, turnaround times, training and competency

documentation, and instrument errors (10). POCT center-specific performance assessment of these QIs is necessary for a large POCT network. Such an approach is exemplified by a recent study that performed a process mapping and risk assessment to establish QIs for point-of-care glucose testing (11). Proficiency testing (PT) facilitates quality assessment of the total testing process, including pre-analytical, analytical, and post-analytical phases, hence serving as a surrogate marker for POCT performance. Although federal guidelines (CLIA) do not require PT for waived tests, POCT centers accredited by CAP should enroll in Health and Human Services (HHS)-approved external PT programs or PT verification programs for regulated analytes (4). Alternatively, PT assessment by split sample testing should be considered at least semi-annually. PT samples should be processed in the same manner as patient samples by a blind operator, and documentation should be retained for a period of 2 years. All acceptable PT results should be reviewed for bias and trends, whereas unacceptable PT results (score <80%) requires investigation to identify sources of error for corrective actions (12). A POCT site-specific manager or educator serves as a liaison and delegates the activities like scheduling and submitting PT results. Furthermore, it is recommended to distribute PT, and quality cross-checks to engage staff from all shifts and departments performing POCT.

2.4 Training and competency assessment:

POCT operators originate from diverse educational backgrounds depending on the testing site, such as nursing, pharmacy, emergency medical technicians and medical assistants. Hence, all qualified POCT users must complete comprehensive training, orientation and competency assessment of each test method prior to patient testing. Online training modules should cover theoretical aspects of test methods, instrument maintenance, calibration, performing internal quality control, patient safety, sample interferences and results reporting. These operators should also undergo in-unit training and competency assessment once every year (12). Since involving every operator in this process is cumbersome, it is crucial to designate one superuser for every ten operators in the testing unit who is knowledgeable about quality control, quality assurance and safety regulations. Additionally, these superusers serve as the primary contact for communication between medical staff and POCT leadership.

POCT competency assessment essentially involves 6 elements as per the CAP checklist: 1. direct observation of the routine patient testing, including patient identification and preparation, and specimen collection, handling, processing and testing, 2. recording and reporting of test results, including critical results, 3. review of quality control records, proficiency testing results, and preventive maintenance records, 4. direct observation of the performance of instrument maintenance and function checks, 5. assessment of test performance through testing previously analyzed specimens like internal blind testing samples or external proficiency testing samples, and 6. evaluation of problem-solving skills (12). For a large POCT network, an online-based learning management system (LMS) may be utilized for the electronic recording of training and competency for each operator.

2.5. Data management:

Poor connectivity is a major barrier to POCT adoption, particularly in remote subspecialty care centers within a large hospital system. Currently, not all POCT devices are interfaced with patient electronic medical records, requiring manual entry of the results by the medical staff. This triggers errors, delays and duplicate testing, further increasing the costs. It is crucial for POCT manufacturers to follow guidelines from the Clinical and Laboratory Standards Institute (POCT01) that enable easy information exchange between POCT devices, electronic medical records, and laboratory information systems (13). Nowadays, POCT middleware providers such as TELCOR are widely used among large hospital systems due to the convenience of data integration, data validation, result routing, and rules or alert set-up. Data management functionalities offered by these middlewares improve quality by preventing untrained operators from accessing the device. This is achieved by mandating the entry of a valid operator identification number before initiating any testing. Additionally, data management lockouts ensure that regular quality control is conducted and prohibit patient testing if quality control measures have not been performed or if the controls exceed an acceptable target range. By implementing these POCT data management strategies, the likelihood of incorrect results reporting from common operational mistakes is significantly reduced (14).

2.6. Continuing education:

A plethora of educational opportunities exist to stay abreast of the evolving technological and regulatory landscape of POCT. POCCs are encouraged to enroll in the point-of-care specialist certification program offered by the Critical and POCT (CPOCT) division of the Association for Diagnostics and Laboratory Medicine. The online course comprises eight learning modules that encompass a range of topics such as instrument selection and validation, quality management, regulations, policies and procedures, connectivity and information technology, education and training, administration, and communication (15). Individuals successfully finishing this program are credentialed as Certified Point-of-Care Testing Professionals (CPPs). More than 1733 laboratory professionals currently hold this certification. However, this involves an enrollment fee. There are several free resources for continuing education credits in POCT. Some of the examples include a one-hour on-demand course offered by the Centers for Disease Control and Prevention (CDC) on CLIA-waived testing requirements, a youtube channel hosted by POCT Center for Teaching and Research at the University of California Davis and the webinars offered by the Whitehat Communications (16). POCT leaders should strive to incorporate these materials and host case conferences as a part of medical staff training and continuing education.

3. Conclusions

In a nutshell, managing a large POCT network requires a multidimensional approach that addresses challenges related to staffing, standardization, quality management, training and competency assessment, data management, and continuing education. As reviewed by our group previously (17), by adhering to evidence-based guidelines and recommendations put forth by professional societies such as the Association for Diagnostics and Laboratory Medicine and the International Federation of Clinical Chemistry and Laboratory Medicine, healthcare organizations can navigate these challenges successfully. Delineation of POCT responsibilities among various stakeholders, such as POCT leadership (director, manager, and test site leaders), POCCs and POCT operators, is key to a successful POCT program. By establishing a robust organizational support system and implementing effective strategies, hospitals can optimize the quality, reliability, and efficiency of their POCT programs, ultimately benefiting patient care and outcomes.

Declarations of interest

None.

CRediT author statement

Anil K Chokkalla: Conceptualization, Investigation, Writing - Original Draft. Brandy Reico: Conceptualization, Investigation, Writing – Review & Editing. Sridevi Devaraj: Conceptualization, Investigation, Writing – Review & Editing, Supervision.

Acknowledgments

Anil K Chokkalla was supported by Ching Nan Ou Fellowship Endowment in Clinical Chemistry.

References

1. Nichols JH. Chapter 19 - Point-of-care testing. In: Clarke W, Marzinke MA, editors. Contemporary Practice in Clinical Chemistry (Fourth Edition): Academic Press; 2020. p. 323-36.

2. Nichols JH. Utilizing Point-of-Care Testing to Optimize Patient Care. Ejifcc. 2021;32(2):140-4.

3. Warade J. Organization of the POCT Unit. Ejifcc. 2015;26(2):125-32.

4. Nichols JH, Alter D, Chen Y, Isbell TS, Jacobs E, Moore N, et al. AACC Guidance Document on Management of Point-of-Care Testing. J Appl Lab Med. 2020;5(4):762-87.

5. Gledhill TR, White SK, Lewis JE, Schmidt RL. A Profile of Point of Care Coordinators: Roles, Responsibilities and Attitudes. Laboratory Medicine. 2019;50(3):e50-e5.

6. Halstead DC, Sautter RL. A Literature Review on How We Can Address Medical Laboratory Scientist Staffing Shortages. Lab Med. 2023;54(1):e31-e6.

7. Khan AI, Pratumvinit B, Jacobs E, Kost GJ, Kary H, Balla J, et al. Point-of-care testing performed by healthcare professionals outside the hospital setting: consensus based recommendations from the IFCC Committee on Point-of-Care Testing (IFCC C-POCT). Clin Chem Lab Med. 2023.

8. James HN. Risk Management for Point-of-Care Testing. Ejifcc. 2014;25(2):154-61.

9. Sciacovelli L, Lippi G, Sumarac Z, West J, Garcia Del Pino Castro I, Furtado Vieira K, et al. Quality Indicators in Laboratory Medicine: the status of the progress of IFCC Working Group "Laboratory Errors and Patient Safety" project. Clin Chem Lab Med. 2017;55(3):348-57.

10. Brun M, Füzéry AK, Henschke B, Rozak K, Venner AA. Identifying sources of error and selecting quality indicators for point of care testing. Pract Lab Med. 2021;25:e00216.

11. Shaw JLV, Arnoldo S, Beach L, Bouhtiauy I, Brinc D, Brun M, et al. Establishing quality indicators for point of care glucose testing: recommendations from the Canadian Society for Clinical Chemists Point of Care Testing and Quality Indicators Special Interest Groups. Clin Chem Lab Med. 2023;61(7):1280-7.

12. Yenice S. Training and Competency Strategies for Point-of-Care Testing. Ejifcc. 2021;32(2):167-78.

13. Ortiz DA, Loeffelholz MJ. Practical Challenges of Point-of-Care Testing. Clin Lab Med. 2023;43(2):155-65.

14. Nichols JH. Quality in point-of-care testing: taking POC to the next level. Accreditation and Quality Assurance. 2006;11(6):273-7.

15. Isbell TS, Colwell E, Frank EL, Karon BS, Luzzi V, Wyer LA. Professional Certification in Point-of-Care Testing. Ejifcc. 2021;32(3):303-10.

16. Kost GJ, Zadran A, Zadran L, Ventura I. Point-Of-Care Testing Curriculum and Accreditation for Public Health-Enabling Preparedness, Response, and Higher Standards of Care at Points of Need. Front Public Health. 2018;6:385.

17. Chokkalla AK, Devaraj S. ASCP Educational Commentary on POCT management for CE credits. ASCP; American Proficiency Institute – 2023 2nd Test Event, pg 1-8.

Monoclonal Light Chains with alpha 2 mobility on Serum Protein Electrophoresis

S. Danalakshmi, Department of Biochemistry, Apollo hospitals, Chennai, Tamil Nadu, India.

Article Info

Author of correspondence:
Dr. S. Danalakshmi, Department of Biochemistry;
E-mail: drdana72@gmail.com;
Address:
Apollo hospitals, 21, Greams lane, Off Greams Road, Chennai, Tamil Nadu- 600006, India.

Keywords
multiple myeloma, monoclonal gammopathy, capillary zone electrophoresis, paraprotein, immunotyping, hypogammaglobinemia.

Abstract

Multiple myeloma (MM) is a neoplasm characterized by malignant proliferation of plasma cells that produce excessive quantities of a single type of immunoglobulin (Ig) called as monoclonal immunoglobulin or M-protein or paraprotein. M-protein produced can be either an intact antibody with both heavy and light-chain components or only light chains or rarely only heavy chains. Presence of M-protein in serum protein electrophoresis (PEP) is useful in the diagnosis, prognosis, and treatment of MM and other plasma cell dyscrasias. These M-proteins are identified commonly in beta and gamma regions and very rarely in alpha 2 region, appearing as a narrow band in agarose electrophoresis or as a sharp symmetric spike (M-spike) or peak in capillary zone electrophoresis. Here, we present an unusual case of monoclonal light chains producing two M- spikes in the alpha 2 globulin region in capillary zone electrophoresis.

Introduction

Multiple myeloma (MM) is a neoplasm of B cell lineage which is characterized by excessive proliferation of abnormal plasma cells. These abnormal plasma cells secrete a specific immunoglobulin type referred to as monoclonal immunoglobulin (Ig) that leads to a condition called monoclonal gammopathy. Serum protein electrophoresis is often used as a preliminary test in suspected cases of multiple myeloma to detect the presence of M-protein, which when present produces a narrow band in agarose electrophoresis or a sharp symmetric spike (M-spike) or peak in capillary zone electrophoresis (1). Serum protein electrophoresis is a technique used to separate serum proteins into different fractions depending on their size and the charge they carry in a buffer medium under the influence of an electric field (2). Zone electrophoresis refers to the migration of charged molecules of proteins in a porous support medium like agarose, cellulose acetate, capillaries, etc., such that each protein zone is sharply separated from the neighbouring zone by a protein free area (3).

Capillary zone electrophoresis (CZE) is a technique in which the buffer filled capillary tube is used as the support medium. By the application of high voltage, serum proteins are separated into 6 fractions viz. albumin, alpha 1(α-1), alpha 2(α-2), beta 1(β-1), beta 2 (β-2), and gamma (γ) globulin regions. The paraproteins (M-proteins), when present, produce a sharp spike in any of the globulin regions in CZE (2, 4). M-protein spike is generally observed in gamma or

beta region and very rarely seen in alpha 2 globulin region (5). Here, we present a case of monoclonal light chain producing M-spikes in alpha 2 globulin region on the protein electrophoresis performed using CZE technique.

Case details

A 64 year-old male presented to the orthopedic out-patient department with complaints of lower backache. He was evaluated at home town where the lumbosacral spine X-ray showed lumbar spondylosis. MRI-lumbosacral spine showed at L4-L5 and L5 level a broad posterior disc osteophyte complex causing canal stenosis with bilateral facetal hypertrophy contributing to lateral recess stenosis and moderate compression of bilateral traversing nerve roots. Abnormal marrow signal intensity in D10 vertebra with bulging of the posterior cortex with epidural soft tissue thickening causing canal stenosis and significant compression of the traversing nerve roots was also noticed. Based on the MRI findings a serum protein electrophoresis was ordered. Serum protein electrophoresis by capillary zone electrophoresis performed in minicap (Sebia) automated analyzer showed, hypogammaglobinemia with two small peaks in alpha 2 region (Figure 1). The laboratory, based on the findings of the protein electrophoresis, suggested to perform serum free light chain assay and serum immunotyping to confirm that the spikes in alpha 2 region in PEP were due to paraprotein. The results of the investigations done are as follows (serum free light chains were performed using Binding site- free lite kit in Optilite analyzer):

- Serum free kappa : 17,052 mg/L (reference interval : 3.3 – 19.4 mg/L)
- Serum free lambda : 11.69 mg/L (reference interval : 5.7 – 26.3 mg/L)
- kappa/lambda ratio : 1458.7 (reference interval : 0.26 -1.65)

Serum immunotyping, performed in minicap (Sebia) automated analyzer, showed both the peaks in alpha 2 region to be kappa free light chains (Figure 2) .The patient was further investigated with PET CT whole body, which showed multiple ill-defined lytic bony lesions in axial skeleton with few lesions showing low-grade FDG uptake. Bone marrow aspiration cytology showed diffuse plasmacytosis (23% plasma cells) composed of mature, few immature plasmablasts, binucleate, occasional trinucleate and quadrinucleate forms consistent with plasma cell myeloma. Bone marrow trephine biopsy revealed a hypercellular marrow with diffuse plasmacytosis consistent with myeloma marrow. The other relevant biochemistry and haematology investigations include, Serum beta-2 microglobulin : 8.54 mg/L (724.34 nmol/L), serum protein : 5.9 g/dL (59 g/L), serum albumin : 4 g/dL (40 g/L), globulin : 1.9 g/dL (19 g/L), serum calcium : 8.2 mg/dL (2.05mmol/L), serum creatinine: 1.1 mg/dL (97.24 micromole/L), serum urea: 43 mg/dL (7.16 mmol/L) and Hemoglobin: 6.5 gm/L. Patient was diagnosed to have multiple myeloma and was started on Chemotherapy using VRD regimen with bortezomib, lenalidamide and dexamethasone, planned for 6 cycles. Each cycle had bortezomib given intravenously on days 1, 8, 15 and 22, dexamethasone given intravenously on days 1, 8, 15, and 22. Lenalidamide was given orally once daily from day 1 to day 22. Follow up serum protein electrophoresis after initiation of 1st cycle of chemotherapy showed that the peak in the alpha 2 region had disappeared (Figure 3) with corresponding serum free kappa level decreasing to 2123 mg/L . Serum immunotyping was negative for kappa free light chains (Figure 4).

Discussion:

Multiple myeloma is the second-most common hematological malignancy and is characterized by excessive proliferation of abnormal plasma cells that secretes monoclonal immunoglobulins (M- protein) (6, 7). Serum protein electrophoresis (PEP) to identify the presence of M- protein and immunofixation electrophoresis (SIFE) or immunotyping to identify the type of monoclonal immunoglobulin that is involved in the disease process are commonly used as screening tests in the diagnosis of multiple myeloma. Other investigations that are usually done include quantification of serum free light chain, urine protein electrophoresis and urine immunofixation electrophoresis for Bence Jones protein. Bone marrow aspirate and trephine biopsy are obtained for carrying out cytogenetic, fluorescent in situ hybridization (FISH) and immunophenotyping studies (8).

Relative reduction or increase in M-protein concentration or their disappearance in serum protein electrophoresis, is one of the parameter of the Uniform Response Criteria (IURC) designed by the International Myeloma Working group (IMWG) to guide the management of patients with multiple myeloma (9). Therefore, identification of paraproteins in PEP is a critical diagnostic step to further guide the investigations and design the management in patients with MM. M-proteins produced in multiple myeloma are either an intact immunoglobulin with both heavy and light-chain components or only light chains or rarely only heavy chains. In the intact immunoglobulin type, the heavy chain is from one of the five immunoglobulin classes G, A, M, D or E, while the light chain is either κ (kappa) or λ (lambda)(10). The commonest immunoglobulin class involved in MM is IgG followed by IgA (as intact Immunoglobulins) and only light chains comprises 15% of patients with MM (7). In electrophoretogram, these M-proteins are commonly observed in gamma-globulin region. They are also observed in β-2 and β-1 regions in diminishing order of frequency and rarely seen in α-2 globulin region. Generally monoclonal intact immunoglobulins of type IgA, IgG and IgM are not observed on the α-2 region and are usually observed in γ, β-1, and β-2 regions (10). Normally the proteins that migrate in α-2 fractions include alpha-2 macroglobulin, ceruloplasmin and haptoglobulin. These are acute phase reactants that are raised in inflammatory conditions. Hence a tall peak in α-2 region in CZE may occur in inflammatory disorders and in nephrotic syndrome (due to increased alpha-2 macroglobulin) closely resembling a monoclonal peak. A radio contrast dye used for imaging studies also produces a spike or a split in this region in CZE, mimicking a paraprotein(11). In vitro hemolysis and haptoglobins of different phenotypes may induce a split in the α-2 globulin region raising

a suspicion for M-protein (12). Although the abnormal patterns observed in α-2 region are more common due to the above mentioned conditions yet, those produced by the paraprotein should not be missed. The case in discussion here had two small spikes in α-2 region and was present along with hypogammaglobinemia. This led to the suspicion of paraprotein which was confirmed to be due to monoclonal kappa light chains by serum immunotyping. Presence of two M-spikes in α-2 globulin region due to monoclonal light chains has not been reported so far. Literature search showed that the M- proteins migrating to α-2 region in electrophoresis are few in numbers and such migrations had IgA preponderance (13, 14). For the case in discussion, the two M spikes on the α-2 region were shown to be because of kappa free light chains (Fig : 2). Most of the time, in light chain myeloma, spikes are often not observed in serum protein electrophoresis, except for the presence of hypogammaglobinemia. In our case too, the peak produced by high concentration of monoclonal kappa (17,052 mg/L) in PEP disappeared and serum immunotyping showed absence of monoclonal gammopathy once the concentration of kappa free light chains decreased to 2123 mg/L. The reason for two M-spikes observed in our case can be attributed to polymerization of monoclonal light chains (kappa) that was present at high concentration (17,052 mg/L) in serum. Their alpha 2 mobility can be reasoned out to be due to their low molecular weight compared to intact immunoglobulins (15).

Conclusion

M- protein with alpha 2 mobility though rare can still occur and produce a spike in alpha 2 region in serum protein electrophoresis. These paraprotein spikes in alpha 2 may be overlooked and mistaken for spikes that are usually produced by components due to various other causes. Paraprotein should be suspected and investigated further when abnormal pattern or small spike in alpha 2 region coexists with hypogammaglobinemia. Immunotyping is performed with specific antisera against IgG, IgA, IgM heavy chains and against free and bound Kappa and Lambda light chains. The superimposition of the antisera pattern allows for visualization of the disappearance and / or decrease of monoclonal fraction on the antisera pattern and to indicate a gammopathy.

Authors' contribution

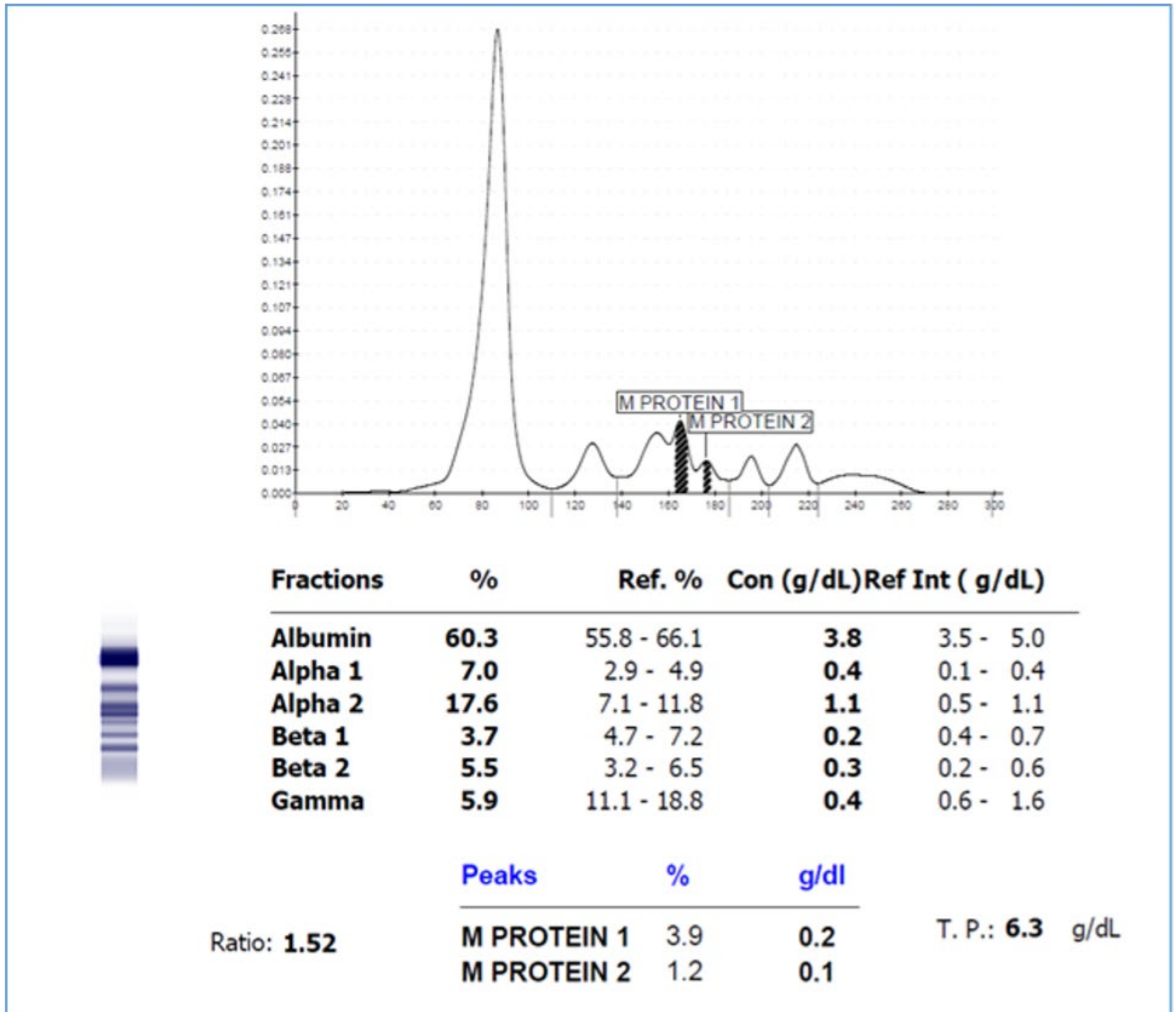
Dr. Danalakshmi S: Conception of the idea, drafting the article, Critical revision and final approval of the version to be published

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.



Fractions	%	Ref. %	Con (g/dL)	Ref Int (g/dL)
Albumin	60.3	55.8 - 66.1	3.8	3.5 - 5.0
Alpha 1	7.0	2.9 - 4.9	0.4	0.1 - 0.4
Alpha 2	17.6	7.1 - 11.8	1.1	0.5 - 1.1
Beta 1	3.7	4.7 - 7.2	0.2	0.4 - 0.7
Beta 2	5.5	3.2 - 6.5	0.3	0.2 - 0.6
Gamma	5.9	11.1 - 18.8	0.4	0.6 - 1.6

Peaks	%	g/dl
M PROTEIN 1	3.9	0.2
M PROTEIN 2	1.2	0.1

Figure 1: Electrophoretogram of capillary zone electrophoresis.
Serum protein electrophoresis, done using Sebia- Minicap, shows two M-protein spikes in alpha 2 region and hypogammaglobinemia

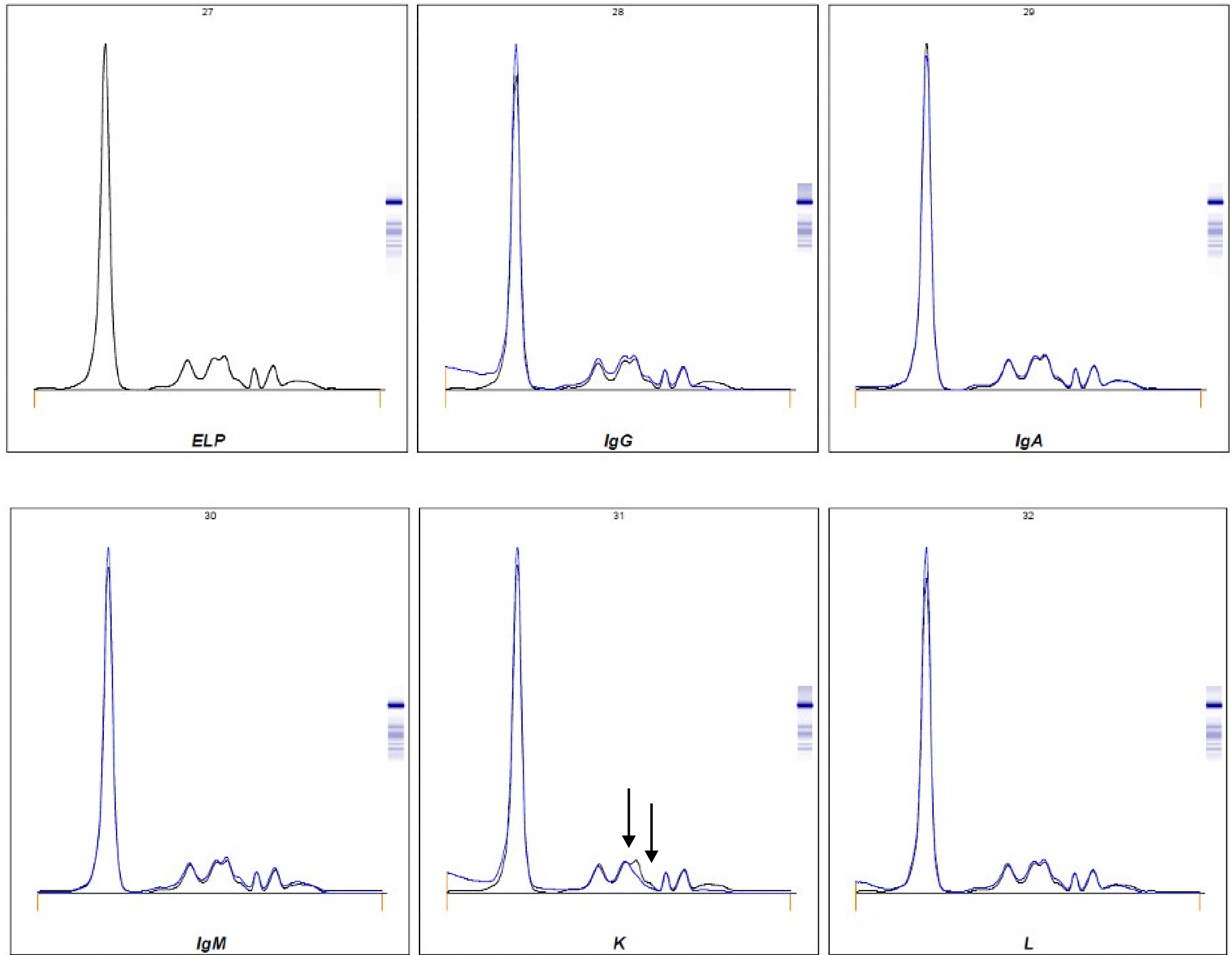
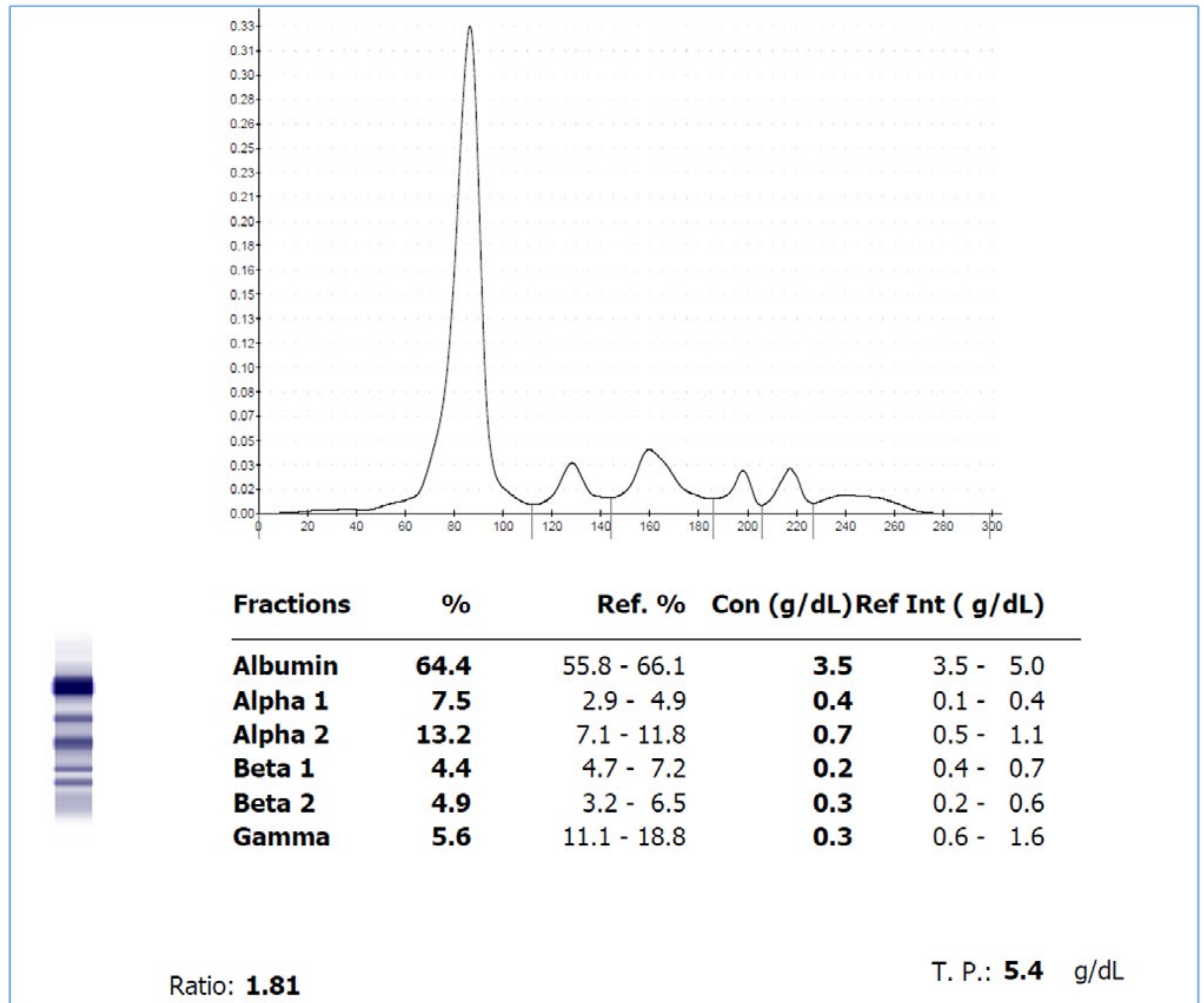


Figure 2: Serum protein immunotyping using Sebia minicap, shows immunosubtraction (two arrow marks) in kappa confirming the spikes to be free kappa light chains (monoclonal Kappa component presenting with disappearance of the kappa antisera pattern as compared to ELP pattern).



Fractions	%	Ref. %	Con (g/dL)	Ref Int (g/dL)
Albumin	**64.4**	55.8 - 66.1	**3.5**	3.5 - 5.0
Alpha 1	**7.5**	2.9 - 4.9	**0.4**	0.1 - 0.4
Alpha 2	**13.2**	7.1 - 11.8	**0.7**	0.5 - 1.1
Beta 1	**4.4**	4.7 - 7.2	**0.2**	0.4 - 0.7
Beta 2	**4.9**	3.2 - 6.5	**0.3**	0.2 - 0.6
Gamma	**5.6**	11.1 - 18.8	**0.3**	0.6 - 1.6

T. P.: **5.4** g/dL

Ratio: **1.81**

Figure 3: Serum protein electrophoresis (done using Sebia- Minicap, capillary electrophoresis) – post initiation of first cycle of chemotherapy, shows absence of M-spikes in alpha 2 globulin region.

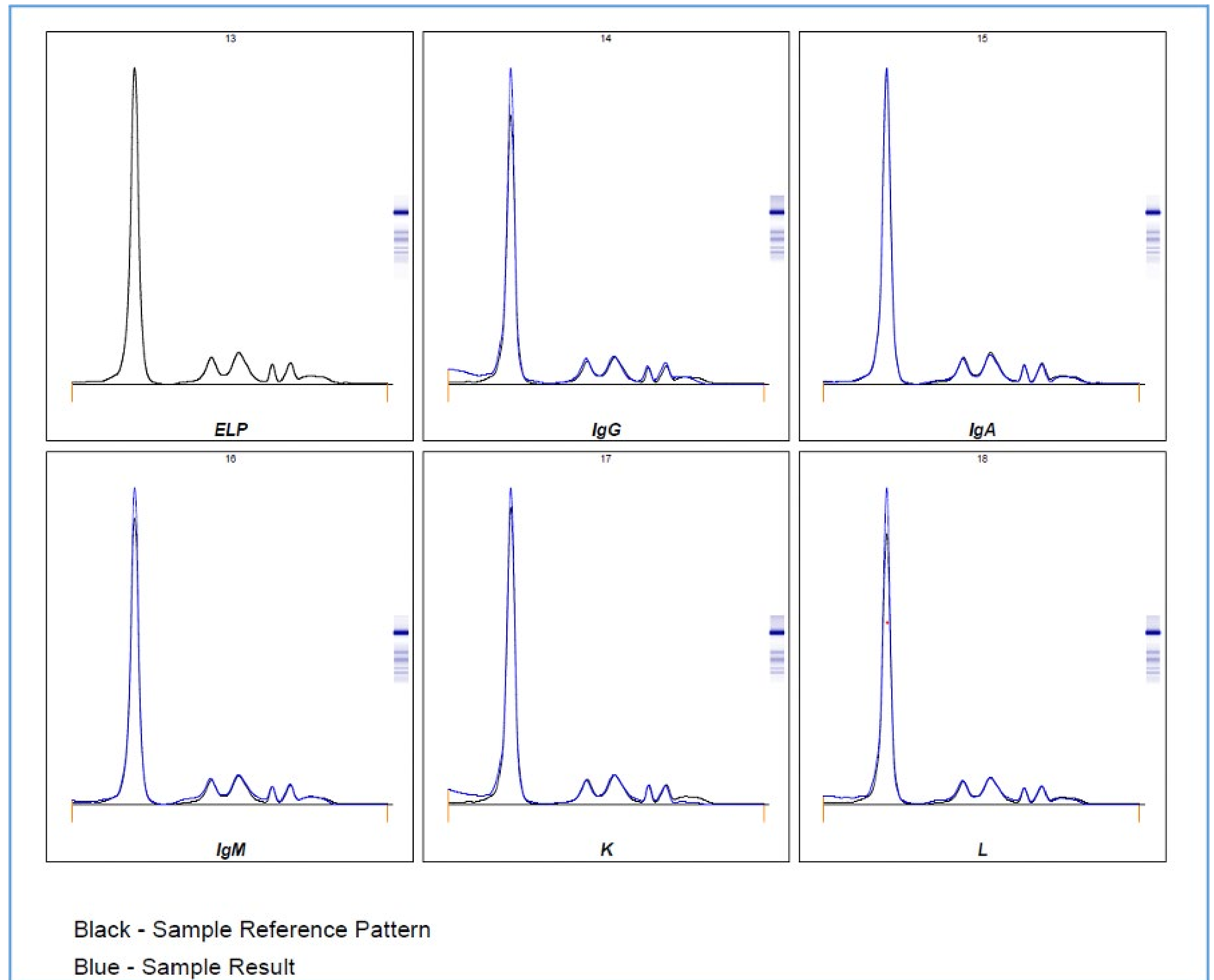


Figure 4: Serum protein immunotyping (post initiation of first cycle of chemotherapy) using Sebia minicap shows absence of monoclonal gammopathy.

References

1. Tripathy S. The role of serum protein electrophoresis in the detection of multiple myeloma: an experience of a corporate hospital. J Clin Diagn Res. 2012 Nov;6(9):1458-61. doi: 10.7860/JCDR/2012/4213.2532. PMID: 23285429; PMCID: PMC3527769.

2. O'Connell TX, Horita TJ, Kasravi B. Understanding and interpreting serum protein electrophoresis. Am Fam Physician. 2005 Jan 1;71(1):105-12. PMID: 15663032

3. Jenkins MA. Serum and urine electrophoresis for detection and identification of monoclonal proteins. Clin Biochem Rev. 2009 Aug;30(3):119-22. PMID: 19841694; PMCID: PMC2755000

4. Yvonne Henskens, John de Winter, Maurits Pekelharing, Gabrielle Ponjee, Detection and identification of monoclonal gammopathies by capillary electrophoresis, Clinical Chemistry, Volume 44, Issue 6, 1 June 1998, Pages 1184–1190, https://doi.org/10.1093/clinchem/44.6.1184

5. Nargis, W. and Ibrahim, M. 2016. A Case of Multiple Myeloma with Unusual Serum Protein Electrophoresis. Pulse. 8, 1 (Jun. 2016), 77–80. DOI:https://doi.org/10.3329/pulse.v8i1.28107.

6. Bora K, Das U, Barman B, Ruram AA. Monoclonal gammopathy with double M-bands: An atypical presentation on serum protein electrophoresis simulating biclonal gammopathy. Indian J Pathol Microbiol. 2017 Oct-Dec;60(4):590-592. doi: 10.4103/IJPM.IJPM_311_17. PMID: 29323083

7. Rafae A, Malik MN, Abu Zar M, Durer S, Durer C. An Overview of Light Chain Multiple Myeloma: Clinical Characteristics and Rarities, Management Strategies, and Disease Monitoring. Cureus. 2018 Aug 15;10(8):e3148. doi: 10.7759/cureus.3148. PMID: 30345204; PMCID: PMC6191009.

8. Gurmukh Singh, MD, PhD, MBA, Serum Free Light Chain Assay and κ/λ Ratio Performance in Patients Without Monoclonal Gammopathies: High False-Positive Rate, American Journal of Clinical Pathology, Volume 146, Issue 2, August 2016, Pages 207–214, https://doi.org/10.1093/ajcp/aqw099

9. Mark, T., Jayabalan, D., Coleman, M., Pearse, R.N., Wang, Y.L., Lent, R., Christos, P.J., Lee, J.W., Agrawal, Y.P., Matthew, S., Ely, S., Mazumdar, M., Cesarman, E., Leonard, J.P., Furman, R.R., Chen-Kiang, S. and Niesvizky, R. (2008), Atypical serum immunofixation patterns frequently emerge in immunomodulatory therapy and are associated with a high degree of response in multiple myeloma. British Journal of Haematology, 143: 654-660. https://doi.org/10.1111/j.1365-2141.2008.07374.x

10. Kyle RA, Rajkumar SV. Plasma cell disorders. In: Goldman L, Ausiello D, editors. Cecil textbook of medicine. Chapter 196. 22. Philadelphia: W. B. Saunders; 2004. pp. 1184–1186

11. Wheeler, Rachel D., et al. "Large Abnormal Peak on Capillary Zone Electrophoresis Due to Contrast Agent." Annals of Clinical Biochemistry, vol. 55, no. 5, Sept. 2018, pp. 608–611, doi:10.1177/0004563217745896.

12. Simon A, Schneider N, Gillery P, Oudart JB. Phénotypes d'haptoglobine : impacts cliniques et biologiques [Clinical and biological features of haptoglobin phenotypes]. Ann Biol Clin (Paris). 2020 Oct 1;78(5):493-498. French. doi: 10.1684/abc.2020.1590. PMID: 33026345.

13. Ajai Chari, Andrew Eisenberger, Michael Pesce,Monoclonal Immunoglobulins of Alpha-2 Mobility on Protein Electrophoresis: A Case Series.,Blood,Volume 110, Issue 11,2007,Page 4753,ISSN 0006-4971,https://doi.org/10.1182/blood.V110.11.4753.4753

14. Bakta IM, Sutarka IN. Biclonal gammopathy in multiple myeloma: a case report. Gan To Kagaku Ryoho. 2000 May;27 Suppl 2:544-8. PMID: 10895208

15. Ellidag HY, Eren E, Aydin O, Yilmaz N. Monoclonal Light Chains Can Remain Unnoticed in Protein Electrophoresis. Indian J Clin Biochem. 2015 Jul;30(3):363-4. doi: 10.1007/s12291-015-0484-2. Epub 2015 Feb 28. PMID: 26089627; PMCID: PMC4469047.

Coffee Colored Serum, Adverse Reaction of Eltrombopag

Carlos Rodríguez Rojas[1*], Celia Juez Santamaría[2*], Oscar David Pons Belda[1*], Emilia Moreno Noguero[1*]

[1]*Clinical Analyses Department, Hospital Can Misses, Eivissa, Spain*
[2]*Pharmacy Department, Hospital Can Misses, Eivissa, Spain*

Article Info

Author of correspondence:
Carlos Rodríguez Rojas, Clinical Analyses Department;
E-mail: carlosrodriguezr91@hotmail.com;
Tel.: +34-697-655476
Address:
Hospital Can Misses, Carrer de Corona, s/n, 07800 Eivissa, Illes Balears, 07800 Spain

Keywords

Brown serum, Eltrombopag, cutaneous hyperpigmentation, clinical laboratory.

Abstract

Serum index and macroscopic characteristics of samples can give valuable information and should be interpreted as a result. Following centrifugation of the sample, on gross inspection it was observed that the serum had a brown color. After ruling out the main causes that can cause a brown coloration, such as intravascular hemolysis or high concentrations of methemoglobin, it was noted that the patient was receiving a high-dose of Eltrombopag therapy. Eltrombopag is a non-peptide thrombopoietin receptor agonist approved for the treatment of severe aplastic anemia (SAA). The drug in solution has a brown color and at high concentrations it is capable of changing the color of the serum and may have different effects in different assays of laboratory. This article describes the case of a patient with brown serum due to the consumption of high doses of Eltrombopag that started to cause cutaneous hyperpigmentation.

Introduction

A fundamental aim in the daily routine of a clinical laboratory in relation to patient safety is to communicate precise and accurate analytical results. Otherwise, serious errors in clinical interpretation would be made. Serum index results are very useful for monitoring the degree of potential interference due to lipemia, hemolysis and jaundice. Nevertheless, in some cases the clinical laboratory must go further and be able to detect certain hidden pathologies using these indices, such as dyslipidemia, hemolysis in vivo, liver disease [1,2] or, as in the present case, adverse drug effects. Endogenous and exogenous constituents in the sample matrix can affect laboratory tests. Some of these potentially interfering factors can be recognized in the pre-analytical phase by colorimetric appearance, turbidity and viscosity, whereas others are detected only by direct analysis. The following case is a description of a patient who presents brown serum due to the consumption of high doses of Eltrombopag that began to cause hyperpigmentation of the skin.

Clinical-diagnostic case

We report a 78-year-old man diagnosed with severe aplastic anemia (SAA) on treatment with cyclosporine 125 mg/12h + eltrombopag 150 mg/24h + allopurinol 300 mg/24h. During the follow-up carried out by the Haematology service, a blood test including a complete blood count and basic biochemistry was obtained. After processing the sample, the hemolysis index was normal, however we detected a high rate of jaundice and lipemia index (Table 1). Due to this result, the hepatic study was included by adding bilirubin determination and the ions were re-analysed by direct potentiometry to correct the negative interference produced in the lipemic samples by indirect potentiometry. After carrying out the corresponding analysis, no increase in total bilirubin was observed in accordance with the patient's jaundiced index. Furthermore, the serum showed a dark brown colour and not a cloudy appearance like a serum with a high lipaemic index would have (Figure 1).

The possible causes of abnormal plasma coloration according to literature could be related with the presence of high levels of metalbumin, myoglobin or methemoglobin. Hemolytic anemia, rhabdomyolysis and exposure to exogenous oxidizing agents could cause the rise of these proteins respectively [3].

To rule out these pathologies, we sequentially added lactate dehydrogenase plus haptoglobin for the assessment of intravascular hemolysis and creatine kinase for the study of rhabdomyolysis. Furthermore, we measured methemoglobin evaluating co-oximetry after requesting a new heparinized sample and rejecting the previous possibilities. The methodology used for the determination of all parameters was the Abbott Alinity assay, except for methemoglobin, which was determined by the radiometer ABL 90. The results shown in Table 1 rule out any of the aforementioned diagnoses.

The next step was checking if any of the drugs the patient was taking could cause abnormal coloration of the serum. According to the eltrombopag data sheet, hyperpigmentation is a reported adverse effect, although infrequently (≥1/1000 to <1/100) [4].

There have been a few small-case reports of patients on treatment with high doses of eltrombopag causing brown discoloration of serum.

One article, described a 52-year-old woman with aplastic anemia who after increasing the dose of eltrombopag to 150 mg/24h, showed a “Reddish-brown serum” [5]. Another one, reported a series of three acute myeloid leukemia patients with eltrombopag doses of 200-300 mg/24h who presented a “brown serum” [6]. The last article related to our topic, described a patient with “brown serum” who was treated with an eltrombopag dose of 150 mg/24h [7].

Discussion

Severe aplastic anemia (SAA) is a bone marrow hypoplasia/aplasia in association with pancytopenia [8]. According to the classification of disease severity and assessment of medical fitness, the Haematology service selected the best treatment approach for the patient. Some guidelines recommend in selected patients a combination of immunosuppressive therapy such as eltrombopag plus cyclosporine [4]. Eltrombopag is a non-peptide thrombopoietin receptor agonist approved for the treatment of SAA, idiopathic thrombocytopenic purpura and chronic hepatitis C associated thrombocytopenia [9]. In SAA we can reach the higher doses of eltrombopag (150 mg/24h) comparing to the other aproved uses [4], which could lead to a major exposure to the drug and a major number and more severe adverse reactions. All the aforementioned cases had doses greater than or equal to 150 mg/24h [5-7]. After detecting this abnormal color of the serum in our patient, he was called for a more exhaustive clinical review. The patient presented a slight cutaneous hyperpigmentation and a yellow sclera of the eye. The cause of the hyperpigmentation is not fully understood. Cases have been described in patients taking similar doses (150 mg/day) [9, 10], this may be of concern to the patient, but there are currently no known adverse clinical sequelae from tissue or plasma pigmentation from the use of eltrombopag [10]. Moreover, the effect is reversible after the drug withdrawal [8]. A differential diagnosis to take into account in patients with hyperpigmentation or a yellow hue is serious liver failure. This is important because the drug is hepatotoxic [11] since it is metabolized and eliminated mainly in the liver [12]. Thereby, we must monitor transaminases and bilirubin for any sign of liver damage in patients taking eltrombopag. This is not an easy task, because as we have described previously, it has the potential to give a brown hue which interferes with some laboratory tests. The interferences found in the spectrophotometric methods are dependent on the technology used. Positive interference in spectrophotometric measurements of total bilirubin has been detected in some laboratories [13-14], which is a serious analytical error because we could lead to a misdiagnosis of liver damage. A study carried out in 2016 evaluated possible interferences of eltrombopag using the Roche Cobas 6000 technology, finding changes greater than 10% in parameters associated with the lipid profile. On the other hand, no interferences were found in the determination of bilirubin or transaminases [15]. Laboratory professionals must know how their determinations are affected so an erroneous diagnosis is not reached. We found out that our patient had total bilirubin measurement of 20,5 umol/L (diazo method, Abbott Alinity) but discordant icterus index (icterus index=3.4, corresponding to 51-85 umol/L of total bilirubin). The unlikely nature of these results led us to investigate the possibility of an analytic interference.

The sample was sent to a reference laboratory that determined total bilirubin using high performance liquid chromatography (HPLC). The result was similar to our laboratory (total bilirubin = 17 umol/L). We can conclude that the measurement of total bilirubin in patients treated with eltrombopag is not affected by the Abbott Alinity assay. The adverse reactions described were reported to the Spanish Pharmacovigilance System through the spontaneous reporting system for adverse drug reactions.

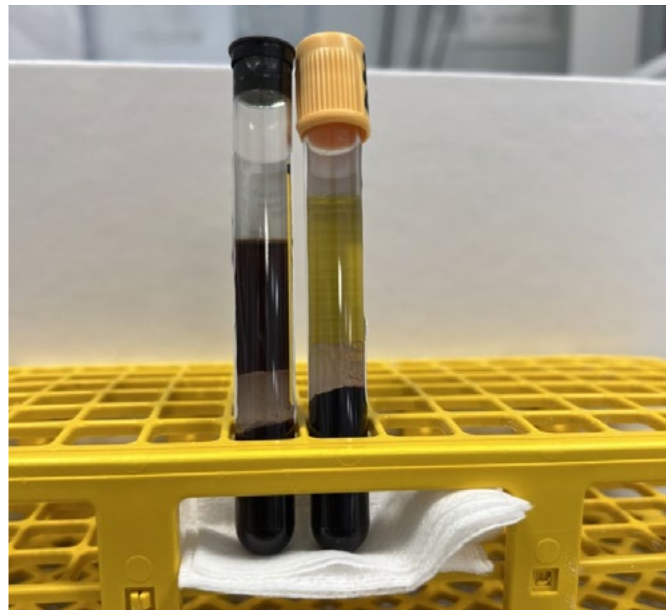


Figure 1: The brown serum on the left belongs to the patient on treatment with eltrombopag and the serum on the right to a control patient without treatment to show different coloration.

Table 1: Results of serial analyses performed in our laboratory

Laboratory test	02/05/23	08/05/23	Reference values
Total bilirubin (umol/L)	20.5	18.8	0.0-20.5
Direct bilirubin (umol/L)	10.2	8.5	0.0-8.5
Indirect bilirubin (umol/L)	10.3	10.3	0.0-11.9
Glumatic oxaloacetic transaminase (U/L)	16	20	5-34
Glumatic pyruvic transaminase (U/L)	12	13	0-45
Lactate dehydrogenase (U/L)	191	184	125-220
Haptoglobin (umol/L)	30	28	2-30
Creatine kinase (U/L)	160	170	30-200
Methemoglobin (%)	0.7	0.6	0.0-1.5

Learning points

• Our objective with this article is to present the importance of serum index. They are normally used as an alert system for possible analytical errors, but they can also contribute to the detection of hidden liver diseases, unstudied dyslipidemias or, as the case may be, adverse pharmacological effects.
• The most common causes of a brown colored serum are the presence of metalbumin, myoglobin, and methemoglobin and should be ruled out first.
• Interferences in the laboratory are usually method dependent and it is the laboratory professional who must know how they interfere with their results.
• Remember that whenever we are faced with a possible pharmacological adverse effect, it must be notified.

Author Contributions:

All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures or Potential Conflicts of Interest

No authors declared any potential conflicts of interest.

References

1. Mainali S, Merrill AE, Krasowski MD. Frequency of icteric interference in clinical chemistry laboratory tests and causes of severe icterus. Pract Lab Med. 2021 Oct 23;27:e00259.

2. Marques-Garcia F. Methods for Hemolysis Interference Study in Laboratory Medicine - A Critical Review. EJIFCC. 2020 Mar 20;31(1):85-97.

3. Dukic L, Maric N, Simundic AM. Dark brown serum and plasma samples: a case report. Biochem Med (Zagreb) 2020;30(2):021002

4. Ficha técnica de Eltrombopag®. Centro de Información online de Medicamentos de la AEMPS. Avaliable on: https://www.aemps.gob.es/cima. (Consulted: Mayo/2023).

5. Al-Samkari H, Goodarzi K. An Eltrombopag Red (Plasma) Alert. Acta Haematol. 2021;144(2):227-228.

6. Cardamone D, Milone MC, Glaser L, Frey NV, Kricka LJ. Eltrombopag and serum of a different hue. Arch Pathol Lab Med. 2013 Sep;137(9):1175.

7. Hager HB, Moksnes M, Brokner M. Brown serum. Tidsskr Nor Laegeforen. 2021 Oct 25;141. English, Norwegian.

8. Hwang YY, Gill H, Chan TSY, Leung GMK, Cheung CYM, Kwong YL. Eltrombopag in the management of aplastic anaemia: real-world experience in a non-trial setting. Hematology. 2018 Aug;23(7):399-404

9. Atay H, Kelkitli E, Turgut M. A rare side effect: eltrombopag associated hyperpigmentation. G Ital Dermatol Venereol. 2020 Dec;155(6):803-804

10. Rodgers GM, Kurtti AL, Gilreath JA. Are eltrombopag plasma and skin hyperpigmentation related? The eyes have it. Am J Hematol. 2019 Mar;94(3):394-395

11. Wong RSM, Saleh MN, Khelif A, et al. Safety and efficacy of long-term treatment of chronic/persistent ITP with eltrombopag: final results of the EXTEND study. Blood. 2017;130(23):2527-2536

12. Kim TO, Despotovic J, Lambert MP. Eltrombopag for use in children with immune thrombocytopenia. Blood Adv. 2018 Feb 27;2(4):454-461

13. Cheng THT, Tsui TKC, Kwok JSS, Lit LCW, Wong EYL, Kam RKT, Grote-Koska D, Staaden A, Okada H, Fuke N, Wong RSM, Li CK, Chan MHM. Comprehensive characterization and resolution of discrepant spectrophotometric bilirubin results in patients on eltrombopag therapy. Clin Chem Lab Med. 2020 Sep 25;58(10):1713-1723

14. Poventud-Fuentes I, Garnett E, Despotovic J, Devaraj S. Interference of eltrombopag with bilirubin measurements on the Vitros 5600 analyzer. J Clin Lab Anal. 2021 Aug;35(8):e23796

15. Gounden V, Zhao Z. Eltrombopag interference in routine chemistry testing. Ann Clin Biochem. 2016 Sep;53(Pt 5):611-4

The parallel lives of prostate specific antigen in cardiac troponin assays

Eleftherios P. Diamandis[1*]

[1]*Senior Scientist, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada.*

Article Info

Author of correspondence:
Dr. Eleftherios P. Diamandis, Senior Scientist, Lunenfeld-Tanenbaum Research Institute;
E-mail: diamandis@lunenfeld.ca;
Address:
Sinai Health System, Toronto, Canada

Serum prostate specific antigen (PSA) is a cancer biomarker that is widely used for the diagnosis and management of prostate cancer (CaP). Serum cardiac troponin is used widely for the diagnosis and monitoring of cardiovascular diseases, including, myocardial infraction (MI). These two biomarkers are among the very best in clinical chemistry testing, in terms of clinical sensitivity and specificity. By comparing the developments of PSA and cardiac troponin biomarkers, it can be realized that they have important analytical and clinical similarities. Below, I mention briefly the analytical and clinical developmental milestones of these two biomarkers over the last 50 years. Please note that due to space limitations, I will not attempt to assign credit to those who played a role in the discovery and validation of either the PSA or cardiac troponin assays. Reviews on this subject have been published elsewhere (1,2).

Prostate Specific Antigen

PSA was isolated from prostate tissue in the 1970's and it was realized at that time that it is prostatic tissue specific (tissue specificity is a major advantage for any biomarker). In the early 1980's, a competitive radioimmunoassay for measuring PSA in serum was developed (sensitivity ~0.1 ng/mL or 100 ng/L, first generation assay) * and used in the initial studies which have shown that PSA is a far superior marker of prostate cancer in comparison to prostatic acid phosphatase, a test that was used at that time. PSA assays were adopted in clinical practice within only 3-4 years, something that is relatively rare in the tumor marker field. A few years later (early 1990's), a more reliable assay based on the "sandwich" principle was developed and it was used by many clinical investigators to study the value of PSA for the diagnosis of prostate cancer and especially, for prostate cancer screening of asymptomatic men. These first-generation PSA assays were able to quantify PSA in the serum of males (concentration is usually 1-4 ng /mL or 1,000-4,000 ng/L). In early 2004, large prospective studies have been initiated to examine the value of PSA in prostate cancer screening, but the outcomes, published a few years later, were, and still are, controversial.

In 1997, our group developed the first third generation PSA assay which was capable of measuring PSA down to 0.001 ng/ml (or 1 ng/L) (3). These third-generation assays opened the possibility that PSA could be used as a highly sensitive marker of prostate cancer relapse, and it is currently used for this purpose (Figure 1). Around the same time, the company MesoScale Diagnostics and others, extended the sensitivity of the PSA assay by two orders of magnitude (fifth generation assays; sensitivity 0.00001 ng/mL or 0.01 ng/L). This development allowed measurement of PSA in the serum of males as well in serum of almost all females (the source of PSA in females is the breast). This development revealed that PSA is an excellent marker of hyperandrogenism in women.

Cardiac troponins

Just like PSA, in the 1970's, researchers isolated troponin T and troponin I from cardiac muscle and they found that this antigen is highly specific to this tissue. Cardiac troponins proved to be superior in specificity to the protein that was used in that time for studying myocardial infraction, which was the CK-MB isoenzyme of creatine kinase. In the early 1990's the first immunoassays for measuring cardiac troponins were developed (first generation) (4) and have shown to have the ability to measure troponin in serum after myocardial infraction (but not in serum of normal men or women). These first-generation assays had sensitivities, like PSA, of around 100 ng/L and were used to confirm the clinical superiority of troponins over immunological CK-MB assays. Like PSA, in the mid-1990's, scientists developed higher sensitivity cardiac troponin assays, which allowed the detection of lower levels of cardiac troponins in the bloodstream, thus enabling detection of smaller and smaller cardiac infarctions. At the same time, many companies have engaged in developing highly sensitive assays for cardiac troponins and eventually reached detectability in the low ng/L range (third, fourth and fifth generation assays) (5). This development necessitated the change of reportable units of troponin from ng/mL to ng/L, to avoid reporting numbers with many zeros, for easier communication with clinicians. Current reference ranges using the 99th percentile in normal individuals, are now up to around 10 ng/L for women and up to 15 ng/L for men. The reference ranges are still different between assays of manufacturers of high sensitivity cardiac troponin assays. The clinical utility of the cardiac troponin assays is so valuable, that the current definition of acute myocardial infarction (AMI) is partly based on the changing levels of cardiac troponins (their kinetics), along with the combination of clinical symptoms. Hundreds of papers have been published on the clinical use of cardiac troponins for diagnosis and monitoring of myocardial infraction and other cardiac diseases. high sensitivity cardiac troponin assays. The clinical utility of the cardiac troponin assays is so valuable, that the current definition of acute myocardial infarction (AMI) is partly based on the changing levels of cardiac troponins (their kinetics), along with the combination of clinical symptoms. Hundreds of papers have been published on the clinical use of cardiac troponins for diagnosis and monitoring of myocardial infraction and other cardiac diseases.

Concluding remarks

PSA and troponin assays have evolved from crude, semiquantitative immunoassays (first generation) to highly sensitive and specific tests (up to fifth generation), that have revolutionized the diagnosis and monitoring of patients with prostate cancer or heart diseases. Harmonization of results between assays from diverse manufacturers, is a major current and future step forward. My prediction is that these assays will likely remain in the menu of clinical chemistry testing for many decades, if not centuries. Do we need sixth-generation PSA and cardiac troponin assays? No, because the fifth-generation assays can comfortably and precisely measure these two biomarkers in almost all healthy and diseased men and women, including men whose prostate has been surgically removed.

Footnote

*Assay generations. Conventionally, the first developed assay for a biomarker is called "first generation". Subsequent generations usually have an approximately 10-fold increase in sensitivity.

References

1. Rittenhouse H, Blase A, Shamel B, Schalken J, Groskopf J. The long and winding road to FDA approval of a novel prostate cancer test: our story. Clin Chem 2012; 59:32-4. doi:10.1373/clinchem.2012.198739.

2. Garg P, Morris P, Fazlanie AL, et al. Cardiac biomarkers of acute coronary syndrome: from history to high-sensitivity cardiac troponin. Intern Emerg Med 2017; 12:147–155. https://doi.org/10.1007/s11739-017-1612-1.

3. Yu H, Diamandis EP, Wong PY, Nam R, Trachtenberg J. Detection of prostate cancer relapse with prostate specific antigen monitoring at levels of 0.001 to 0.1 microG./L. J Urol. 1997;157:913-8. doi:10.1016/s0022-5347(01)65083-3.

4. Katus HA, Remppis A, Looser S, et al. Enzyme linked immuno assay of cardiac troponin T for the detection of acute myocardial infarction in patients. J Mol Cell Cardiol 1989; 21:1349–1353. doi:10.1016/0022-2828(89)90680-9.

5. Mueller M, Celik S, Biener M, Vafaie M, Schwoebel K, Wollert KC, et al. Diagnostic and prognostic performance of a novel high-sensitivity cardiac troponin T assay compared to a contemporary sensitive cardiac troponin I assay in patients with acute coronary syndrome. Clin Res Cardiol 2012; 101:837–45.

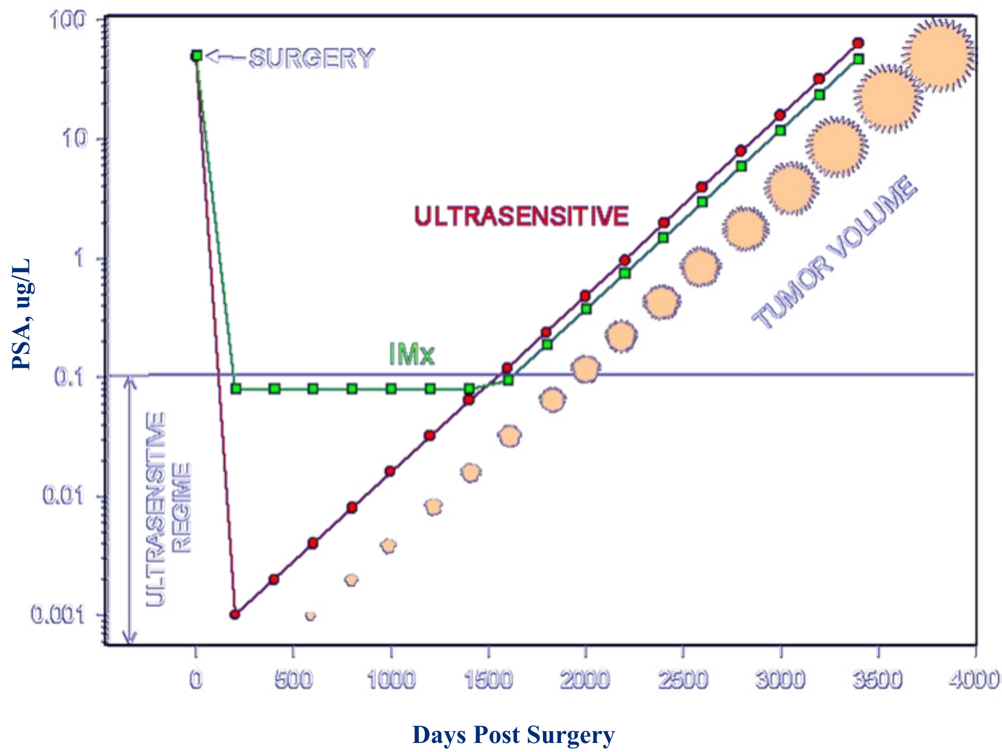


Figure 1: A hypothetical example of monitoring the progression/relapse of a prostate cancer patient after radical prostatectomy (surgery). Green dots represent a first-generation PSA assay (Abbott IMx). Red dots represent a third-generation assay, described in Ref. 3.

Editor-in-chief

János Kappelmayer
Department of Laboratory Medicine
University of Debrecen, Hungary

Assistant Editor

Harjit Pal Bhattoa
Department of Laboratory Medicine
University of Debrecen, Hungary

Editorial Board

Khosrow Adeli, The Hospital for Sick Children, University of Toronto, Canada
Borut Božič, University Medical Center, Lubljana, Slovenia
Edgard Delvin, CHU Sainte-Justine Research Center, Montréal, Québec, Canada
Nilda E. Fink, Universidad Nacional de La Plata, Argentina
Ronda Greaves, Biochemical Genetics, Victorian Clinical Genetics Services, Victoria, Australia
Mike Hallworth, Shrewsbury, United Kingdom
Ellis Jacobs, EJ Clinical Consulting, LLC, USA
Allan S. Jaffe, Mayo Clinic, Rochester, USA
Bruce Jordan, Roche Diagnostics, Rotkreuz, Switzerland
Gábor L. Kovács, University of Pécs, Hungary
Evelyn Koay, National University, Singapore
Tamas Kőszegi, University of Pécs, Hungary
Janja Marc, University of Ljubljana, Slovenia
Gary Myers, Joint Committee for Traceability in Laboratory Medicine, USA
Tomris Ozben, Akdeniz University, Antalya, Turkey
Maria D. Pasic, Laboratory Medicine and Pathobiology, University of Toronto, Canada
Maria del C. Pasquel Carrera, College of Chemists, Biochemists and Pharmacists, Pichincha, Ecuador
Béla Nagy Jr, University of Debrecen, Hungary
Oliver Racz, University of Kosice, Slovakia
Rosa Sierra Amor, Laboratorio Laquims, Veracruz, Mexico
Sanja Stankovic, Institute of Medical Biochemistry, Clinical Center of Serbia, Belgrade, Serbia
Danyal Syed, Ryancenter, New York, USA
Grazyna Sypniewska, Collegium Medicum, NC University, Bydgoszcz, Poland
Peter Vervaart, LabMed Consulting, Australia
Stacy E. Walz, Arkansas State University, USA

Publisher: IFCC Communications and Publications Division (IFCC-CPD)

Copyright © 2023 IFCC. All rights reserved.

The eJIFCC is a member of the **Committee on Publication Ethics (COPE).**

The eJIFCC (Journal of the International Federation of Clinical Chemistry)
is an electronic journal with frequent updates on its home page. Our articles, debates, reviews and editorials are addressed to clinical laboratorians. Besides offering original scientific thought in our featured columns, we provide pointers to quality resources on the World Wide Web.

This is a Platinum Open Access Journal distributed under the terms of the *Creative Commons Attribution Non-Commercial License* which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Published by:

ifcc
International Federation
of Clinical Chemistry
and Laboratory Medicine

www.ifcc.org