July 2023 ISSN 1650-3414 Volume 34 Number 2

eJIFCC

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

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

In this issue

In this issue

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Basics of laboratory statistics

Vivek Pant, Santosh Pradhan, Keyoor Gautam

Samyak Diagnostic Pvt Ltd, Bench to Clinic Research Center, Kathmandu, Nepal

ARTICLE INFO

Corresponding author:
Dr. Vivek Pant
Consultant Biochemist
Samyak Diagnostic Pvt Ltd
Bench to Clinic Research Center
Kathmandu
Nepal
Phone: +977 9841486789
E-mail: drvpant@gmail.com

Key words:
standard deviation, coefficient of variation, quality control, verification, reference interval

ABSTRACT

The strict monitoring of examinations and evaluation of newer methods or instruments is a daily routine in clinical laboratory. The automated analyzers accumulate an enormous amount of data from patients' examinations and quality control procedures. This laboratory data is meaningless if it does not generate the information that we can extend to the population of our interest. In an analytical work, the most important operation is the comparison of data, to quantify accuracy and precision and to generate meaningful explanation for clinician and patients queries. Most of the information needed in the regular laboratory work can be obtained with the use of simple convenient statistical tools. This article describes the basics of laboratory statistics, the knowledge of which answers about the application of quality control in laboratory, accuracy and diagnostic power of our examinations, variability in reports, comparison of different methods and derivation of a biological reference interval for an analyte.

INTRODUCTION

In the clinical laboratory, statistics are used to verify and monitor the performance of analytical methods and to guide the clinical interpretation of laboratory data. Laboratory statistics can be broadly described under following headings

1. Quality control and statistics
2. Diagnostic power of a laboratory test
3. Variability in Reports
4. Method Comparison
5. Reference Interval

QUALITY CONTROL

Quality control is the analysis of control materials, comparing the results with a predefined acceptable limit and plotting a result in a chart. Internal quality control data is best visualized using a Levey-Jennings control chart where the dates of analyses are plotted along X-axis and control values are plotted on Y-axis. The mean and one, two and three standard deviation (SD) limits are also marked on the Y-axis. Inspecting the pattern of plotted points provides a simple way to detect random error and shifts or trends in the calibration. Daily repeating the same control sample should produce a normally distributed set of data. This means, approximately 66% of values should fall between +/-1SD ranges and be evenly distributed on either side of mean. Similarly, 95% and 99 % of values should fall within +/-2SD and +/-3SD limits respectively. (Figure 1) A calculation of mean, standard deviation and coefficient of variation (CV) of this

Figure 1 Normal distribution curve

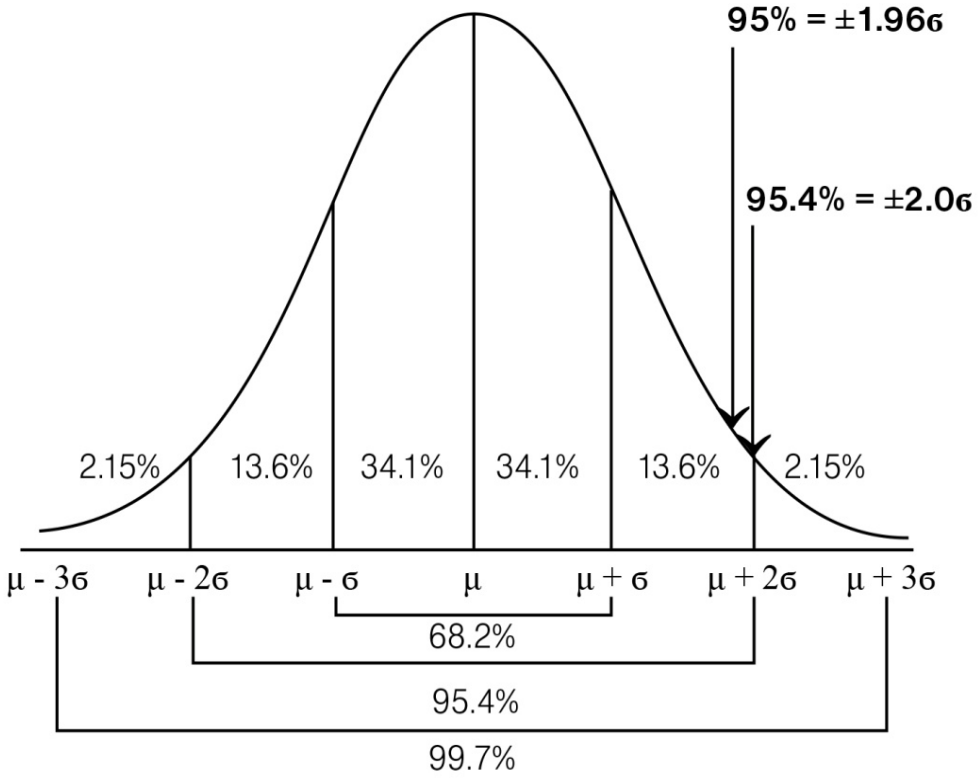
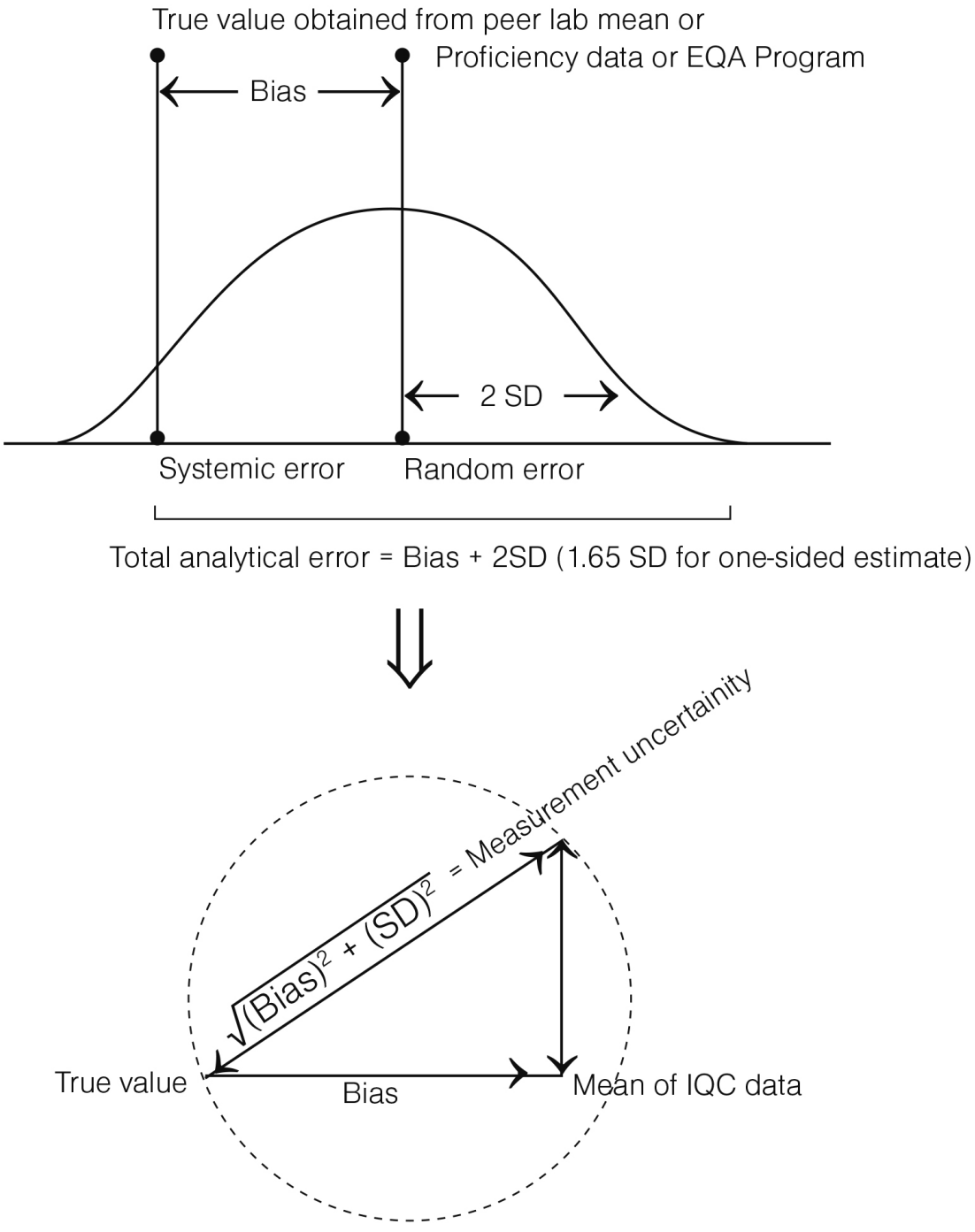


Figure 2 **Illustration of systemic and random error and concept of measurement of uncertainty**



dataset is useful for further calculation and derivation of other laboratory entities (Figure 2). Mean is the average value of measurements and SD is the primary measure of dispersion or variation of the individual result from the mean value. To derive SD, we calculate the deviation from the mean for each observation; square those results, sum them, divide by the number of observations minus one, and finally take a square root. CV is the SD expressed as a percent of the mean. Acceptable CV needs to be defined for each analyte based on medical significance.

Quality control rules are designed to detect two types of error, systemic error or bias and random error or imprecision. Precision is the agreement with replicate measurements and therefore the imprecision is caused by increased random error. Accuracy is the agreement between best estimate of the mean of results and its true value, therefore inaccuracy is caused by increased systemic error. These two errors when combined give a total analytical error (Figure 3). [1] In practice, replicate measurements can reduce, but not completely eliminate systematic and random errors, and therefore total error cannot be exactly known. [2] It follows that the true value of a measured quantity cannot be exactly known either. This assumption is fundamental to the measurement of uncertainty (MU) approach.

MU approach focuses on identifying the dispersion of results that might have been obtained for an analyte if a sample had been measured repeatedly. To do this, the MU approach uses available data about repeated measurements from a given measuring system to define an interval of values within which the true value of the measured analyte is believed to lie, with a stated level of confidence. This can be simply estimated from the CV calculated from repeated measurements of internal quality control sample. (Figure 2 and 3) In the MU concept, a measurement result can comprise two uncertainties (i) that associated with a bias correction, and (ii) the uncertainty due to random effects. [3] Both these uncertainties are expressed as SDs which, when combined together, provide the combined standard uncertainty for the procedure. (Figure 2 and 3)

External quality control (EQC) refers to the process of controlling the accuracy of an analytical method by interlaboratory comparisons. Two of the most important comparison statistics of an Interlaboratory program are the coefficient of variation ratio (CVR) and standard deviation index (SDI), which are consensus-based metrics of imprecision and bias, respectively.

The CVR allows evaluating imprecision relative to the consensus group and is expressed mathematically by the formula: Lab CV/ Consensus group CV

If the labs imprecision is equal to the imprecision of consensus group, then CVR will be 1.0.

The SDI or Z-score is a useful parameter for evaluating bias relative to the consensus group and is expressed mathematically by the formula:

(Lab mean-Consensus group mean)/ Consensus group SD

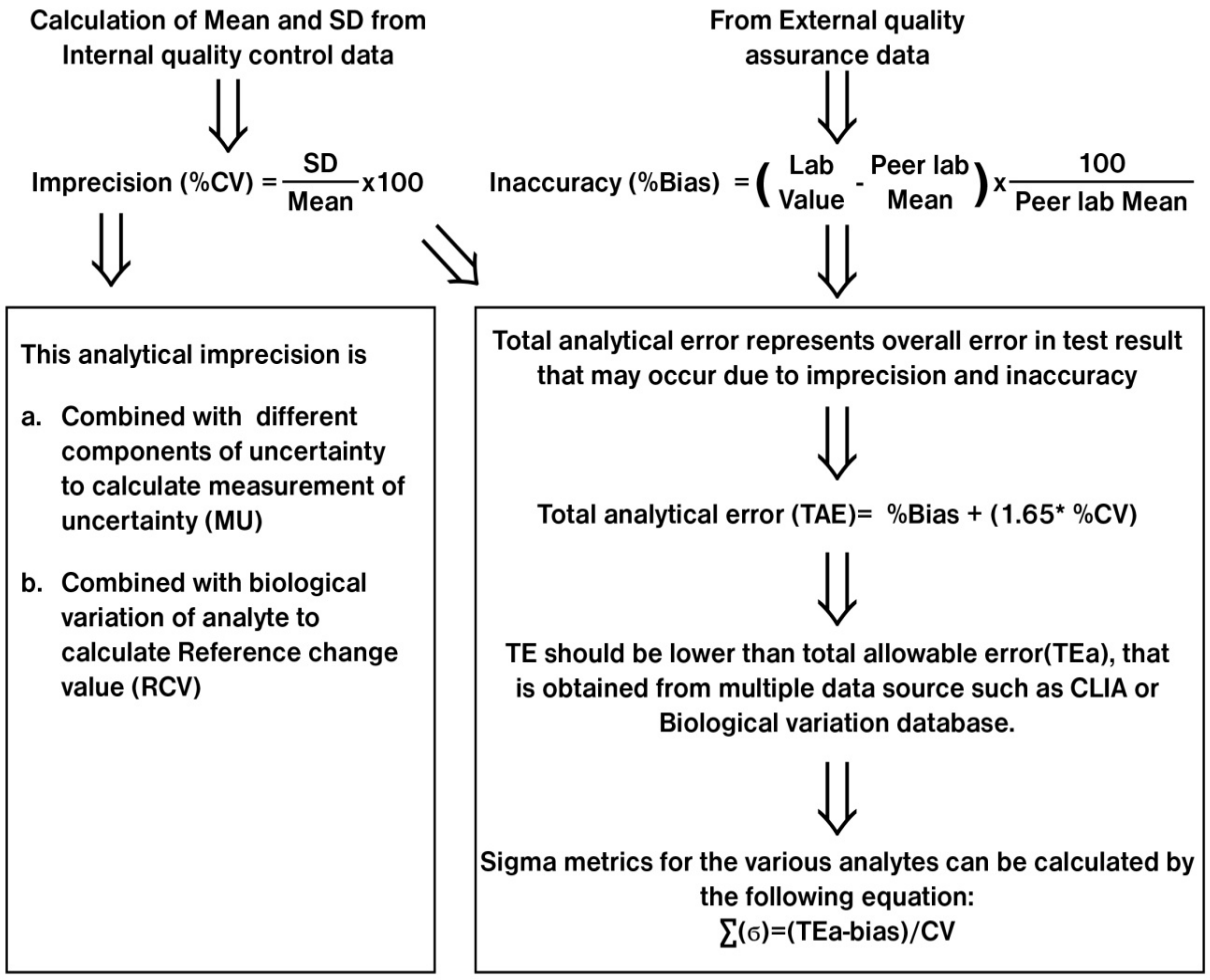
The target SDI is 0.0, which indicates that the labs mean is identical to the consensus group mean. A positive or negative deviation from this target statistic may indicate a bias compared to the consensus group mean.

DIAGNOSTIC POWER OF A TEST

Any user of the laboratory report wants to know the probability of disease given a positive or negative test result. There is no such ideal test which can achieve a perfect discrimination for non-diseased and diseased individuals.

Diagnostic accuracy of a test is measured by calculating the tests' sensitivity, specificity, and predictive values (Figure 4 and 5); these can be

Figure 3 Use of internal and external quality control data



further utilized to construct a Receiver Operating Characteristics (ROC) curve.

Limit of Detection (LoD), and Limit of Quantitation (LoQ) are terms used to describe the smallest concentration of a measurand that can be reliably measured by an analytical procedure. (Figure 5) LoD is the lowest analyte concentration at which detection is feasible. LoQ is the lowest concentration at or above the concentration of LoD and this concentration must be sufficient to produce analytical signals that meet predefined targets for bias, imprecision and total error. LoD is important for tests used to discriminate between the presence and absence of an analyte (e.g. drugs, troponin-I, human chorionic gonadotrophin). Likewise LoQ is important to reliably measure low levels of analyte (e.g. TSH, CRP) for clinical diagnosis and management.

Sensitivity and specificity are not absolute. They are affected by the prevalence of disease and may vary among different populations. Each laboratory test has its defined sensitivity and specificity by the manufacturer and it should be

taken into the clinical consideration for appropriate application of the test.

If a test has high sensitivity, it would not miss a disease, but will also yield false positive results. If a test has high specificity, it will find patients who do not have disease but there will be people who have disease and will be tested negative. This is more dangerous if the investigations are related to infectious diseases. The threshold for a given test is determined by examining the ROC curve, where the sensitivity is plotted as the function of the 1-Specificity for different cut off points. (Figure 6) The area under the ROC curve reflects the diagnostic ability of a test to differentiate people with and without disease of interest.

For example, if the area under the ROC curve is 96%, then there is a 96% chance that a randomly selected diseased person would have a more abnormal result than a randomly selected

Figure 4 **Formulas to calculate diagnostic power of a test**

$$\textbf{Sensitivity} = \frac{\text{True positive}}{\text{True positive + False negative}}$$

$$\textbf{Specificity} = \frac{\text{True negative}}{\text{False positive+ True negative}}$$

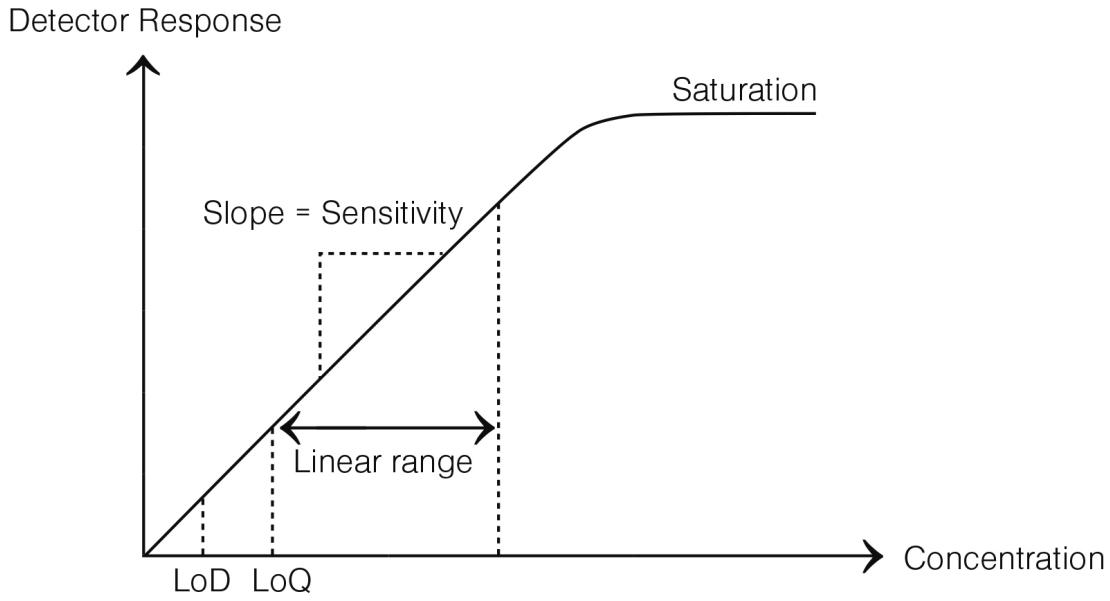
$$\textbf{Positive predictive value} = \frac{\text{True positive}}{\text{True positive+ False positive}}$$

$$\textbf{Negative Predictive value} = \frac{\text{True negative}}{\text{True negative+ False negative}}$$

$$\textbf{Positive Likelihood ratio} = \frac{\text{Sensitivity}}{\text{1-Specificity}}$$

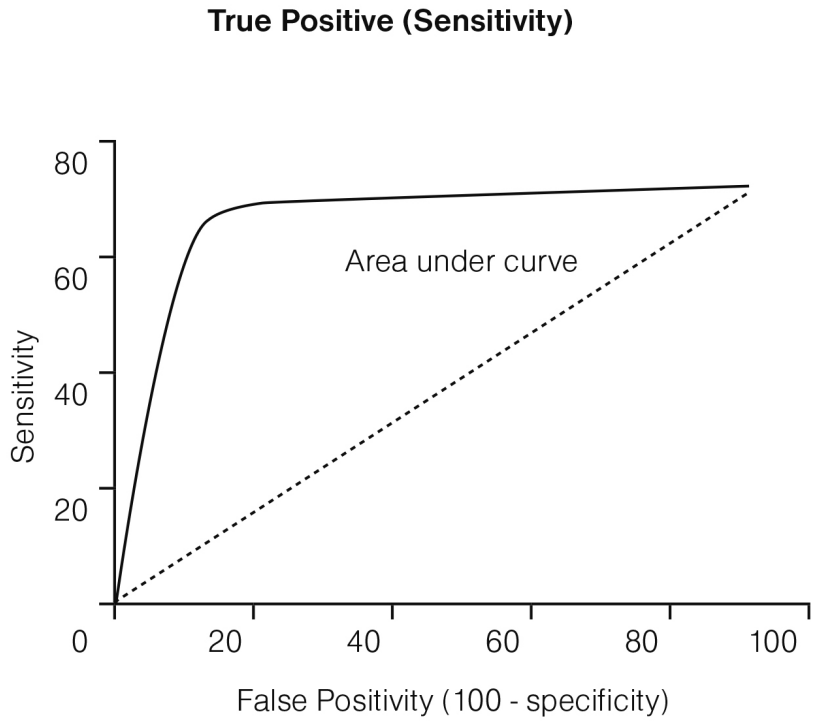
$$\textbf{Negative Likelihood ratio} = \frac{\text{1-Sensitivity}}{\text{Specificity}}$$

Figure 5 **The slope of the concentration versus detector response signifies the sensitivity of the test**



LOD- Limit of detection, LOQ- Limit of quantitation.

Figure 6 **ROC curve**



non-diseased person. The ROC curve also allows comparing the curves (diagnostic accuracy) generated from two or more tests.

Clinicians are more interested to know the predictive value of a test. The predictive value denotes the overall performance of a diagnostic laboratory test in terms of its ability to accurately distinguish the presence of a disease state with a positive test result from the absence of a disease state with a negative test result. (Figure 4). The negative predictive value can be regarded as a reassurance number - when it is very high, the patient can be assured that they don't have disease.

To calculate the predictive values, the 2x2 table is constructed. Predictive values are affected by outcome prevalence. The lower the disease prevalence, lower will be the positive predictive value and this will raise the negative predictive value. Thus, positive predictive value, even for a good test with a high sensitivity, can be poor when there are few persons with the disease. We can also calculate the predictive value using Bayes' Theorem which describes the probability of occurrence of an event related to any condition. [4]

For a laboratory screening tests, particularly where the results of the individual tests are highly variable, a statistical entity known as Multiple of the Median (MoM) is used to report the results. MoM is helpful to estimate the risk for pregnancy complication such as Down's syndrome, neural tube defect, preeclampsia in various weeks of gestation.

Example- Alpha feto protein (AFP) testing is used to screen for a neural tube defect during the second trimester of pregnancy. Because AFP concentrations normally increase during pregnancy, MoM is used to normalize the test result. The MoM is a measure of how far an individual test result deviates from the median value of a large set of AFP results obtained from unaffected pregnancies. For example, if the median AFP result at 16 weeks of gestation is 20 ng/mL and a pregnant woman's AFP result at that same gestational age is 60 ng/mL, then her AFP MoM is equal to 60 divided by 20 (60/20) or 3.0. In other words, her AFP result is 3 times higher than normal.

Calculation for MoM is done by dividing the patient result of particular biomarker by the median result of same biomarker determined by the laboratory. The Mom cut off for each parameter varies by laboratory as it depends on the population characteristics and medical history as well as the analyzer used for making the measurements.

METHOD VERIFICATION

All the invitro diagnostic instruments and reagents that are available must be documented and approved by an official agency. In Europe, the documentation must get a CE mark, and in the United States, an approval procedure by FDA is mandated.

Validation of the products is done at the manufacturer's level to show that the device/reagent is fit for its purpose. This includes measurement of trueness and precision, linearity, chemical interferences, carryover, and risk appraisal. [5] Clinical laboratories usually limit the verification process to compare claims regarding trueness and precision. The other verification criteria may be regarded as inherent to the method/instrument however it depends on the accreditation bodies.

To verify the precision, at least 5 observations during 5 days, in a patients sample or a reference material, are suggested. [6] When the imprecision is obtained from repeated measurements of the same sample and unchanged conditions, it is called the repeatability or within-series variation. If conditions change between estimating the imprecision, for example, from one day to

another or after recalibration of the measurement procedure, the imprecision is characterized as between-series imprecision. Using both these imprecision, the combined or intralaboratory imprecision can be obtained. Statistically, an ANOVA test can also be used to estimate the within- and between-series variation and provides a method to estimate the within-laboratory variation.

To verify bias, laboratories compare a new measurement procedure with previous ones by splitting samples into aliquots. At least 20 numbers of samples in the entire measuring interval is chosen and measured by both methods. [6] Before the statistical evaluation is performed, the scatter plot and difference plots should be carefully studied to identify outliers and are deleted. Statistically, the significance of the difference between the methods is evaluated by the Student t test. This data is used for various more advanced calculations, for example, the regression function, that is, the slope and intercept, and the correlation coefficient. This is discussed further in the method comparison section below.

METHOD COMPARISON

It is mandatory to evaluate analytical methods in the laboratory before their use for patient examinations. In addition to determining experiments for measuring accuracy and precision, it is also necessary to compare the new method to be introduced and other methods in use.

Method comparison involves testing patient samples during a number of different analytical runs by both the new and current methods. In most of the cases, comparison method is the existing method in one's own laboratory or a reference laboratory.

The comparison aims to estimate the constant and proportional differences between the two methods. Various statistical approaches can be used for method comparison procedures. Pearson's correlation coefficient is often used for such comparisons but does not provide appropriate conclusions. The correlation describes the linear relationship between two data sets, but not their agreement, and does not reveal whether there is a constant or proportional difference between the two data sets.

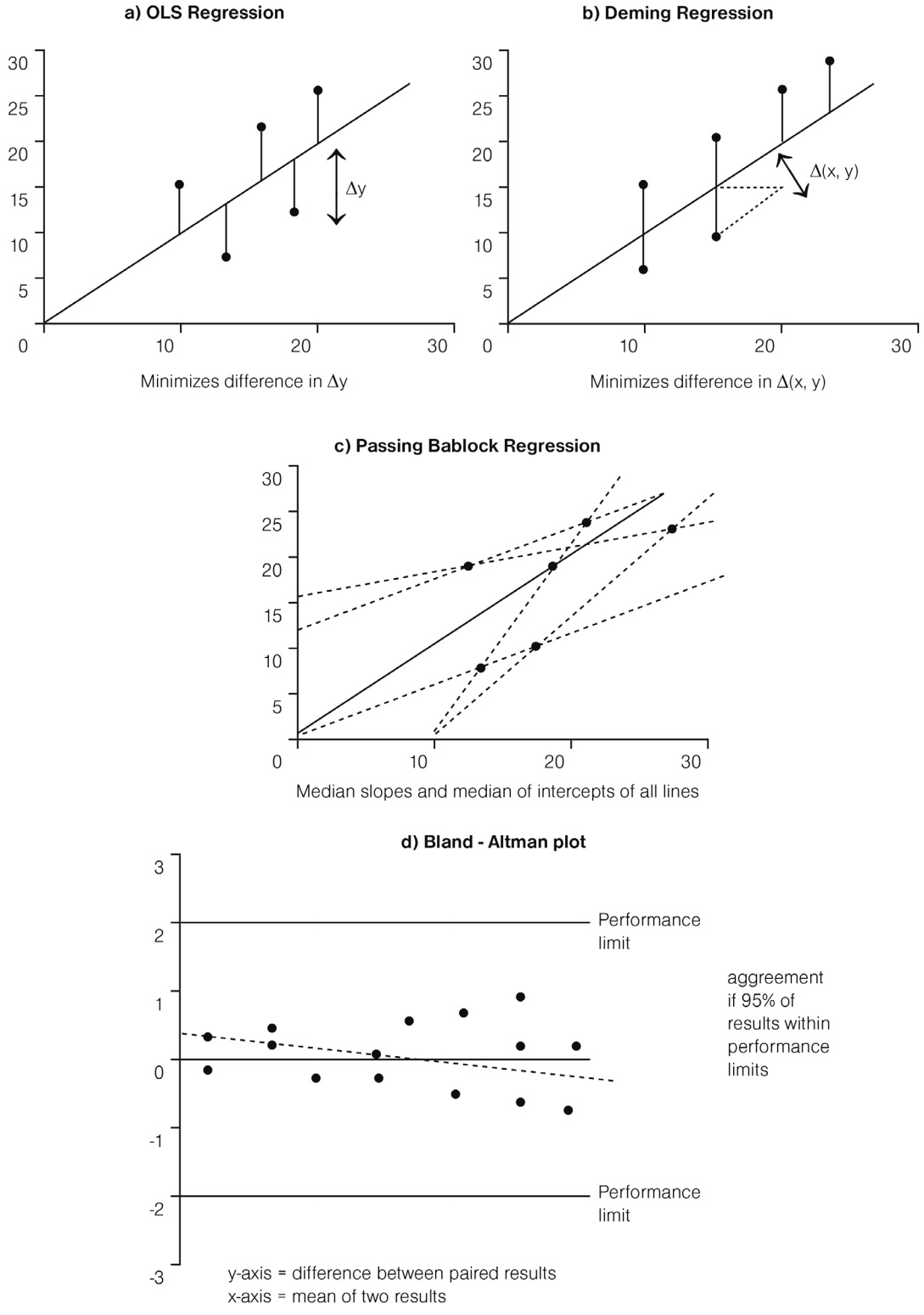
There are various ways to construct the function that binds two variables. To evaluate the equivalences between two methods, a regression function is used. A straight line can adequately describe the relationship between the two variables.

For this purpose, at least 40 patient samples should be analyzed by both methods with at least 2 reagent lots on each analyzer. [7] The analytical concentration should span the entire analytical range. The results are plotted on the Y-axis (dependent) and the reference method (existing) on the X-axis (independent). A linear regression line is inserted through the data points and the slope and Y intercept are calculated. (Figure 7a) There are a number of spreadsheets available that can automatically calculate and plot regression graphs which can be used by the laboratory. [8, 9] The best fit line is defined by the equation; $Y=mx + b$, where m is the slope and b is the Y intercept. A perfect correlation will have all points lying on a line at a 45^{o} angle to the X-axis.

This line will have a Y-intercept of zero and slope of 1. The correlation coefficient (R^2) will be 1.00 and the standard error will be 0.

The common model of this simple linear regression is easy but often may not be suitable for our daily evaluations. The linear regression assumes that the variable x is error-free and that the error of the test method, variable y, is distributed normally and is constant throughout the range of concentrations studied. (Figure 7a) We rarely meet these assumptions in practice.

Figure 7 Illustration of various regression and method comparison models



Thus, other statistical methods for comparing methods have been developed, such as the Passing-Bablok regression, Deming regression, Mountain plot, Bland and Altman plot. (Figure 7b-7d)

Deming regression does not assume that the reference method is free from error and it is the best approach to use when two methods are expected to be identical and the data is normally distributed without outliers. Passing-Bablok regression is used for nonparametric data and performs better when outliers are present. However, Passing-Bablok is computationally intensive and unreliable for small sample sizes.

VARIABILITY IN REPORTS

Serial measurements of laboratory parameter are often required to monitor patient's health. However, repeated laboratory measurements are seldom identical. The change in laboratory result may be due to biological variation, analytical imprecision or a change in patients' health condition. The minimum change required to conclude that two serial measurements are likely different is termed as the reference change value (RCV). A good clinical laboratory should have sufficient data to calculate RCV which are based on the estimates of biological variation (BV) data and analytical variation (AV) data. The BV data are mostly taken from the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) BV database, which delivers real time BV data for numerous analytes. [10] This database is based on results from systemic reviews and published studies by the BV data critical appraisal checklist. [11]

When the pre-analytical conditions are unvarying, the RCV formula becomes:

$$RCV = \sqrt{2} \times Z \times \sqrt{(CVA^2 + CVI^2)}$$

Where, Z indicates the number of standard deviations appropriate to the desired probability, 1.96 for $P < 0.05$; CVA, analytical imprecision; and CVI, within subject biological variation. The CVA of each test is provided by imprecision testing in laboratory.

Acceptable CV or analytical precision needs to be defined for each analyte based on medical significance. Generally, the precision should be equal to or less than one half of the within subject biological variation.

Therefore, analytes with larger biological variation do not require as much analytical accuracy as analytes with small biological variation. For example, BV of fasting triglyceride is 20%; therefore, analytical variation can be as high as 10% without significantly affecting medical decision making.

REFERENCE INTERVAL

When developing reference intervals (RI), clinical laboratories must consider what data sources and statistical methods to use. RI for the same measurements and instruments may differ between laboratories because of the differences in:

a. Operating conditions
b. Criteria for selection of healthy subjects
c. Patient populations
d. Geographical areas in relation to temperature, altitude, barometric pressure and humidity
e. Subject preparation and sample collection

The RI is defined as the interval corresponding to the central 95% of values of a reference population, including the two boundary limits: upper reference limit (+ 2SD) and lower reference limit (-2SD). (Figure 1)

It is recommended that medical laboratories determine their own RIs to cover the variability of their local populations and their specific

analytic methods and devices. For the process of RI determination, the Clinical Laboratory Standards Institute (CLSI) recommends "direct" approach, where well defined reference subjects are selected with pre-defined criteria and the measurements are done afterwards. Direct method is hard to apply for every laboratory in routine practice for it demands much time and money. The alternative approach is the "indirect" method where test results of patients that were ordered for screening, diagnosis or follow-up purposes are derived from laboratory information system (LIS) and used to determine the RIs. This method is faster and cheaper. Besides, the results obtained by the indirect method take into account the analytical and biological variability of the analyzed parameter. Recently, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Committee on Reference Intervals and Decision Limits encourages the use of indirect methods to establish and verify reference intervals.

Both parametric and non-parametric approaches may be taken when analyzing reference range data. The parametric approach involves calculating the mean and standard deviation to determine the range of values that fall within the 95% confidence interval. The non parametric approach involves establishing the values falling at the 2.5 and 97.5 percentiles of the population as the lower and upper reference limits. Outliers can have substantial effect on the calculation of reference ranges by this method and should be removed. Mathematically, outliers are results that differ from the mean by more than 3SD or differ from other results by more than 30%.

Consensus RI for some analytes is determined by medical experts based on the result of clinical outcome studies. Whenever, the consensus RI is available, clinical laboratories should report these values instead of determining their own RI. Example of consensus groups: American Diabetes association, American Heart Association, IFCC etc.

For an FDA approved test method, the clinical laboratories can adopt the manufacturers stated RI. However it should be verified in healthy cohort of samples. Ideally, 40 healthy samples (20 men and 20 women) should be tested and if 95% of the results fall within the published reference range, it can be accepted for use.

CONCLUSION

In this article the basic laboratory statistics is explained in its simplest form. This offers guidance to understand and employ basic statistical controls and methods required by the clinical laboratory.

However, the authors suggest to refer other sources for step-by- step guidance to the quality control, method development, validation/ verification and comparison of test methods.

REFERENCES

1. https://www.aacc.org/cln/articles/2013/september/total-analytic-error Accessed on 19/03/2023

2. Westgard JO, Westgard SA. Total analytic error. From concept to application. 2013.

3. White GH. Basics of estimating measurement uncertainty. The Clinical Biochemist Reviews. 2008 Aug; 29 (Suppl 1):S53.

4. Kallner A. Bayes' theorem, the ROC diagram and reference values: Definition and use in clinical diagnosis. Biochemia medica. 2018 Feb 15;28(1):16-25.

5. Westgard JO. Basic method validation, Westgard QC. Inc., Madison, WI. 2008:168-74.

6. Clinical and Laboratory Standards Institute, User Verification of Precision and Estimation of Bias: Approved Guideline, third ed., CLSI, Wayne, Pennsylvania, 2014. CLSI Document EP15-A3.

7. Clinical and Laboratory Standards Institute, Measurement Procedure Comparison and Bias Estimation Using Patient Samples: Approved Guideline, third ed., Clinical and Laboratory Standards Institute, Wayne, PA, 2013. CLSI document EP09-A3.

8. https://www.westgard.com/. Accessed on 19 March 2023.

9. https://www.acb.org.uk/.Accessed on 19 March 2023.

10. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, Jonker N, Minchinela J, Simon M, Braga F, Perich C, Boned B, Roraas T, Marques-Garcia F, Carobene A, Aslan B, Barlett WA, Sandberg S. The EFLM Biological Variation Database. https://biologicalvariation.eu/.

11. Aarsand AK, Røraas T, Fernandez-Calle P, Ricos C, Díaz-Garzón J, Jonker N, Perich C, González-Lao E, Carobene A, Minchinela J, Coşkun A. The biological variation data critical appraisal checklist: a standard for evaluating studies on biological variation. Clinical Chemistry. 2018 Mar 1;64(3):501-14.

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

CA125, Galectin-3 and FGF-23 are interrelated in heart failure with reduced ejection fraction

Damien Gruson[1,2], Diane Maisin[2], Anne-Catherine Pouleur[3], Sylvie A. Ahn[3], Michel F. Rousseau[3]

[1] *Pôle de recherche en Endocrinologie, Diabète et Nutrition, Institut de Recherche Expérimentale et Clinique, Cliniques Universitaires St-Luc and Université Catholique de Louvain, Brussels, Belgium*

[2] *Department of Clinical Biochemistry, Cliniques Universitaires St-Luc and Université Catholique de Louvain, Brussels, Belgium*

[3] *Division of Cardiology, Cliniques Universitaires St-Luc and Pôle de recherche cardiovasculaire, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Brussels, Belgium*

ARTICLE INFO

Corresponding author:
Dr. Damien Gruson
Pôle de recherche en Endocrinologie,
Diabète et Nutrition, (EDIN)
Université catholique de Louvain
Tour Claude Bernard
54 Avenue Hippocrate
B-1200 Brussels
Belgium
Phone: +32-(0)2-7646747
Fax: +32-(0)2-7646930
E-mail: damien.gruson@uclouvain.be

Key words:
CA125, Galectin-3, inflammation, NT-proBNP, biomarker, heart failure

ABSTRACT

Background

Carbohydrate Antigen 125 (CA125) is the most widely used biomarker in ovarian cancer screening. In p atients with heart failure (HF), increased levels of CA125 have been observed and related to disease severity. Our objective was to determine the association of CA125 levels with two biomarkers of adverse remodeling in HF patients with reduced ejection fraction (HFrEF).

Methods

CA125 circulating levels were determined with an electrochemiluminscent immunoassay. Concentrations of B-type natriuretic peptide (BNP), N-terminal proBNP (Nt-proBNP), Galectin-3 and Fibroblast Growth Factor 23 (FGF23) were also measured by immunoassays.

Results

CA125 levels were increased in HFrEF, were associated to disease severity according NYHA classes. Median CA125 concentration was also significantly related to cardiovascular mortality. CA125 concentrations were positively and significantly associated to Galectin-3 and FGF23.

Conclusions

Concentrations of CA125 are increased in patients with HFrEF, associated to disease severity and adverse cardiovascular outcomes. CA125 levels are also correlated to Galectin-3 and FGF-23, two biomarkers related to fibrosis and cardiovascular remodeling.

❖❖❖❖❖

The burden of heart failure (HF) is recognized worldwide (1,2). The sub-phenotyping of HF patients is important to anticipate potential adverse outcomes and adapt treatment. Biomarkers play an important role in the diagnosis of HF and natriuretic peptides, N-terminal pro B-type natriuretic peptide (NT-proBNP) and B-type natriuretic peptide (BNP), remain the first-choice biomarkers (1). Different classes of other biomarkers might inform on different aspects on HF development and provide additional valuable information about patients' risk (2). To this end, several tumor markers such as Carbohydrate Antigen 125 (CA125), CA 15-3, CA 19-9, carcinoembryonic antigen, alpha-feto protein and chromogranin, have been explored in HF (3).

The evidence about the involvement of CA125 in the pathophysiology of HF is accumulating. CA125, also known as MUC16, is a large glycoprotein synthesized by mesothelial cells and is the most widely used biomarker in ovarian cancer screening (3). In patients with HF, increased levels of CA125 have been observed, strongly associated with right-sided HF parameters, and related to disease severity as such with diagnostic and prognostic perspectives (3,4).

The hypothesis beyond the elevation of CA125 in HF patients include congestion and inflammation (5). CA125 is non-linearly and positively associated with intrarenal venous flow (IRVF) measured by Doppler ultrasound, a potential surrogate marker of renal congestion (6). Inflammatory process and remodeling might also trigger CA125 in HF (4).

Other biomarkers like Galectin-3 (Gal-3) and Fibroblast growth factor 23 (FGF-23) have been related to inflammation and adverse remodeling in HF. FGF-23 is produced by osteocytes, regulates phosphate homeostasis and has also been evaluated in HF and linked to adverse outcomes, inflammation, and fibrosis (7).

Our objectives were to assess CA125 levels in a group of HF patients with reduced ejection fraction (HFrEF) and to evaluate its association with Gal-3 and FGF-23.

This study is retrospective and blood samples were collected in 102 HF patients with reduced left ventricular ejection fraction. Each patient gave informed consent, and the protocol was approved by the local institutional review board. Demographic information including medical history (New York Heart Association [NYHA] class), clinical signs and standard laboratory data were recorded. All HF patients received optimal therapy and none of the female patients had ovarian cancer. CA125 concentrations were determined with a two-sites electrochemiluminescent automated assay on the Cobas® 8000 platform (Roche Diagnostics, Mannheim, Germany). The upper limit of the reference interval (URL) for the CA125 assay is 35 U/mL. N-terminal proBNP (NT-proBNP) was measured with automated electrochemiluminescent immunoassay also on the Cobas® 8000 platform. Gal-3 and FGF-23 concentrations

were determined with enzyme-linked immunosorbent assays as previously described (8).

Biomarkers were modelled as continuous variables. The non-parametric Spearman rank correlation coefficients were used to assess the relationships between biomarkers, age, EF and GFR. Multiple regression analysis was performed to test the independent associations between age, gender and the different biomarkers. Statistical analysis was performed using Medcalc software version 20.111 (Medcalc Software Ltd).

Patients' characteristics were as followed: mean age: 69 ± 13 years; males n=89; females n=23; NYHA II-IV; etiology: ischemic n=86, dilated cardiomyopathy n=26; mean left ventricular ejection fraction (EF): 23 ± 6%). The median circulating levels of NT-proBNP and BNP were 3356 ng/mL [76-33020] and 532 ng/L [range: 21-5017] respectively. Median circulating levels of Gal-3 and FGF23 were 17 ng/mL [range: 8-50] and 102 RU/mL [20-15000], respectively.

The mean CA125 in HF patients was 135 U/mL [range: 5-2587]. CA125 concentrations above the URL were observed in 57 % of the HFrEF patients. CA125 concentrations were significantly related to NYHA classes ($p<0.001$, Figure 1A) and geometric means were 23 U/mL in NYHA class II (n=45), 77 U/mL (n=44) in NYHA class III and 246 U/mL in NYHA class IV (n=12). CA125 significantly and negatively correlated to left ventricular ejection fraction ($r=-0.27$, $p<0.001$) and higher CA125 concentrations were related to the lowest survival rate (Figure 1B).

CA125 concentrations were positively and significantly associated to Galectin-3 ($r=0.31$, $p<0.001$) and FGF23 ($r=0.38$, $p<0.001$) (Figure 2A and Figure 2B). CA125 was also significantly related to natriuretic peptides. With multiple regression analysis the independent determinant of CA125 levels were age, BNP and Galectin-3.

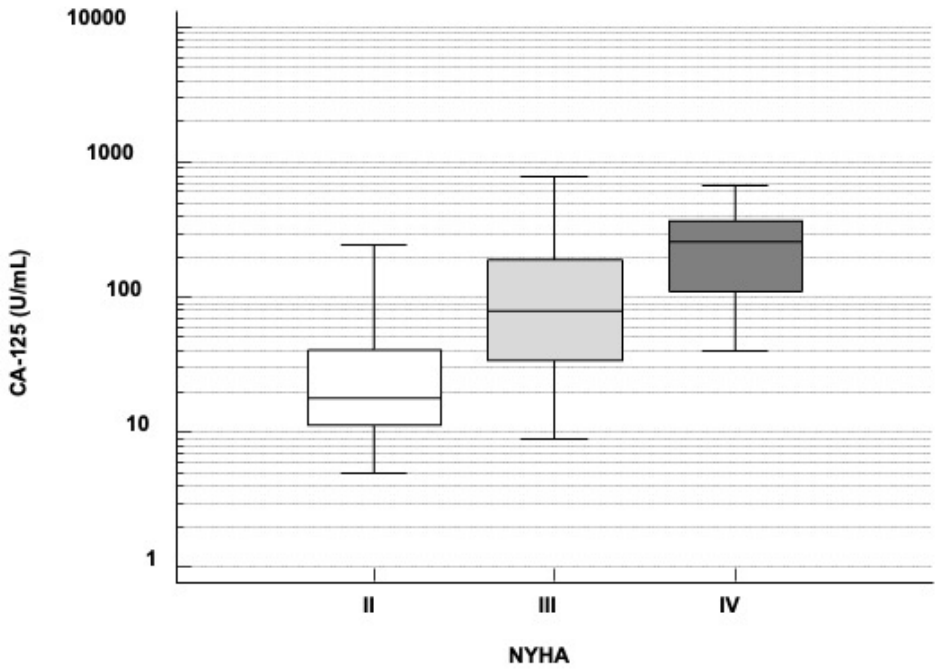
Our study showed, as already evident from literature, a raise of CA125 concentrations in HFrEF patients and association of CA125 with disease severity and prognosis. Interestingly, our study unrevealed significant and positive relationships between CA125 levels and two biomarkers of cardiovascular remodeling Gal-3 and FGF23.

The available evidence of the role of CA125 in the pathophysiology of HF is increasing with related perspectives to forward the diagnosis of HF (4). This was observed in our study with a significant proportion of HFrEF patients having significant increase of CA125. Two hypotheses are formulated in literature to explain such elevation of CA125 in HF, congestion, and inflammation. Data show that CA125 can predict the presence of a congestive intrarenal venous flow in patients with acute HF (6). The involvement of CA125 in the inflammatory process and remodeling in HF is also documented in the literature. Experimental data have also suggested a potential molecular interaction between CA125 and Gal-3; however, the biological and clinical relevance of this interaction is still uncertain (9). We reinforce this hypothesis through the significant relationships that we found in our study between CA125 and both Gal-3 and FGF-23. The participation of CA125 to such remodeling pathways can confirm interactions representing therapeutical targets. This is already a perspective for high-grade serous ovarian cancers and other MUC16/CA-125-expressing malignancies where targeting Gal-3 with a high-affinity antibody has been proposed (10).

CA125 testing offers several advantages in HF, which are presented in figure 3. Nevertheless, if the analytical and clinical value can be estimated as high, as different assays for CA125 are available but not standardized, it is important to mention that the decision limits need to be adapted to each method.

Our study is preliminary and has several limitations. A first one is a clear limited number of

Figure 1 **Association of CA125 circulating levels and New York Heart Association (NYHA) classes (A) and survival of HF patients (B)**



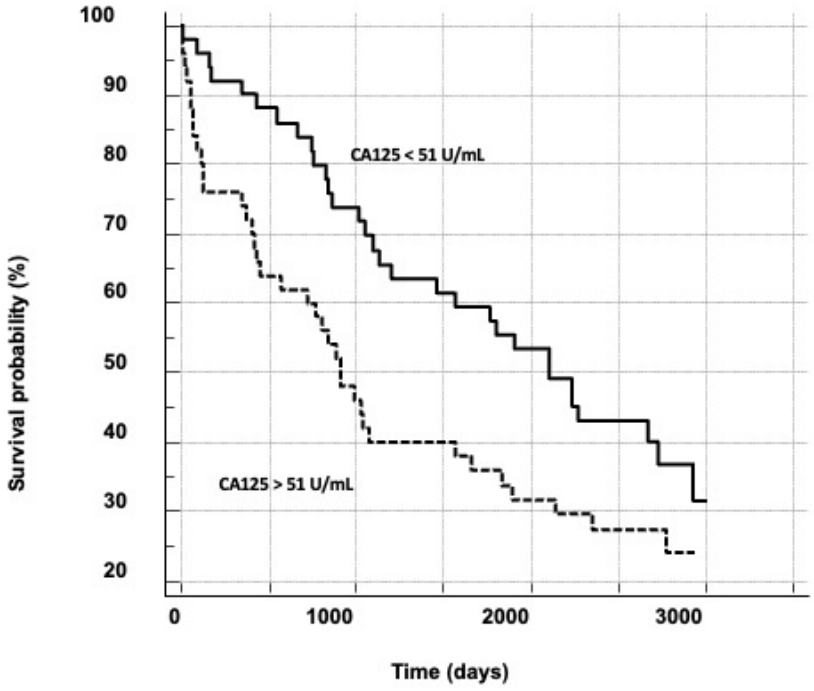
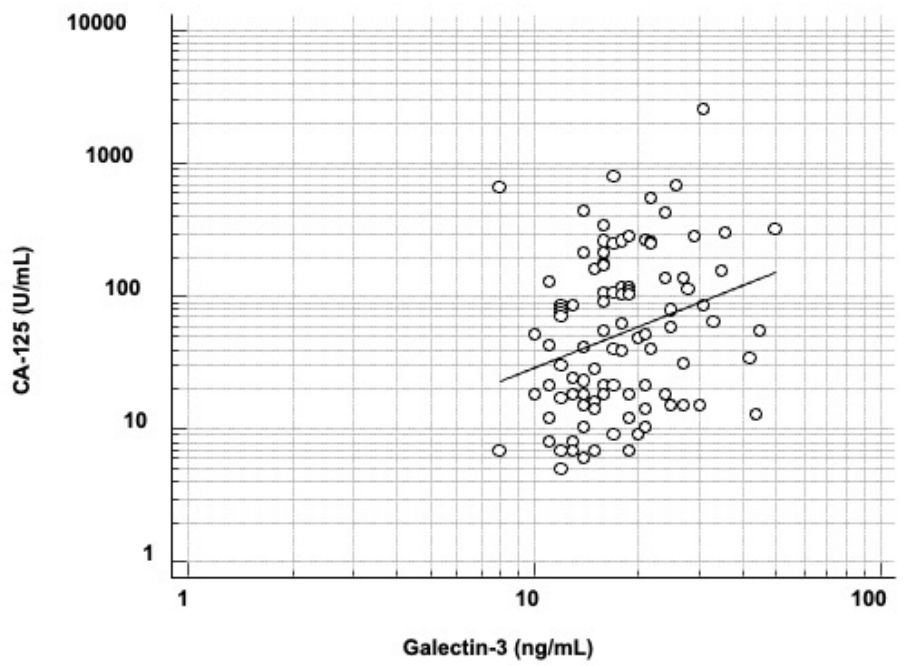


Figure 2 **Relationships between Galectin-3 and CA125 concentrations (A) and FGF-23 and CA125 concentrations (B) in HFrEF patients**



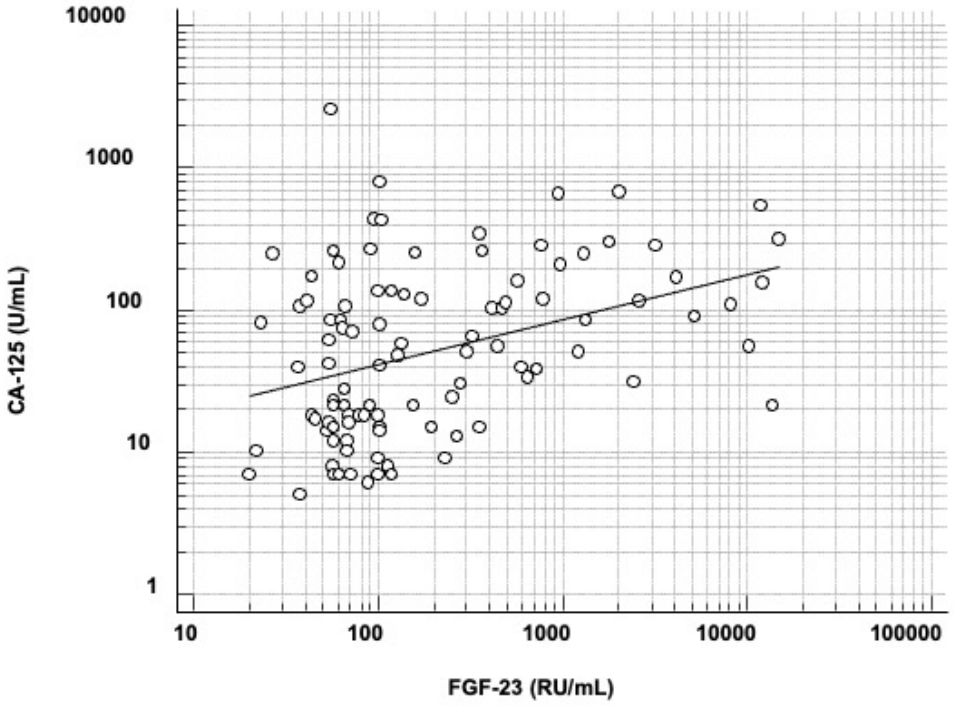
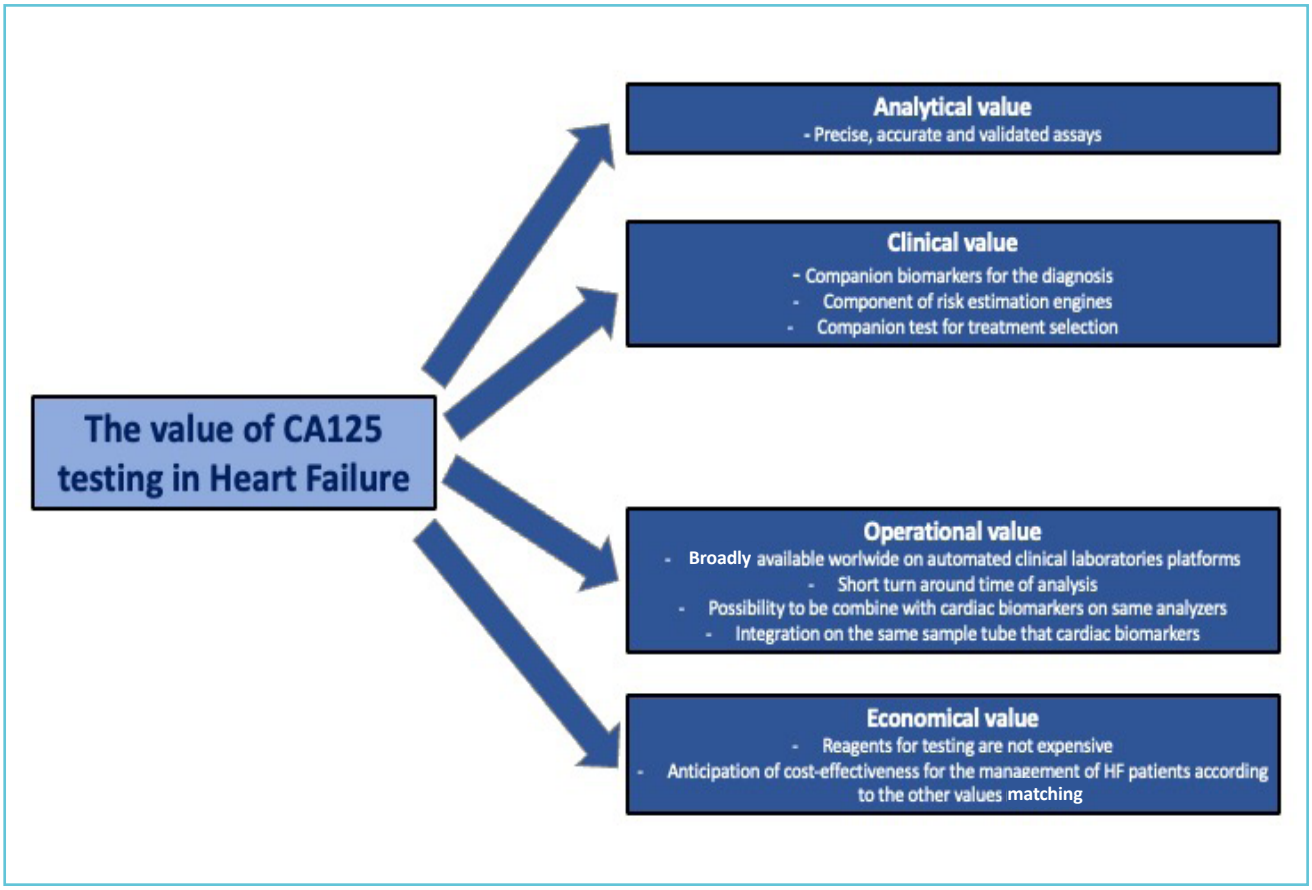


Figure 3 The potential value of CA125 testing in HF according to analytical, clinical, operational, and economic criteria



patients, even if the cohort was homogenous. Our observations will have to be confirmed on larger cohorts of patients. A second, is the lack of data from imaging to correlate fibrosis and remodeling to blood biomarkers.

In conclusion, our study showed increased CA125 concentrations in patients with HFrEF and a relation with disease severity. CA125 is also significantly and positively correlated to Galectin-3 and FGF-23, two biomarkers related to fibrosis and cardiovascular remodeling.

❖❖❖❖❖

Author contributions

Damien Gruson performed experimental design, conducted the experiments, analyzed and interpreted the results. Michel Rouseau, Sylvie Ahn and Anne-Catherine Pouleur were involved in the experimental design and read and approved the final version of the paper. Diane Maisin helped with analyses and interpretation of the results, and approved the final version.

Conflict of interest statement

The authors confirm that this paper content has no conflict of interests.

❖❖❖❖❖

REFERENCES

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart

failure Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J [Internet]. 2021 Sep 21 [cited 2022 Jul 30];42(36):3599–726. Available from: https://academic.oup.com/eurheartj/article/42/36/3599/6358045

2. González A, Richards AM, de Boer RA, Thum T, Arfsten H, Hülsmann M, et al. Cardiac remodelling – Part 1: From cells and tissues to circulating biomarkers. A review from the Study Group on Biomarkers of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail [Internet]. 2022 Jun 1 [cited 2022 Jul 30];24(6):927–43. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ejhf.2493

3. Yilmaz MB, Nikolaou M, Solal AC. Tumour biomarkers in heart failure: is there a role for CA-125? Eur J Heart Fail [Internet]. 2011 Jun [cited 2022 Jan 21];13(6):579–83. Available from: https://pubmed.ncbi.nlm.nih.gov/21525015/

4. Núñez J, de la Espriella R, Miñana G, Santas E, Llácer P, Núñez E, et al. Antigen carbohydrate 125 as a biomarker in heart failure: a narrative review. Eur J Heart Fail [Internet]. 2021 Sep 1 [cited 2022 Jan 21];23(9):1445–57. Available from: https://pubmed.ncbi.nlm.nih.gov/34241936/

5. Higgins A, Tang WHW. Carbohydrate antigen 125 in heart failure: congestive kidneys or beyond? Eur Heart J Acute Cardiovasc Care [Internet]. 2021 Jun 30 [cited 2022 Jan 21];10(5):484–6. Available from: https://pubmed.ncbi.nlm.nih.gov/33948631/

6. Núñez-Marín G, de la Espriella R, Santas E, Lorenzo M, Miñana G, Núñez E, et al. CA125 but not NT-proBNP predicts the presence of a congestive intrarenal venous flow in patients with acute heart failure. Eur Heart J Acute Cardiovasc Care [Internet]. 2021 Jun 30 [cited 2022 Jan 21];10(5):475–83. Available from: https://pubmed.ncbi.nlm.nih.gov/33829233/

7. Duran A, daSilva-deAbreu A, Joury A, Ventura HO. FGF23 predicts outcomes in heart failure but questions remain unanswered. Int J Cardiol [Internet]. 2021 Sep 1 [cited 2022 Jul 30];338:145–6. Available from: http://www.internationaljournalofcardiology.com/article/S0167527321010342/fulltext

8. Gruson D, Ferracin B, Ahn SA, Rousseau MF. Comparison of fibroblast growth factor 23, soluble ST2 and Galectin-3 for prognostication of cardiovascular death in heart failure patients. Int J Cardiol [Internet]. 2015 Jun 15 [cited 2022 Jul 30];189(1):185–7. Available from: http://www.internationaljournalofcardiology.com/article/S0167527315007937/fulltext

9. Núñez J, Rabinovich GA, Sandino J, Mainar L, Palau P, Santas E, et al. Prognostic value of the interaction between galectin-3 and antigen carbohydrate 125 in acute heart failure. PLoS One [Internet]. 2015 Apr 13 [cited 2022 Jan 21];10(4). Available from: https://pubmed.ncbi.nlm.nih.gov/25875367/

10. Stasenko M, Smith E, Yeku O, Park KJ, Laster I, Lee K, et al. Targeting galectin-3 with a high-affinity antibody for inhibition of high-grade serous ovarian cancer and other MUC16/CA-125-expressing malignancies. Sci Rep [Internet]. 2021 Dec 1 [cited 2022 Jan 21];11(1). Available from: https://pubmed.ncbi.nlm.nih.gov/33580170/

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

A framework for implementing best laboratory practices for non-integrated point of care tests in low resource settings

Lena Jafri[1,2], Sibtain Ahmed[1], Hafsa Majid[1], Farooq Ghani[1], Tahir Pillay[3,4], Aysha Habib Khan[1], Imran Siddiqui[1], Shahid Shakeel[1], Shuja Ahmed[1], Saba Azeem[1], Adil Khan[5,6]

[1] *Department of Pathology and Laboratory Medicine, The Aga Khan University, Karachi, Pakistan*

[2] *Corresponding member for Pakistan-IFCC Committee on Point of Care Testing*

[3] *Department of Chemical Pathology, University of Pretoria, Pretoria, South Africa*

[4] *Chair, IFCC Communications and Publications Division*

[5] *Point-of-Care Testing & Clinical Chemistry, Department of Pathology and Laboratory Medicine, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, USA*

[6] *Chair, IFCC Committee on Point of Care Testing*

ARTICLE INFO

Corresponding author:
Lena Jafri
Associate Professor & Section Head
Chemical Pathology
Department of Pathology and
Laboratory Medicine
The Aga Khan University
Karachi
Pakistan
E-mail: lena.jafri@aku.edu

Corresponding member for Pakistan-IFCC
Committee on Point of Care Testing

Key words:
framework, POCT, non-integrated,
low resource settings

ABSTRACT

The method we respond to pandemics is still inadequate for dealing with the point of care testing (POCT) requirements of the next large epidemic. The proposed framework highlights the importance of having defined policies and procedures in place for non-integrated POCT to protect patient safety. In the absence of a pathology laboratory, this paradigm may help in the supply of diagnostic services to low-resource centers. A review of the literature was used to construct this POCT framework for non-integrated and/or unconnected devices. It also sought professional advice from the Chemical Pathology faculty, quality assurance laboratory experts and international POCT experts from the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). Our concept presents a comprehensive integrated and networked approach to POCT with direct and indirect clinical laboratory

supervision, particularly for outpatient and inpatient care in low-resource health care settings.

BACKGROUND

Point-of-care testing (POCT), also referred to as bedside testing or near patient testing (NPT), is a field of laboratory medicine that is developing rapidly in terms of analytical quality and clinical reach (1). The POCT solutions provide the clinician with a fast turnaround time of diagnostic results thereby enhancing patient care (2) (3). The technological developments with POCT, such as instrument miniaturization, ease of use and improved accuracy, have not been complemented by a coordinated approach to data management, connectivity and device software interoperability (4). Non-integrated POCT can be defined as a category of POCT that are conducted outside of a traditional laboratory setup but are not integrated into a unified device, system, or LIS. These tests typically rely on test reagent strips/kits and involve the interpretation of visual cues by a healthcare professional. Many clinically useful commercially available POCT devices are non-integrated or have limited interface with the Laboratory Information System (LIS). Some examples of non-integrated POCT include rapid influenza tests, rapid HIV tests, pregnancy tests, rapid malaria tests and SARS-CoV-2 antigen test. An integrated point-of-care test (POCT) device with connectivity is a medical diagnostic tool that can transmit the test results to a centralized database, LIS, or a healthcare provider's electronic medical record system through wireless or wired connectivity. This allows for real-time monitoring of patients, timely interventions, and remote consultations with healthcare providers.

Knowledge in good laboratory practices by POCT end-users, including physicians and allied health professionals working in patient care areas using these devices is limited (5). Furthermore, these end-users must demonstrate a commitment to quality assurance (QA) and quality control (QC) since this is essential for POCT reliability (6) (7). Despite the relative ease of POCT, regulatory bodies such as Joint Commission International (JCI) and the College of American Pathologists (CAP) recommend oversight by the central clinical laboratory for all hospital based POCT (8) (9). The current POCT program for integrated POCT instruments at our institution has strict oversight by the clinical laboratory for both QA and QC (10)(4). Testing performed using POCT devices that lack connectivity to a middleware system or LIS raises concerns regarding reliability. This is because there is no information captured regarding QC performance, the person who performed the testing, transmission of POCT results to the LIS. (11). Furthermore, it is probable that this kind of testing is carried out without standardization of training or supervision by qualified laboratory personnel, which is risky. Subsequently, POCT has not been appropriately utilized in these settings employing non-integrated devices. In order to cater for this crisis, we propose the current framework, considering all regulatory requirements as a practical guide to initiate non-integrated POCT at in-patient and out-patient health care settings. The framework is developed bearing in mind the challenge of POCT-related QA practices and regulatory compliances (12).

In the authors' experience, laboratory QA/QC instruction for non-integrated POCT devices in Chemical Pathology curricula in the national residency programs in Pakistan or most other countries is scant. This creates problems in educating future chemical pathologists on how to establish, evaluate and maintain the quality of in-clinic or inpatient POCT testing using such instruments. Furthermore, the clinical laboratory receives frequent requests for initiating POCT by devices which are not integrated. Acknowledging this void, this framework is outlined given the numerous POCT tests that can

be performed on non-integrated POCT instruments i.e., standalone devices with no IT connectivity and result transmission. This POCT framework for non-integrated and/or unconnected devices was created through a review of the literature using popular search engines such as PubMed, Cochrane, Embase, and Web of Science. It also sought expert consensus from the Chemical Pathology faculty, the CAP director of Aga Khan University (AKU), Karachi Pakistan, the POCT coordinators, the QA team at AKU and international POCT experts from International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). Our model proposes an extensive approach to POCT that has direct and indirect supervision by the clinical laboratory especially for outpatient and inpatient care in low resource health care settings.

SCOPE

This framework predominantly applies to handheld POCT non-integrated devices measuring single or multiple analytes in hospital and outpatient NPT settings. This model may also be used in fieldwork, research settings, in rural or low resource settings. Along with "direct" bedside testing it can be established in a "satellite laboratory" located close to an emergency unit or units for acute care. After several debates and consultations amongst authors the framework was drafted, the prospective framework was distributed amongst the authors for comment. Revisions were made to address each comment, and the final guidance document was approved prior to publication. As additional scientific studies become available and POCT instruments and analytical performance capability evolve, this framework may change; revision is anticipated approximately every two years. This framework is not intended to be all-inclusive; rather, it provides a minimum standard for maintenance of these non-integrated POCT instruments in the clinical setting.

NON-INTEGRATED POCT MANAGEMENT AND RESPONSIBILITIES

Hospitals with a competent POCT coordination should provide organizational and administrative structures for POCT test/device selection, method validation, data management, quality control, continuous trainings, and competency assessments (13). Like all other POCT program the non-integrated POCT devices must be linked to the existing or new POCT program of the hospital or academic medical centre.

Although the end users of POCT may be familiar with its routine administration and delivery of results, the clinical laboratory director or, ideally, a chemical pathologist with training in this area, must bear overall responsibility for the programme. All non-integrated POCT should be managed by central laboratory's POCT team managed by a POCT coordinator, an experienced medical technologist or scientist from a clinical laboratory (7). The POCT coordinator should provide leadership to all POCT users and POCT sites in the following four domains: POCT test introduction, quality assurance, education, and administration. The minimal objectives of providing oversight should be as follows:

- To provide high quality (accurate and precise) of all non-integrated POCT devices
- To assess the need of non-integrated POCT devices before introduction into clinical practice
- To ensure that non-integrated POCT devices are cost-effective
- To train and assess the competency of POCT users
- To provide written policies and standard operating procedures for POCT devices being used at those ancillary sites
- To provide faster turnaround times with minimal inconvenience to the patient

- To outline the billing system on POCT sites
- To ensure all logs are maintained even if interface with LIS is missing (quality control logs, temperature logs, maintenance logs)
- To ensure compliance with policies and procedures by conducting audits. For non-integrated POCT more frequent audits and visits to POCT sites are recommended

For a new non-integrated POCT device to be induced in a clinical setting and practice the test must be reviewed and approved by the POCT interdisciplinary committee (14). This committee would be led by the laboratory director and chemical pathologist (trained POCT expert) and should have members from various sections and departments such as microbiology, haematology, molecular pathology, blood bank and transfusion services. Requests and demands for new non-integrated POCT must be made through this committee. Attention must be paid to need assessment and whether the non-integrated POCT meets quality standards (15). The committee must ensure that it meets the safety and security requirements in relation to protecting data, patient confidentiality and risk management. Once the non-integrated POCT is in place and the responsibility of ongoing problems and compliance issues can be handled by the POCT end user committee (led by the POCT coordinator with pathologists or subject experts, POCT site supervisors, nursing managers, IT, and biomedical experts as members). Nurse directors or nursing managers qualify as POCT site supervisors. POCT site supervisors need to be vigilant for non-integrated POCT (16). They must be made responsible to establish and maintain a system where audits are performed to ensure quality control is being performed and documented and corrective action is being done for outlier results, according to written policies. In the event of a lack of IT connectivity, it is also necessary to manually update employee listings and training and competency records into a spreadsheet (17). Request for training and for competency assessment can be made to the POCT coordinator through these POCT site supervisors. POCT site supervisors should be made responsible for POCT in-house inventory and for administration of the daily operation of POCT at their respective site. Furthermore, audits should be performed to determine if critical results are being documented into patient charts and handled appropriately.

SELECTION AND EVALUATION OF NON-INTEGRATED POCT

Before bringing any non-integrated POCT into the POCT Program a clinical needs assessment should be conducted. A standard approach must be carried out for every new request of non-integrated POCT by answering some basic questions:

- What is the diagnostic caveat that clinicians are anticipating solving by using this non-integrated POCT?
- Is this non-integrated POCT cost-effective?
- Based on clinical requirements, what is the unacceptable turnaround time for each non-integrated POCT under evaluation?
- What are the potential risks to the patients because of non-integration with LIMS?
- How will the clinical laboratory control these risks?

Waived tests are excluded from method evaluation under Clinical Laboratory Improvement Amendment of 1988 (CLIA), although it is acceptable laboratory practice to confirm the manufacturer's declared performance standards. However, CAP does not entirely adhere to the CLIA way of categorizing tests and instead uses the POCT checklist to ensure compliance with CAP requirements. The CAP defines POCT as waived and nonwaived tests that are only

performed close to the location where the patients are. In comparison to moderately complex tests, waived tests have distinct requirements for quality control, reagents, competency assessment, and calibration. Both waived and moderately complicated tests must meet the same standards for proficiency testing, quality management, procedure manuals, specimen handling, results reporting, POCT instruments, personnel training and certifications, and safety (18).

The protocol for non-integrated POCT method validation according to CAP and CLIA'88 standards must include accuracy, precision, verification of cut-offs, reportable range and analytical measuring range, POCT inter-instrument comparison and comparison with bench top analyzers placed in the central laboratory(19) (20) (21). Reagent shipments and lot numbers must be validated (22). To determine the appropriate use of non-integrated POCT, an evaluation of each test is necessary to establish the unacceptable turnaround time based on clinical requirements. POCT tests, such as beta hCG and SARS-CoV-2 antigen test are recommended to have a turnaround time of no more than 20 minutes in emergency situations. This allows healthcare professionals to make timely and informed medical decisions.

Management of consumables and reagents should be procured in a cost-effective manner for each POCT site. POCT costing must include the fixed capital cost (instrument, proficiency survey cost, service contract for vendor, ancillary infrastructure, etc.) and variable cost (reagent consumption, internal controls, consumables, cartridges) (10) (23).

POCT POLICIES AND PROCEDURES FOR NON-INTEGRATED POCT

As per CLSI guidelines, a quality management system (QMS) approach must be followed for the development of standards and policies for non-integrated POCT. The laboratory director or designee should take responsibility for QC, QA, and test utilization of non-integrated POCT. Every POCT site that performs non-integrated POCT must have written policies and procedures available at the testing sites. The POCT training curriculum should be outlined for every non-integrated POCT by the pathologists or the subject experts and shared with POCT interdisciplinary committee for approval and feedback. Every non-integrated POCT at the institute, as well as its adherence to legal requirements, must be handled by the central laboratory. The central laboratory is responsible for ensuring that the necessary training, quality control (QC), proficiency testing (PT), and validation processes are carried out, confirmed, and documented initially and then on a regular basis.(24).

It should be ensured that the purpose of POCT, i.e. prompt results for prompt patient management, must not be lost and the processes should be simple and easy to follow (25). As noted by Harvey, the mean turnaround time expected by clinicians managing patient in in critical care areas ranges from 5-15 minutes(26). Hence the policies and processes need to be carefully designed keeping this challenge of turnaround time in mind.

NON-INTEGRATED POCT DATA CAPTURE

Healthcare regulatory bodies and accreditation agencies, such as the CLIA'88, the JCI, CAP, emphasize the importance of monitoring POCT operator competency and instrument quality as these will lead to reliability of results (20) (27) (28). These regulations stress the need for laboratory oversight and review of POCT QC and patient data. The labour and resources that must be devoted to the POCT locations in order to achieve regulatory compliance with quality assurance, including record keeping, archiving,

billing, and data entry into the electronic medical record or LIMS, will increase with manual, non-integrated POCT devices (29).

The advantages of POCT are multiplied when patient and QC results are directly downloaded into a LIMS with minimal human intervention (30) (31). For accreditation and patient safety, trail (manual or electronic) must link each patient result to the POCT user, user's training and competency records, the reagents or cartridges utilized and its validation, and the device validation and maintenance even if manual entries or manual logs must be kept (32). Where possible connectivity of POCT device to POCT data management system and to the LIMS need to be established. Before bringing non-integrated POCT into practice evaluation of data security, processes, risk assessment must be carried out and reviewed by the IT support and POCT teams.

STAFF TRAINING AND COMPETENCY ASSESSMENT

For all non-integrated POCT, a thorough POCT training plan and curriculum will have to be developed in line with the CLIA'88 and CAP standards by the subject experts (for example by chemical pathologists for Beta- Human Chorionic Gonadotropin testing). *"Evaluating the competency of all testing professionals and ensuring that staff maintains their competency to perform test procedures and report test findings promptly, accurately, and competently"* are two CLIA'88 requirements for competency evaluation (33). The purpose of the curriculum would be to identify and control potential serious medical errors attributable to non-integrated POCT. Training curriculum must include all phases of the testing process and consist of (not limited to) the following:

- Direct observation of routine patient test performance
- Testing previously analyzed specimens, internal or external QC samples
- Recording and reporting of patient test results
- Recording and reporting of QC results
- Interpretation of patient test results, QC results
- Demonstration of POCT device maintenance
- Assessment of problem-solving skills

Training and certification of all POCT users on non-integrated devices with no interface with LIMS must be done separately from integrated POCT. The record of training and certification must be available from the POCT coordinator and site supervisors. If possible barcoded identification must be provided to the certified POCT users in the institute. Competency assessment should be performed annually for waived tests or for non-waived tests, after 6 months from the first test on hire and then annually thereafter. Records of competency must be maintained via the online connectivity server or in the form of manual logs.

INDIVIDUALIZED QUALITY CONTROL PLAN (IQCP)

The analytical goals for non-integrated POCT are equivalent to those used for the central laboratory. In order to ensure that the use of non-integrated POCT does not compromise standard of patient care and clinical decision-making, Individualized Quality Control Plan (IQCP) ought to be outlined and followed (34). The proposed IQCP aims to provide clinical laboratories with the framework to implement it when appropriate and offer flexibility to design a QC plan that meets the needs of the laboratory. A process to identify and mitigate errors will be required by each POCT site using the non-integrated POCT devices. The overall intent of IQCP at POCT sites for non-integrated POCT is to help ensure that

clinical laboratories and hospitals remain in compliance with regulations (35). The proposed IQCP covers includes risk assessment, quality control plans and quality assessment monitoring. It demonstrates how laboratories providing oversight to non-integrated POCT can perform a risk assessment to evaluate and record their current quality activities using the IQCP guide, create a quality control plan (QCP) from the risk assessment information, and establish a QA for the test system being evaluated for an IQCP.

IQCP-RISK ASSESSMENT STRATEGY

A risk assessment strategy is required for non-integrated POCT devices with no interface with LIS (36). The process of identifying and evaluating the potential failures and errors has been laid down that could occur during all the phases of POCT testing in Tables 1-3 (37). Risk is, by definition, the product of two factors: the likelihood that harm may occur and its seriousness (38). The goal of the risk assessment is to examine every step of the non-integrated POCT process, from preanalytical to analytic to post analytic, and identify any potential points of error that could endanger the patient if they are not caught. It includes evaluation of the five components of POCT testing: specimen, test system, reagents/cartridges, POCT site environment and POCT users or testing personnel. The table 1 describes the risk assessment including some common sources of errors and solutions encountered in non-integrated POCT program before the actual analysis.

Specimen preparation, reagent handling and test analysis (Table 2) present its own set of risks and problems, which must be identified and mitigated in order to assure the overall safety and reliability of the POCT process. By implementing approaches to reduce the risk of errors and inconsistencies in analysis, POCT can help enhance patient outcomes.

Table 1 Risk assessment for non-integrated POCT in pre-analytical testing phase

Risk Assessment Components	Possible sources of error	Risk Mitigation: Manage/Prevent Errors from Occurring
Patient or specimen identification	Wrongly identified throughout the testing process A functioning barcode reader is not available in the test system to ensure positive patient identification Erroneous patient/specimen information entry	• Train POCT users to follow dual identification criteria and ensure correct patient, correct side and correct POCT • Outline criteria for specimen rejection • Provision of barcoded samples. System should be in place with barcodes to create entry of patient information and test request on LIMS
Specimen collection	Insufficient specimen volume	• POCT system will not perform the test if the specimen volume does not meet the minimum volume requirement
	Incorrect vacutainer or POCT cartridge Wrong labelling Criteria for specimen rejection is missing	• Train POCT users to the correct selection of vacutainer or POCT cartridge • Ensure specimen rejection criteria is present in policy and processes and is taught and assess to POCT users

Table 2 Risk assessment plan for non-integrated POCT in analytical testing phase

Risk Assessment Components	Possible sources of error	Risk Mitigation: Manage/Prevent Errors from Occurring
Testing Personnels' Training and Competency	Not trained POCT user Not competent POCT user Improper training or expired competency records Proper POCT technique/ system maintenance /QC not being accurately followed	• Central laboratory to provide proper initial training and competency • Competency assessment frequency to be decided based on risk assessment • Compliance monitoring • Follow established policies and procedures • Testing personnel training and competency records • Traceability of POCT users
Test system- Calibration	Calibration omitted Calibration out and patient results reported There is no mechanism, such as an operator lockout, to ensure only trained personnel use the test system	• Ensure POCT users follow standard operating procedures through trainings, competency assessment, audits • Test process flow charts • POCT user manual/ electronically shared QC and calibration records with POCT coordinator
Test system- Maintenance	Maintenance procedures not consistent with manufacturer's instructions	• Train POCT users for POCT device maintenance
Test system- Reagent	Expired reagents Wrong lot of reagents	• Train POCT users to check reagent expiry prior to use • Manufacturer's instructions to be followed for reagents storage
Test system- QC	Expired QC Wrong lot of QC Unable to interpret QC Unable to trouble shoot if QC is out	• Prevent QC degradation during storage and use • Training of POCT users on interpreting QC before release of patient results • Maintain QC logs (manual if electronic not possible) • Assure POCT user review each QC before patient result release – through trainings • Training of POCT users on troubleshooting
Environment-room Temperature	Room temperature not appropriate	• Daily room temperature logs • POCT coordinator to sign off logs periodically
Environment- Storage conditions	Sample / reagents/ cartridge/ strips storage conditions not appropriate	• Daily refrigerator temperature logs where POCT items are stored. • POCT coordinator to sign off logs periodically

Table 3 Risk assessment plan for non-integrated POCT in analytical testing phase

Risk Assessment Components	Possible sources of error	Risk Mitigation: Manage/Prevent Errors from Occurring
Test Results reporting and archiving	Incorrect result reading Incorrect result entry Delay in signing out Data security Data storage loss Patient report lost to retrieval No correlation between initially generated/finally recorded result	• Training and competency of POCT users for result reporting and entry into LIS to be ensured by the central lab • Frequency of competency assessment may be increased • Even if instrument interface with LIS is missing provision of manual entry of results should be given to POCT users • Manual records of all tests be entered in log sheets and signed off by central laboratory on periodically • Provision of LIMS data entry should be limited to the POCT on which the POCT user is certified • Missing data audits to be conducted periodically • Periodic scanning/ digitalization of manual patient results' logs
Reference range or cutoff	Incorrect Reference range or cutoff Incorrect units	• Reference range or cutoff must be verified by the central lab • Reference range/ cutoff must be built in LIS • Units on reports must be checked by the clinical lab, built in LIS
Interpretation	Incorrect interpretation	• POCT report includes interpretation verified by the central laboratory and built-in LIS • Distinction that the test was not done in central lab on patient report
	False positive and false negative results	• The report includes a disclaimer regarding the false negatives and false positive causes to aid clinical interpretation
	Critical results not informed	• Ensure POCT users read/ follow critical results policy • Critical results policy should be readily available at POCT sites

The significance of the post-analytical phase in POCT stems from its ability to influence clinical decision-making and patient outcomes. The accuracy and reliability of the test results are checked during this phase, and the results are interpreted in the context of the patient's clinical history and current state. Incorrect or insufficient interpretation of test data might result in erroneous treatment decisions or therapy delays, both of which can have a detrimental impact on patient outcomes. Table 3 shows the risk assessment plan for non-integrated POCT during post-analytical testing phase.

Once an IQCP is developed and is acceptable with no risk to the patients then the non-integrated POCT should be introduced in clinical areas. Regular IQCP audits and risk management must follow starting with specimen and patient identification, specimen collection, specimen container, transport, etc., and moving through the other components, the lab would identify and list potential places where mistakes could occur, evaluating each one for risk of error. A historical analysis is necessary for the risk management. Risk management for non-integrated POCT can be a proactive project to identify potential flaws in new, altered, or complex processes, a reactive project to respond to an incidence or finding, or a continuous assessment based on daily events and observation. Non-integrated POCTs can involve several complex processes that require specialized knowledge, expertise, and training to ensure accurate and reliable results. Some of such complex processes are as follows:

- Some non-integrated POCTs require specialized knowledge to interpret results accurately, particularly rapid tests for infectious diseases like influenza or COVID-19, which require experience in detecting subtle changes in color or signal intensity to determine positive or negative results.
- For non-integrated POCT devices, such manual QC procedures may add extra work as they may involve manual checks of equipment performance, tracking reagent quality, and regular checks to maintain consistency and reliability of test results.
- POCTs generate significant amounts of data that require management and tracking. This data includes patient identification, test results, quality control data, and instrument maintenance logs. Manual data management is essential in non-integrated POCT
- Complying with these regulations can be more complicated and may require specialized knowledge and expertise, particularly for non-integrated POCT devices.

The manufacturers' package inserts, pertinent policies and procedures, QC, corrected reports, physician complaints, employee training and competency records, PT results, and temperature records are the documents that must be reviewed periodically. Table 4 is an illustration of a practical checklist that can be used, particularly when writing the IQCP for non-integrated POCT. The review of historical non-integrated POCT data will then determine the frequency of occurrence of errors and the impact of harm to a patient.

Table 5 shows a template of the 'Risk Matrix' for the non-integrated POCT that can be followed. This will determine if the non-integrated POCT can be continued or should be removed from the POCT Program.

IQCP-QUALITY CONTROL PLAN

A QCP should be outlined by the central laboratory for each non-integrated POCT device describing the practices and procedures to reduce the chance of possible failures and errors in the test processes. The QCP must ensure that the accuracy and reliability of test results

Table 4 Checklist for IQCP risk assessment for non-integrated POCT

Review	Available	Not Available	Not Applicable
Need of non-integrated POCT justified			
Process Map			
Manufacturer QC requirements			
Manufacturer or package inserts			
Manufacturer alerts			
QC Certificates			
Calibration			
Method Validation/ Verification			
Policies and Processes			
Historical QC data			
Proficiency survey results and corrective action documentation			
POCT testing personnel Training and competency records			
Instrument maintenance logs			
Temperature charts			
Specimen rejection logs			
Physician/client complaints			

Table 5 Risk matrix example to assess severity of harm from non-integrated POCT

Risk Category	**Occurrence of Errors**	**Severity scale for probability of harm**				
		Negligible	**Minor**	**Serious**	**Critical**	**Catastrophic**
Expired Reagents use	Frequent (Once per week)					
	Probable (Once per month)					
	Occasional (Once per year)					
	Remote (Once every few years)					
	Improbable (Once in the life of the measuring system)					

Key:

Unshaded is Acceptable, shaded is Unacceptable.

Severity scale for probability of harm:

- Negligible: Could result in inconvenience or temporary discomfort
- Minor: Could result in temporary injury or impairment not requiring professional medical intervention
- Serious: Could result in injury or impairment requiring professional medical intervention
- Critical: Could result in permanent impairment or life-threatening injury
- Catastrophic: Could result in patient death

from non-integrated POCT, are appropriate for patient care. The QCP for each non-integrated POCT may at least include, electronic controls, internal QC, external QC or PT, calibration, maintenance and training and competency assessment (6). The main QC requirements must be addressed which include the following:

- Procedure established for internal QC (39)
- Internal QC material procurement
- Correction of nonconformities and availability of trouble shooting guide
- PT processes and policies
- Periodic comparison of results from non-integrated POCT device and the gold standard or working instrument for same analyte placed in central laboratory.
- Comparison of results and performance across different POCT sites
- Sub-optimal performance in internal QC and/ or PT to be brought to the immediate attention of the POCT committees

For analytes for which PT surveys are not available or are not accessible, an in-house scheme can be established using split patient samples (40). In split patient each specimen can be split and analyzed in the same manner with the non-integrated POCT method and then with the central laboratory method or another POCT site or by another POCT user. If findings agree within the analyte's allowed performance range, bias between results can be determined and reviewed for acceptance. Criteria for acceptance can be obtained from literature or published guidelines or using ± 2 or 3 standard deviations from the mean from QC data for quantitative assays (41).

IQCP-QUALITY ASSURANCE

For continuous monitoring of the QCP effectiveness for non-integrated POCT a QA plan should be in place (42). Practices, processes, and resources to consider for monitoring effectiveness of a QCP must include clinical audits and review of the following:

- Policies and standard operating procedures
- Logs of training and competence assessment
- Logs of internal QC reviews
- PT performance reviews
- Turnaround time reports
- Logs of critical results informed
- Complaint reports
- Logs of maintenance
- Logs of breakdowns

All POCT programs need to be observed and evaluated periodically to assure that the program is meeting the needs of patients, testing personnel and hospital. All POCT sites must be periodically audited and assessed for compliance of policy, procedure, and protocols, along with POCT users' knowledge, skills, and practices (43).

CONCLUSIONS

In low resource healthcare settings, our approach to non-integrated POCT involves both direct and indirect supervision by the clinical laboratory. This comprehensive model ensures effective POCT management. Major recommendations included in the current proposed framework are taking a formalized approach to POCT within the facility, use of written policies, standard operating procedures, forms, and logs, POCT end user training, including periodic competency assessments, POCT devices performance evaluation and use of both statistical QC and PT programs, use of properly established or validated reference intervals or cutoffs and ensuring accurate patient results reporting. This paradigm may aid with the delivery of diagnostic services to low resource centers in the absence of a pathology

laboratory and may satisfy the demands of POCT specialists, particularly in developing nations. The suggested framework emphasizes how crucial it is for non-integrated POCT to have clear policies and procedures in place to guard or gatekeep patient safety.

REFERENCES

1. Luppa PB, Müller C, Schlichtiger A, Schlebusch H. Point-of-care testing (POCT): Current techniques and future perspectives. TrAC Trends Anal Chem. 2011 Jun 1;30(6): 887–98.

2. Quinn AD, Dixon D, Meenan BJ. Barriers to hospital-based clinical adoption of point-of-care testing (POCT): A systematic narrative review. Crit Rev Clin Lab Sci. 2016 Jan 2;53(1):1–12.

3. Collinson P. Cardiac biomarkers by point-of-care testing – back to the future? J Lab Med. 2020 Apr 1;44(2):89–95.

4. Singapore S. Point-of-care Testing (POCT): Requirements for Quality and Competence. ISO; 2016.

5. Shaw JL. Practical challenges related to point of care testing. Practical laboratory medicine. 2016 Apr 1;4:22-9.

6. Martin CL. Quality Control Issues in Point of Care Testing. Clin Biochem Rev. 2008 Aug;29(Suppl 1):S79–82.

7. Jones BA, Meier FA. Patient safety in point-of-care testing. Clin Lab Med. 2004;24(4):997–1022.

8. O'Leary D. College of American Pathologists Conference XXVIII on alternate site testing: keynote address. Global view of how alternate site testing fits in with medical care. Arch Pathol Lab Med. 1995 Oct;119(10):877–80.

9. Nichols JH. The National Academy of Clinical Biochemistry Laboratory Medicine Practice Guidelines: evidence-based practice for point of care testing. Point Care. 2007;6(4):213–4.

10. Khan AH, Shakeel S, Hooda K, Siddiqui K, Jafri L. Best practices in the implementation of a point of care testing program: experience from a tertiary care hospital in a developing country. EJIFCC. 2019 Oct;30(3):288.

11. Bardach NS, Huang J, Brand R, Hsu J. Evolving health information technology and the timely availability of visit diagnoses from ambulatory visits: A natural experiment in an integrated delivery system. BMC Med Inform Decis Mak. 2009 Jul 17;9(1):35.

12. Wiencek J, Nichols J. Issues in the practical implementation of POCT: overcoming challenges. Expert Rev Mol Diagn. 2016 Apr 2;16(4):415–22.

13. Kosack CS, Page AL, Klatser PR. A guide to aid the selection of diagnostic tests. Bull World Health Organ. 2017 Sep 1;95(9):639–45.

14. Warade J. POCT: Taking control in uncontrolled premises, part 2. 2016.

15. Šprongl L. Quality assurance for the POCT systems. In-Quality Management and Practices 2012 Apr 27. IntechOpen.

16. Poe S, Nichols J. Quality assurance, practical management, and outcomes of point-of-care testing: Nursing perspectives, Part II. Clin Lab Manage Rev. 2000;14(1):12–8.

17. Warade J. Organization of the POCT Unit. EJIFCC. 2015 Mar 10;26(2):125–32.

18. Klein RD, Campbell S, Howanitz PJ. CAP Point of Care Checklist: Frequently Asked Questions. Point Care. 2005; 4(2):75–85.

19. Farmer T. Toward a culture shift in laboratory quality: application of the full ISO 15189 standard. MLO Med Lab Obs. 2015 May;47(5):38–9.

20. EP15 C. A3: User verification of precision and estimation of bias; approved guideline. 2014.

21. Westgard JO, Burnett RW. Precision requirements for cost-effective operation of analytical processes. Clin Chem. 1990 Sep;36(9):1629–32.

22. Nichols JH. Quality in point-of-care testing. Expert Rev Mol Diagn. 2003 Sep 1;3(5):563–72.

23. Broughton PM, Woodford FP. Benefits of costing in the clinical laboratory. J Clin Pathol. 1983 Sep 1;36(9): 1028–35.

24. Dyhdalo KS, Howanitz PJ, Wilkinson DS, Souers RJ, Jones BA. Documentation of quality control and operator training at point-of-care testing: a College of American Pathologists Q-Probes study of 106 institutions. Arch Pathol Lab Med. 2014;138(11):1444–8.

25. Lewandrowski K. Point-of-care testing: an overview and a look to the future (circa 2009, United States). Clin Lab Med. 2009;29(3):421–32.

26. Harvey M. Point-of-care laboratory testing in critical care - ProQuest. Am J Crit Care. 1999 Mar;8(2):72.

27. Medicare M. CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA)-HCFA. Final rule with comment period. Fed Regist. 1992;57(40):7002–186.

28. Ehrmeyer SS, Laessig RH. Regulatory requirements (CLIA '88, JCAHO, CAP) for decentralized testing. Am J Clin Pathol. 1995 Oct 1;104(4 Suppl 1):S40-9.

29. Salka L, Kiechle FL. Connectivity for point-of-care glucose testing reduces error and increases compliance. Point Care. 2003;2(2):114–8.

30. Quinn AD, Dixon D, Meenan BJ. Barriers to hospital-based clinical adoption of point-of-care testing (POCT): a systematic narrative review. Crit Rev Clin Lab Sci. 2016; 53(1):1–12.

31. Schimke I. Quality and timeliness in medical laboratory testing. Anal Bioanal Chem. 2009 Mar;393(5):1499–504.

32. Lone R, Sequeira S, Shantakumari N. Compliance with Accreditation and Standardization of Point of Care Testing-from Vision to Action. J Int Dent Med Res. 2021; 14(2):860–4.

33. Department of Health and Human Services, Centers for Medicare and Medicaid Services. Clinical laboratory improvement amendments of 1988; final rule. Fed Regist. 2003 Oct 1;1065-66 [42CFR493.1453] and 1053-4 [42CFR493.1413].

34. James HN. Risk Management for Point-of-Care Testing. EJIFCC. 2014 Sep 4;25(2):154–61.

35. Briggs C, Guthrie D, Hyde K, Mackie I, Parker N, Popek M, Porter N, Stephens C. Guidelines for point-of-care testing: haematology. British journal of haematology. 2008 Sep;142(6):904-15.

36. Bruno LC. IQCP: Guideline and Helpful Tools for Implementation. Lab Med. 2016;47(4):e42–6.

37. Nichols JH, Christenson RH, Clarke W, Gronowski A, Hammett-Stabler CA, Jacobs E, et al. Executive summary. The National Academy of Clinical Biochemistry Laboratory Medicine Practice Guideline: Evidence-based practice for point-of-care testing. Clin Chim Acta. 2007 Apr 1; 379(1):14–28.

38. Aita A, Padoan A, Antonelli G, Sciacovelli L, Plebani M. Patient safety and risk management in medical laboratories: theory and practical application. J Lab Precis Med. 2017;2(2):75–75.

39. Holt H, Freedman DB. Internal quality control in point-of-care testing: where's the evidence? Ann Clin Biochem. 2016;53(2):233–9.

40. Guiñón L, García-Villoria J, Ribes A, Gort L, Molina A, Soler A, et al. External quality assessment in the absence of proficiency testing: A split-sample testing program experience. Clin Biochem. 2021 Nov 1;97:78–81.

41. Nichols JH, Alter D, Chen Y, Isbell TS, Jacobs E, Moore N, Shajani-Yi Z. AACC guidance document on management of point-of-care testing. The Journal of Applied Laboratory Medicine. 2020 Jul;5(4):762-87.

42. Wang Q, Du J, Chen L, Du YD, Luo W. Comparison of POCT glucose meters and analysis of the interference factor. Journal of Laboratory Medicine. 2022 Jun 27;46(3): 195-201.

43. 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.

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Pooled analysis of diagnostic performance of the instrument-read Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA)

Giuseppe Lippi[1,2], Brandon M. Henry[1,3], Mario Plebani[1,4]

[1] *IFCC Task Force on COVID-19, Milano, Italy and IFCC WG on SARS-CoV-2 variants.*

[2] *Section of Clinical Biochemistry and School of Medicine, University of Verona, Verona, Italy*

[3] *Clinical Laboratory, Division of Nephrology and Hypertension, Cincinnati Children's Hospital Medical Center, Cincinnati OH, USA*

[4] *Department of Medicine, University of Padova, Padova, Italy*

ARTICLE INFO

Corresponding author:
Prof. Giuseppe Lippi
Section of Clinical Biochemistry
University Hospital of Verona
Piazzale L.A. Scuro, 10
37134 Verona
Italy
Phone: 0039-045-8122970
Fax: 0039-045-8124308
E-mail: giuseppe.lippi@univr.it

Key words:
SARS-CoV-2, COVID-19, immunoassay, diagnosis, antigen

ABSTRACT

Background

This article presents a critical literature review and meta-analysis of diagnostic performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA), a rapid diagnostic antigen test (RDT-Ag) adapted for automatic reading with portable instruments, thus potentially combining the advantages of point-of-care testing with those of a laboratory-based immunoassay.

Methods

We conducted an electronic search in PubMed and Scopus with the keywords "Quidel" OR "SOFIA" AND "Antigen" AND "SARS-CoV-2" OR "COVID-19" up to March 24, 2023, for identifying articles containing data on accuracy of Quidel Sofia SARS antigen FIA for diagnosing acute SARS-CoV-2 infections. We selected those where test accuracy was compared to that of

a reference SARS-CoV-2 molecular assay, and with sufficient information for constructing a 2×2 table.

Results

A total number of 18 articles (48165 samples; 9.8% positive at molecular testing) were included in this meta-analysis, averaging 24 sample cohorts. The diagnostic accuracy (summary area under the curve), sensitivity and specificity were 0.980, 0.76 and 1.00 in all samples, 0.981, 0.81 and 0.99 in samples collected from symptomatic patients, 0.931, 0.55 and 1.00 in those taken from asymptomatic patients, and 0.960, 0.77 and 0.99 in samples from mixed cohorts of patients, respectively. Minor and clinically negligible differences of accuracy could be found by comparing test results in nasal and nasopharyngeal swabs.

Conclusion

Quidel Sofia SARS Ag FIA meets the minimum performance criteria of accuracy for SARS-CoV-2 antigenic testing, thus combining satisfactory diagnostic performance with the advantages of being potentially used as a portable device.

❖❖❖❖❖

INTRODUCTION

Three years after the World Health Organization (WHO) declared the outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) a pandemic, coronavirus disease 2019 (COVID-19) is still considered a public health emergency of international concern [1]. This is mostly due to the fact that the number of infections continues to grow irrespective of immunity and environmental conditions, thus no longer following the typical seasonal pattern that has characterized the early phase of the pandemic [2]. Along with a constant number of daily infections comes the still relevant impact that COVID-19 has on the most vulnerable parts of the population, especially comprising older people, immunocompromised patients, and those with underlying health conditions such as cancer, cardiovascular and pulmonary diseases, diabetes, obesity, and other chronic illness [3].

According to the WHO [4], a confirmed case of SARS-CoV-2 infection could be an individual with (i) a positive test result of a nucleic acid amplification test (NAAT) irrespective of other clinical or epidemiological criteria, or (ii) a positive test result of a professional used or self-test SARS-CoV-2 antigen (Ag) assay, meeting specific clinical (i.e., being symptomatic) or epidemiological (i.e., being a contact of a COVID-19 case or directly linked to a cluster) criteria. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) has recently endorsed similar recommendations, stating that the diagnosis of an acute SARS-CoV-2 infection can be made by either molecular or Ag testing, reserving the use of the second approach to specific clinical settings (i.e., especially in those at lower risk of having an acute SARS-CoV-2 infection or for specific epidemiological purposes) [5]. Two recent economic analyses revealed that an approach based on sequential testing (SARS-CoV-2 Ag testing first, followed by NAAT in those testing negative) is not only clinically safe, but also is more cost-effective than molecular testing alone [6,7]. As concerns the specific diagnostic performance of SARS-CoV-2 Ag testing, both the WHO [8] and the IFCC [5,9] mandate that minimum performance criteria shall be met by SARS-CoV-2 Ag immunoassays, either rapid diagnostic tests (RDT-Ag) or laboratory based, in that they should display ≥0.80 sensitivity and ≥0.97 specificity, respectively, when used in suspected COVID-19 cases (i.e.,

symptomatic subjects). Recent literature review revealed that although most laboratory-based tests seem to fulfil these performance limits [10], the diagnostic accuracy of RDT-Ag varies broadly, with average sensitivity of 0.73 (95%CI, 0.69-0.76) in symptomatic subjects, decreasing to 0.55 (95%CI, 0.48-0.62) in those without symptoms [11]. Importantly, according to the Cochrane COVID-19 Diagnostic Test Accuracy Group, the vast majority of tests failed to meet the WHO and IFCC minimum sensitivity criterion of ≥0.80, thus raising serious doubts about their reliability and safety [11].

The diagnostic sensitivity of all SARS-CoV-2 Ag tests is influenced by a widely heterogeneous analytical sensitivity (i.e., the limit of detection; LoD) [12], as well as by a kaleidoscope of pre-analytical and post-analytical variables [13], among which accuracy of test reading and interpretation play the lion's share [14]. Thus, the possibility to standardize and/or automate this last but highly relevant step of RDT-Ag performance now allowed by some commercial tests may help eliminate a very important source of variability in test performance.

For this purpose, the aim of this investigation is to provide a critical literature review and meta-analysis of the diagnostic performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA), a widely used RDT-Ag immunoassays adapted for being automatically read by a portable instrument, thus potentially combining the advantages of point-of-care (POC) testing with those of a laboratory-based immunoassay.

MATERIALS AND METHODS

Assay description

The Quidel Sofia SARS antigen Fluorescent Immunoassay has been specifically developed for qualitative detection of SARS-CoV-1 and SARS-CoV-2 nucleocapsid (n) protein. The test, included within the category of lateral flow immunofluorescent sandwich assays, has been specifically adapted for use with the portable Sofia, Sofia 2 and Sofia Q analyzers, thus enabling to achieve objective and automated test results within 15 min. According to manufacturer's indications, the assay should be specifically used for SARS-CoV-2 testing using direct nasal swabs collected from symptomatic patients within the first 5 days of symptoms onset, or for serial testing of asymptomatic patients (in such cases within 24-36 hours between repeated tests). The test has been cleared for being used as a POC, under a CLIA Certificate of Waiver, Certificate of Compliance, or Certificate of Accreditation.

The test sample is initially placed in a reagent tube (i.e., the swab is rotated for at least 3 times, pressing the head against the bottom and side of the tube for enabling optimal mixing with the buffer) for disrupting viral particles (thus enabling nucleoproteins exposition). A fixed sample volume (i.e., 120 uL) is then pipetted into a test cassette sample well, from where the sample migrates throughout the test strip. In the "WALK AWAY Mode" the cassette is immediately inserted into the portable analyzer, where test results could be displayed after 15 min, whilst in the "READ NOW Mode" the cassette in maintained outside of the analyzer for 15 min, then inserted and immediately read (i.e., within 1 min). When either SARS-CoV-1 or SARS-CoV-2 viral N antigens are present (the test does not differentiate between the two coronaviruses), they are sequestered within a specific site. The analyzer then scans the test strip and measures the fluorescent signal, transforming the fluorescent measure in antigen concentration by means of a method-specific algorithm.

Search strategy

We planned an electronic search in Medline (PubMed interface) and Scopus, using the keywords "Quidel" OR "SOFIA" AND "Antigen" AND "SARS-CoV-2" OR "COVID-19" in all search fields, without language or time constrains (i.e., up to March 24, 2023), for identifying published articles that contained data on accuracy of Quidel Sofia SARS antigen FIA for diagnosing COVID-19. Two authors (G.L. and B.M.H.) screened all articles originally detected based on the predefined search criteria, selecting those with the following inclusion criteria: (i) Quidel Sofia SARS antigen FIA diagnostic performance was compared versus a reference molecular technique; (ii) data on true positive (TP), true negative (TN), false positive (FP) and false negative (FN) rates could be extracted from the text of the article, or could be otherwise provided by the authors after direct request (i.e., by emailing the corresponding authors).

After extraction, data were used for constructing a 2×2 table, which enabled the estimation of pooled accuracy (based on a Summary Receiver Operating Characteristic Curve; SROC), sensitivity and specificity with their respective 95% confidence interval (95%CI). Separate analyses were conducted according to the respiratory sample type (i.e., nasal or nasopharyngeal swab) and the population enrolled (asymptomatic, symptomatic, mixed). The Mantel-Haenszel test and random effects model were used for finally pooling the data, while the heterogeneity was calculated with χ^2 test and I^2 statistics. The statistical analysis was performed with Meta-DiSc 1.4 (Unit of Clinical Biostatistics team of the Ramón y Cajal Hospital, Madrid, Spain) [15].

This analysis was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA Checklist, available as Supplementary File 1), conducted in agreement with the Declaration of Helsinki and within the terms of local legislation. No ethical committee approval was required for performing this critical literature review and meta-analysis.

RESULTS

Our digital search in PubMed and Scopus based on the aforementioned criteria allowed to initially identify 70 articles after eliminating redundancy between the two scientific databases. We then excluded 52 articles, for the following reasons: 36 studies which did not report any data on diagnostic testing, 5 were unsuitable for constructing the 2×2 table (including no response after delivering a specific request to the authors for the data), 6 were literature reviews, 2 did not contain specific data on Quidel Sofia SARS antigen FIA, 2 were focused on performance of SARS-CoV-2 antigen manual assay, and 1 that reported data on a duplicate cohort included in a large subsequent investigation. Thus, a total of 18 articles (totalling 48165 samples; range, 43-23462; 9.8% NAAT positive) meeting our inclusion criteria were finally included in this meta-analysis, equating to 24 sample cohorts (Table 1) [16-33]. Specifically, 4 studies included mixed cohorts of asymptomatic and symptomatic subjects, 5 included two separate cohorts of asymptomatic or symptomatic patients, 7 included only symptomatic patients, and 2 studies included only asymptomatic subjects. As concerns the type of the sample, one study included a single cohort of patients with double sample collection (i.e., nasal and nasopharyngeal), in 16 cohorts only a nasal swab was collected and in 6 cohorts a single nasopharyngeal swab was taken.

The overall diagnostic performance of Quidel Sofia SARS antigen FIA in all samples (i.e., nasal and/or nasopharyngeal) is summarized in figure 1 and table 2, displaying 0.980 (with 0.01

SE) area under the curve (AUC), 0.76 (95%CI, 0.74-0.78; I^2, 95%) sensitivity and 1.00 (95%CI, 1.00-1.00; I^2, 86%) specificity. The corresponding values of AUC, sensitivity and specificity in the reference nasal swab were 0.987 (with 0.01 SE), 0.72 (95%CI, 0.69-0.75; I^2, 89%) and 1.00 (95%CI, 1.00-1.00; I^2, 81%). In samples taken from symptomatic cohorts (Figure 2), the cumulative AUC (0.981 with 0.02 SE) and sensitivity (0.81; 95%CI, 0.77-0.83; I^2, 22%) were predictably higher, whilst the specificity remained almost unvaried (0.99; 95%CI, 0.99-0.99; I^2, 0%). Nearly identical results were found when limiting the analysis to the reference nasal swab, displaying 0.963 (with 0.05 SE) AUC, 0.80 (95%CI, 0.77-0.83; I^2, 0%) sensitivity and 0.99 (95%CI, 0.99-1.00; I^2, 0%) specificity. These performances obviously decreased in samples taken from asymptomatic subjects (Figure 3), AUC being 0.931 (with 0.01 SE), 0.55 (95%CI, 0.48-0.61; I^2, 93%) the sensitivity and 1.00 (95%CI, 1.00-1.00; I^2, 89%) the specificity. Using the nasal reference sample collected from asymptomatic subjects the AUC was 0.888 (with 0.07 SE), the sensitivity 0.45 (95%CI, 0.37-0.52; I^2, 93%) and the specificity 1.00 (95%CI, 1.00-1.00; I^2, 91%). Finally, in the four studies which included mixed cohorts of asymptomatic and symptomatic patients (all except one using nasopharyngeal samples; and study excluded due to lack of negative controls) (Figure 4), the AUC was 0.960 (with 0.03 SE), the sensitivity 0.77 (95%CI, 0.75-0.80; I^2, 99%) and the specificity 0.99 (95%CI, 0.99-1.00; I^2, 93%). Table 3 synthesizes the diagnostic performance

Table 1 **Summary of the characteristics of the studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections**

Study	Country	Sample matrix	Sample size	Population	Reference test
Alonaizan et al., 2022 [16]	Saudi Arabia	Nasal swab	76	Asymptomatic	RT-PCR (Cepheid GeneXpert GX-XVI SARS-CoV-2)
Alonaizan et al., 2022 [16]	Saudi Arabia	Naso-pharyngeal swab	76	Asymptomatic	RT-PCR (Cepheid GeneXpert GX-XVI SARS-CoV-2)
Bachman et al., 2021 [17]	USA	Nasal swab	170	Symptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Beck et al., 2021 [18]	USA	Nasal swab	346	Symptomatic	RT-PCR (Hologic Aptima Panther SARS-CoV-2 TMA test)
Bornemann et al., 2022 [19]	Germany	Naso-pharyngeal swab	7859	Asymptomatic + symptomatic	RT-PCR (Multiple assays)

Černila et al., 2023 [20]	Slovenia	Naso-pharyngeal swab	804	Asymptomatic + symptomatic	RT-PCR (unspecified)
Černila et al., 2023 [20]	Slovenia	Naso-pharyngeal swab	132	Symptomatic	RT-PCR (unspecified)
Epling et al., 2022 [21]	USA	Nasal swab	117	Symptomatic	RT-PCR (unspecified)
Ford et al., 2021 [22]	USA	Nasal swab	865	Asymptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Ford et al., 2021 [22]	USA	Nasal swab	266	Symptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Freeman et al., 2022 [23]	USA	Nasal swab	138	Asymptomatic	RT-PCR (Cepheid Xpert Xpress SARS-CoV-2)
Freeman et al., 2022 [23]	USA	Nasal swab	249	Symptomatic	RT-PCR (Cepheid Xpert Xpress SARS-CoV-2)
Hahn et al., 2021 [24]	USA	Naso-pharyngeal swab	60	Asymptomatic + symptomatic	RT-PCR (New York SARS-CoV-2 RT-PCR)
Harmon et al., 2021 [25]	USA	Nasal swab	23462	Asymptomatic	RT-PCR (Multiple assays)
Harris et al., 2021 [26]	USA	Nasal swab	885	Symptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Jääskeläinen et al., 2021 [27]	Finland	Nasal swab	148	Symptomatic	RT-PCR (In-house)
Mack et al., 2021 [28]	USA	Naso-pharyngeal swab	10982	Asymptomatic + symptomatic	RT-PCR (Multiple assays)
Mitchell et al., 2021 [29]	USA	Nasal swab	144	Asymptomatic	RT-PCR (Cepheid Xpert Xpress SARS-CoV-2)
Mitchell et al., 2021 [29]	USA	Nasal swab	104	Symptomatic	RT-PCR (Cepheid Xpert Xpress SARS-CoV-2)
Porte et al., 2021 [30]	Chile	Naso-pharyngeal swab	64	Symptomatic	RT-PCR (Primerdesign COVID-19 Genesig)

Pray et al., 2021 [31]	USA	Nasal swab	871	Asymptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Pray et al., 2021 [31]	USA	Nasal swab	53	Symptomatic	RT-PCR (CDC 2019-nCoV RT-PCR Diagnostic Panel)
Smith et al., 2021 [32]	USA	Nasal swab	43	Asymptomatic + symptomatic	RT-PCR (Abbott Alinity)
Young et al., 2020 [33]	USA	Nasal swab	251	Symptomatic	RT-PCR (BD MAX real-time SARS-CoV-2 PCR assay)

Figure 1 **Summary of the diagnostic performance (area under the curve [AUC], sensitivity and specificity) of the studies which cumulatively explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections**

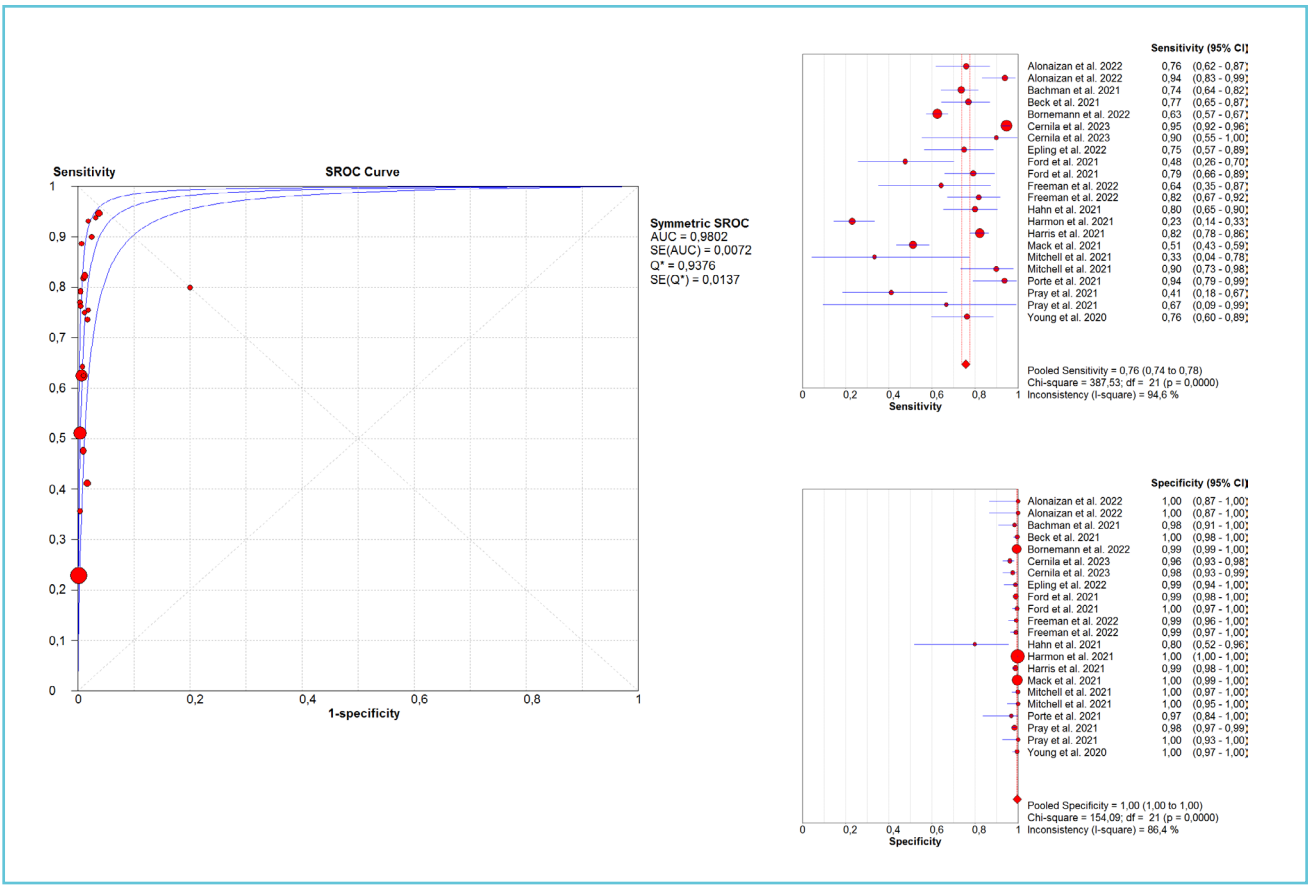


Figure 2 **Summary of the diagnostic performance (area under the curve [AUC], sensitivity and specificity) of the studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections in samples taken from symptomatic subjects**

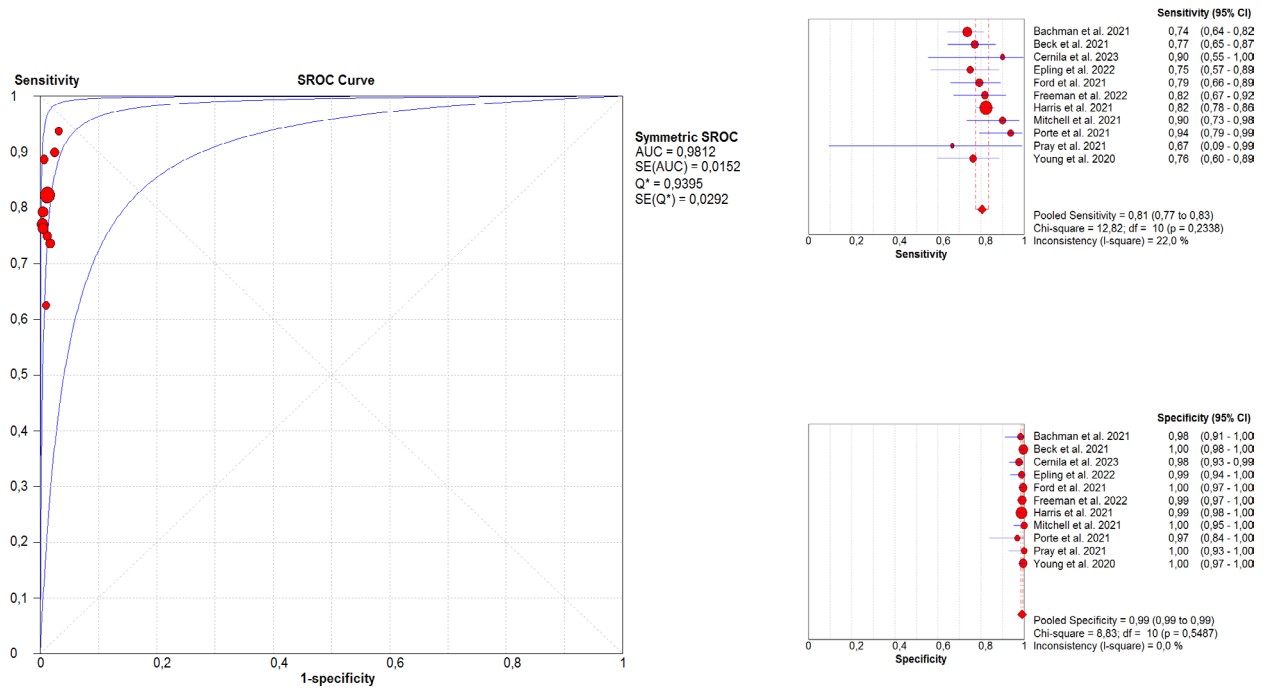


Table 2 **Summary of the diagnostic performance of the studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections**

Cohort	AUC (SE)	Sensitivity (95%CI)	Specificity (95%CI)
All samples	0.980 (0.01)	0.76 (0.74-0.78)	1.00 (1.00-1.00)
All samples (nasal swab)	0.987 (0.01)	0.72 (0.69-0.75)	1.00 (1.00-1.00)
Symptomatic patients	0.981 (0.02)	0.81 (0.77-0.83)	0.99 (0.99-0.99)

Symptomatic patients (nasal swab)	0.963 (0.05)	0.80 (0.77-0.93)	0.99 (0.99-1.00)
Asymptomatic patients	0.931 (0.01)	0.55 (0.46-0.61)	1.00 (1.00-1.00)
Asymptomatic patients (nasal swab)	0.888 (0.07)	0.45 (0.37-0.52)	1.00 (1.00-1.00)
Mixed cohorts	0.960 (0.03)	0.77 (0.75-0.80)	0.99 (0.99-1.00)

Figure 3 **Summary of the diagnostic performance (area under the curve [AUC], sensitivity and specificity) of the studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections in samples taken from asymptomatic subjects**

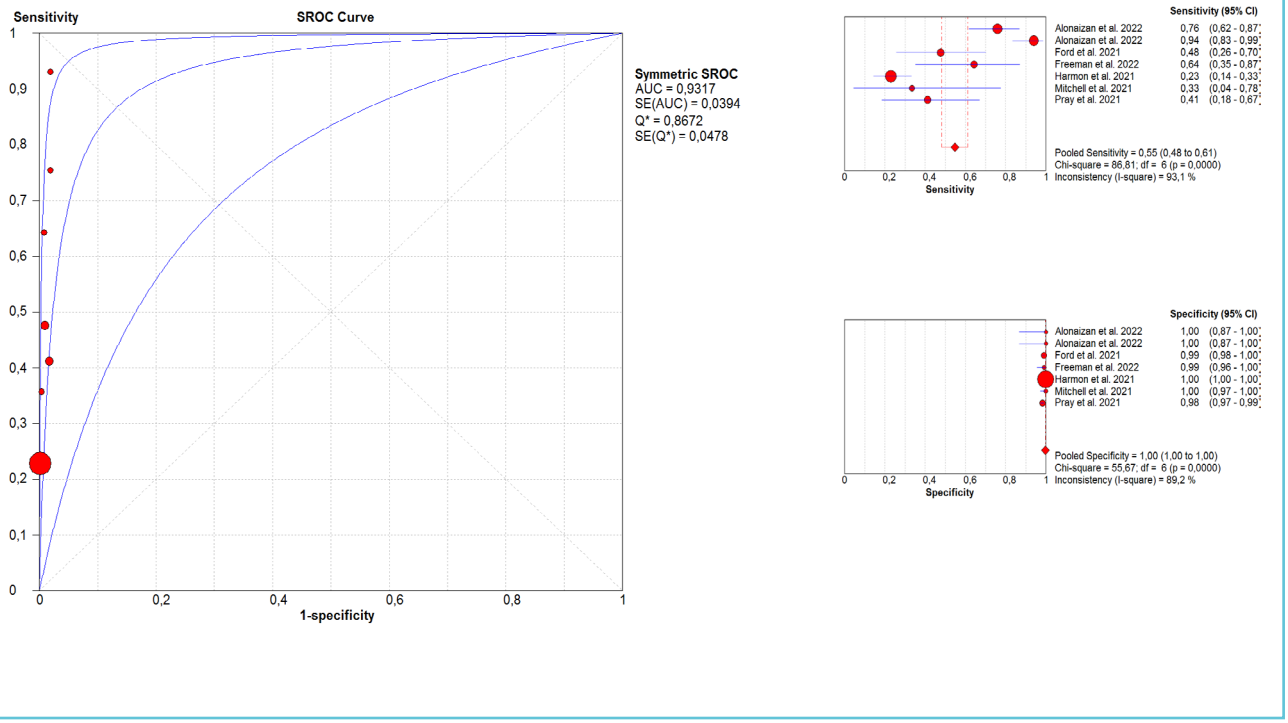


Figure 4 Summary of the diagnostic performance (area under the curve [AUC], sensitivity and specificity) of the studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections in samples taken from mixed cohort of asymptomatic and symptomatic subjects

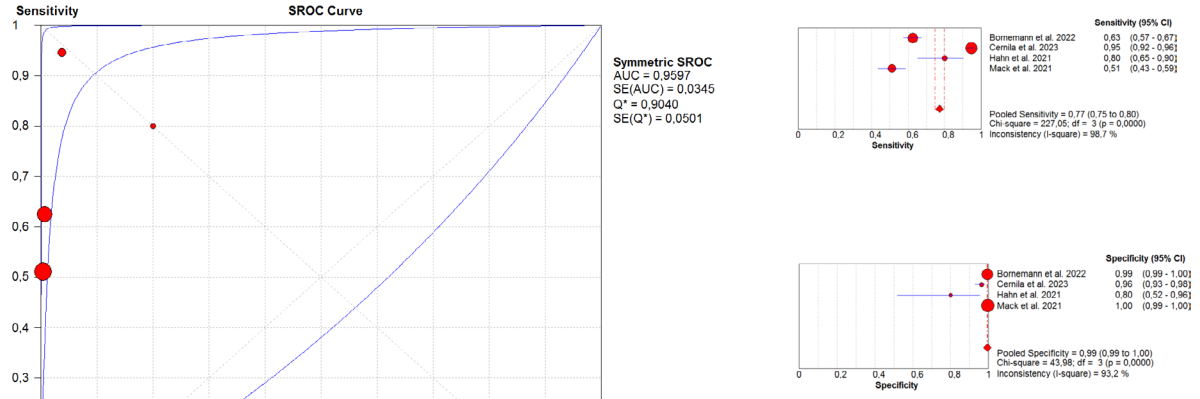


Table 3 Synthesis of the diagnostic performance of studies which explored the performance of Quidel Sofia SARS antigen Fluorescent Immunoassay (FIA) for diagnosing acute SARS-CoV-2 infections and ought to be excluded from the meta-analysis due to unavailability of data for constructing a 2x2 table

Authors	Cohort	Sensitivity (95% CI)	Specificity (95% CI)
Agard et al., 2022 [34]	Low-risk	0.26 (-)	1.00 (-)
Agard et al., 2022 [34]	High-risk	0.37 (-)	1.00 (-)
Al-Alawi et al., 2021 [35]	Symptomatic patients	0.64 (0.50-0.77)	0.97 (0.95-0.98)

Brihn et al., 2021 [36]	Asymptomatic patients	0.60 (0.50-0.71)	1.00 (0.99-1.00)
Brihn et al., 2021 [36]	Symptomatic patients	0.72 (0.61-0.83)	0.99 (0.97-1.00)
Schroeder et al., 2022 [37]	Asymptomatic patients	0.60 (0.45-0.71)	-
Schroeder et al., 2022 [37]	Symptomatic patients	0.77 (0.56.-0.85)	-

of four other studies [34-37] which reported quantitative data on the diagnostic accuracy of Quidel Sofia SARS antigen FIA, but were excluded due to unavailability of sufficient information for constructing a 2×2 table.

DISCUSSION

Due to the ongoing surge of infections and the predictable transformation of COVID-19 into an endemic disease, SARS-CoV-2 testing remains of paramount importance for a variety of reasons beyond diagnosing an acute viral infection, thus including the anticipation of local outbreaks [38], predicting future pressure on healthcare systems [39], and timely detection of changes in viral biology and its interaction with the host (i.e., emergence of new variants) [40]. In this problematic scenario, the availability of easy, rapid, affordable, and reliable tests is central to the paradigm for the future management of COVID-19.

Despite recent endorsements by both the WHO and IFCC, which paved the way to diffuse usage of SARS-CoV-2 Ag testing at the population level, concerns have grown as to whether most of these rapid tests would display sufficient accuracy for being used for screening, especially in symptomatic subjects. The recent meta-analysis of the Cochrane COVID-19 Diagnostic Test Accuracy Group revealed that even in high-risk (i.e., symptomatic) populations, the accuracy of such tests is extremely heterogeneous, exhibiting a pooled diagnostic accuracy of 0.76 (95%CI, 0.70-0.81), that only approximates the minimum performance criterion of ≥0.80 set by the WHO even at the upper limit of the 95%CI [11], and decreasing further to 0.72 (95%CI, 0.69-0.75) when data from "sensitivity-only" investigations were included. Not surprisingly, the diagnostic sensitivity fell well below the WHO sensitivity limit when the analysis included asymptomatic cohorts (i.e., 0.57; 95%CI, 0.48-0.65), becoming the lowest when these tests are used for purposes of large population screening (i.e., 0.45; 95%CI; 0.36-0.54) [11]. Many reasons have been highlighted for justifying the lower diagnostic performance of SARS-CoV-2 RDT-Ag compared to NAATs and even to laboratory-based immunoassay, including the fact that the visual reading of test results, often performed by the patients themselves, may lead to inaccurate interpretation [41], an issue which could be theoretically overcome using analyzer-read SARS-CoV-2 RDT-Ag [42].

The results of our meta-analysis of studies which explored the performance of Quidel Sofia SARS Ag FIA for diagnosing acute SARS-CoV-2 infections reveal that the overall performance of this instrument-read test satisfactory met the WHO threshold of ≥0.80 and ≥0.97 diagnostic sensitivity and specificity in symptomatic individuals (i.e., being 0.81 and 1.00), thus achieving satisfactory accuracy for being used for the WHO and IFCC intended purposes, irrespectively of the type of sample being tested (i.e., nasal or nasopharyngeal swab; table 2). Notably,

the diagnostic performance was also found to be nearly optimal in the mixed cohorts of patients (i.e., 0.77 sensitivity and 0.99 specificity), whilst the diagnostic sensitivity remained definitively low in cohorts of asymptomatic subjects (i.e., 0.55, decreasing to 0.45 when using nasal swabs). Similar results were reported in the four studies whose results could not be pooled in our analysis, with values of diagnostic sensitivity in samples taken from symptomatic individuals comprised between 0.64-0.77 and specificity always ≥0.97. Expectedly, even in these investigations the diagnostic sensitivity of Quidel Sofia SARS Ag FIA was found to be remarkably decreased in samples taken from asymptomatic or mixed cohorts of subjects (i.e., between 0.26-0.60). These results are hence aligned to those earlier published by the Cochrane COVID-19 Diagnostic Test Accuracy Group, which pooled the results of only 4 studies (with 1064 samples) and calculated an overall diagnostic sensitivity of 0.80 (95%CI, 0.72-0.86) and an overall diagnostic specificity of 0.99 (95%CI, 0.99-1.00) for Quidel Sofia SARS Ag FIA. Importantly, the article by Ford et al. provided additional information on the use of such test, showing that the diagnostic sensitivity parallels the likelihood of obtaining a positive viral culture, thus enabling a very accurate identification of contagious subjects [22]. Two additional studies, excluded from our pooled analysis because they lacked clinical performance data deserve to be briefly mentioned. Deil et al. carried out a preliminary analysis by constructing a mathematical model for estimating the economical burden of sample-and-stay strategy in German healthcare workers based on the use of Quidel Sofia SARS Ag FIA, and concluded that sequential testing was effective to significantly lower the cumulative hospital expenditure due to shortage of quarantined hospital staff [43]. In a subsequent investigation, the same authors explored the economic impact of using the Quidel Sofia SARS Ag FIA compared to that based on clinical judgement and NAAT for diagnosing COVID-19 in a cohort of German adult patients presenting to the emergency department, concluding that the RDT-AG test enabled to substantially reduce hospital costs by over 200 € for each patient tested [44].

In conclusion, the results of this critical literature review and meta-analysis suggest that the modest but significant improvement shown by the instrument-read Quidel Sofia SARS Ag FIA over more traditional "optically only"-read RDT-Ag would straightforwardly align its diagnostic accuracy to that exhibited cumulatively by laboratory-based SARS-CoV-2 immunoassays, (i.e., 0.76 vs. 0.73 sensitivity and 1.00 vs. 0.98 specificity) [10]. This test may hence combine satisfactory diagnostic performance with the advantages of being potentially used as a POC. On the other hand, the still insufficient diagnostic sensitivity emerged from our analysis in samples taken from asymptomatic patients would suggest to discourage its usage – as with most other SARS-CoV-2 Ag immunoassays - for diagnosis of acute SARS-CoV-2 infection in low-probability subjects. However, in such settings, it could be theoretically used to identify those with higher viral load, who may be responsible for a substantially higher burden of transmission.

❖❖❖❖❖

Conflicts of interest

The authors declare no conflict of interest.

❖❖❖❖❖

REFERENCES

1. McVernon J, Liberman J. WHO keeps covid-19 a public health emergency of international concern. BMJ 2023; 380:p504.

2. Mattiuzzi C, Henry BM, Lippi G. Regional Association between Mean Air Temperature and Case Numbers of Multiple SARS-CoV-2 Lineages throughout the Pandemic. Viruses 2022;14:1913.

3. Lippi G, Sanchis-Gomar F, Henry BM. Coronavirus disease 2019 (COVID-19): the portrait of a perfect storm. Ann Transl Med 2020;8:497.

4. World Health Organization. Public health surveillance for COVID-19: interim guidance - Interim guidance. Available at: https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.2. Last accessed, March 23, 2023.

5. Lippi G, Favresse J, Gromiha MM, SoRelle JA, Plebani M, Henry BM. Ad interim recommendations for diagnosing SARS-CoV-2 infection by the IFCC SARS-CoV-2 variants working group. Clin Chem Lab Med 2022;60:975-81.

6. Pighi L, Henry BM, Mattiuzzi C, De Nitto S, Salvagno GL, Lippi G. Cost-effectiveness analysis of different COVID-19 screening strategies based on rapid or laboratory-based SARS-CoV-2 antigen testing. Clin Chem Lab Med. 2023 Feb 28. doi: 10.1515/cclm-2023-0164. Epub ahead of print.

7. Bernasconi L, Neyer P, Oberle M, Schmid B, Martin E, Fankhauser H, et al. Efficiency evaluation of a SARS-CoV-2 diagnostic strategy combining high throughput quantitative antigen immunoassay and real time PCR. Clin Chem Lab Med. 2023 Mar 22. doi: 10.1515/cclm-2023-0087. Epub ahead of print.

8. World Health Organization. Antigen-detection in the diagnosis of SARS-CoV-2 infection. Available at: https://www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays. Last accessed: March 23, 2023.

9. Favresse J, Douxfils J, Henry B, Lippi G, Plebani M. Clinical Chemistry and Laboratory Medicine celebrates 60 years - narrative review devoted to the contribution of the journal to the diagnosis of SARS-CoV-2. Clin Chem Lab Med. 2022 Nov 24. doi: 10.1515/cclm-2022-1166. Epub ahead of print.

10. Lippi G, Henry BM, Plebani M. Pooled analysis of laboratory-based SARS-CoV-2 antigen immunoassays. Clin Chem Lab Med. 2023 Jan 12. doi: 10.1515/cclm-2022-1321. Epub ahead of print.

11. Dinnes J, Sharma P, Berhane S, van Wyk SS, Nyaaba N, Domen J, et al. Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev 2022;7:CD013705.

12. Cubas-Atienzar AI, Kontogianni K, Edwards T, Wooding D, Buist K, Thompson CR, et al. Limit of detection in different matrices of 19 commercially available rapid antigen tests for the detection of SARS-CoV-2. Sci Rep 2021;11:18313.

13. Lippi G, Henry BM, Plebani M. An overview of the most important preanalytical factors influencing the clinical performance of SARS-CoV-2 antigen rapid diagnostic tests (Ag-RDTs). Clin Chem Lab Med 2022;61:196-204.

14. Ang GY, Chan KG, Yean CY, Yu CY. Lateral Flow Immunoassays for SARS-CoV-2. Diagnostics (Basel) 2022;12:2854.

15. Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.

16. Alonaizan F, AlHumaid J, AlJindan R, Bedi S, Dardas H, Abdulfattah D, et al. Sensitivity and Specificity of Rapid SARS-CoV-2 Antigen Detection Using Different Sampling Methods: A Clinical Unicentral Study. Int J Environ Res Public Health 2022;19:6836.

17. Bachman CM, Grant BD, Anderson CE, Alonzo LF, Garing S, Byrnes SA, et al. Clinical validation of an open-access SARS-COV-2 antigen detection lateral flow assay, compared to commercially available assays. PLoS One 2021;16:e0256352.

18. Beck ET, Paar W, Fojut L, Serwe J, Jahnke RR. Comparison of the Quidel Sofia SARS FIA Test to the Hologic Aptima SARS-CoV-2 TMA Test for Diagnosis of COVID-19 in Symptomatic Outpatients. J Clin Microbiol 2021;59: e02727-20.

19. Bornemann L, Dähne T, Fomenko A, Kaup O, Kleideiter J, Kulis-Horn R, et al. Performance of the Sofia SARS-CoV-2 rapid antigen test as frontline test in a university hospital, Germany. Diagn Microbiol Infect Dis 2022;103: 115663.

20. Černila M, Logar M, Možina H, Osredkar J. Comparison between the Sofia SARS Antigen FIA Test and the PCR Test in Detection of SARS-CoV-2 Infection. Lab Med 2023;54:e44-8.

21. Epling J, Lowery M, Mahoney A, Baffoe-Bonnie A, Locklear T, Tenzer M, et al. Diagnostic accuracy of a new COVID-19 antigen test obtained by mid-turbinate swab. Ann Fam Med. 2022 Apr 1;(20 Suppl 1). doi: 10.1370/afm.20.s1.2920.

22. Ford L, Lee C, Pray IW, Cole D, Bigouette JP, Abedi GR, et al. Epidemiologic Characteristics Associated With Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Antigen-Based Test Results, Real-Time Reverse Transcription Polymerase Chain Reaction (rRT-PCR) Cycle Threshold Values, Subgenomic RNA, and Viral Culture Results From University Testing. Clin Infect Dis 2021;73:e1348-55.

23. Freeman MC, Freeman TJ, Iagnemma J, Rasmussen J, Heidenreich K, Wells A, et al. Performance of the Sofia SARS-CoV-2 Rapid Antigen Test in Symptomatic and

Asymptomatic Pediatric Patients. J Pediatric Infect Dis Soc 2022;11:417-21.

24. Hahn M, Olsen A, Stokes K, Fowler RC, Gu R, Semple-Lytch S, et al. Use, Safety Assessment, and Implementation of Two Point-of-Care Tests for COVID-19 Testing. Am J Clin Pathol 2021;156:370-80.

25. Harmon K, de St Maurice AM, Brady AC, Swaminathan S, Aukerman DF, Rueda MA, et al. Surveillance testing for SARS-COV-2 infection in an asymptomatic athlete population: a prospective cohort study with 123 362 tests and 23 463 paired RT-PCR/antigen samples. BMJ Open Sport Exerc Med 2021;7:e001137.

26. Harris DT, Badowski M, Jernigan B, Sprissler R, Edwards T, Cohen R, et al. SARS-CoV-2 Rapid Antigen Testing of Symptomatic and Asymptomatic Individuals on the University of Arizona Campus. Biomedicines 2021;9: 539.

27. Jääskeläinen AE, Ahava MJ, Jokela P, Szirovicza L, Pohjala S, Vapalahti O, et al. Evaluation of three rapid lateral flow antigen detection tests for the diagnosis of SARS-CoV-2 infection. J Clin Virol 2021;137:104785.

28. Mack CD, Osterholm M, Wasserman EB, Petruski-Ivleva N, Anderson DJ, Myers E, et al. Optimizing SARS-CoV-2 Surveillance in the United States: Insights From the National Football League Occupational Health Program. Ann Intern Med 2021;174:1081-9.

29. Mitchell SL, Orris S, Freeman T, Freeman MC, Adam M, Axe M, et al. Performance of SARS-CoV-2 antigen testing in symptomatic and asymptomatic adults: a single-center evaluation. BMC Infect Dis 2021;21:1071.

30. Porte L, Legarraga P, Iruretagoyena M, Vollrath V, Pizarro G, Munita J, et al. Evaluation of two fluorescence immunoassays for the rapid detection of SARS-CoV-2 antigen-new tool to detect infective COVID-19 patients. PeerJ 2021;9:e10801.

31. Pray IW, Ford L, Cole D, Lee C, Bigouette JP, Abedi GR, et al. Performance of an Antigen-Based Test for Asymptomatic and Symptomatic SARS-CoV-2 Testing at Two University Campuses - Wisconsin, September-October 2020. MMWR Morb Mortal Wkly Rep 2021;69:1642-7.

32. Smith RL, Gibson LL, Martinez PP, Ke R, Mirza A, Conte M, et al. Longitudinal Assessment of Diagnostic Test Performance Over the Course of Acute SARS-CoV-2 Infection. J Infect Dis;224:976-82.

33. Young S, Taylor SN, Cammarata CL, Varnado KG, Roger-Dalbert C, Montano A, et al. Clinical Evaluation of BD Veritor SARS-CoV-2 Point-of-Care Test Performance Compared to PCR-Based Testing and versus the Sofia 2 SARS Antigen Point-of-Care Test. J Clin Microbiol 2020; 59:e02338-20.

34. Agard A, Elsheikh O, Bell D, Relich RF, Schmitt BH, Sadowski J, et al. Clinical comparison and agreement of PCR, antigen, and viral culture for the diagnosis of COVID-19: Clinical Agreement Between Diagnostics for COVID19. J Clin Virol Plus 2022;2:100099.

35. Al-Alawi S, Al-Hinai H, Al-Kindi N, Al-Rashidi M, Al-Kindi H, Al-Shukri I, et al. Evaluation of Four Rapid Antigen Tests for Detection of SARS-CoV-2 virus. Oman Med J 2021;36:e297.

36. Brihn A, Chang J, OYong K, Balter S, Terashita D, Rubin Z, et al. Diagnostic Performance of an Antigen Test with RT-PCR for the Detection of SARS-CoV-2 in a Hospital Setting - Los Angeles County, California, June-August 2020. MMWR Morb Mortal Wkly Rep 2021;70:702-6.

37. Schroeder LF, Bachman MA, Idoni A, Gegenheimer-Holmes J, Kronick SL, Valdez R, et al. Predicting Direct-Specimen SARS-CoV-2 Assay Performance Using Residual Patient Samples. J Appl Lab Med 2022;7:661-73.

38. Dehesh P, Baradaran HR, Eshrati B, Motevalian SA, Salehi M, Donyavi T. The Relationship Between Population-Level SARS-CoV-2 Cycle Threshold Values and Trend of COVID-19 Infection: Longitudinal Study. JMIR Public Health Surveill 2022;8:e36424.

39. COVID-19 APHP-Universities-INRIA-INSERM Group. Early indicators of intensive care unit bed requirement during the COVID-19 epidemic: A retrospective study in Ile-de-France region, France. PLoS One 2020;15:e0241406.

40. Han AX, Toporowski A, Sacks JA, Perkins MD, Briand S, van Kerkhove M, et al. SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs. Nat Genet 2023;55:26-33.

41. Karlafti E, Tsavdaris D, Kotzakioulafi E, Kaiafa G, Savopoulos C, Netta S, et al. The Diagnostic Accuracy of SARS-CoV-2 Nasal Rapid Antigen Self-Test: A Systematic Review and Meta-Analysis. Life (Basel) 2023;13:281.

42. Keskin AU, Ciragil P, Topkaya AE. Clinical Accuracy of Instrument-Read SARS-CoV-2 Antigen Rapid Diagnostic Tests (Ag-IRRDTs). Int J Microbiol 2022;2022: 9489067.

43. Diel R, Hittel N, Nienhaus A. Point-of-Care COVID-19 Antigen Testing in Exposed German Healthcare Workers-A Cost Model. Int J Environ Res Public Health 2021; 18:10767.

44. Diel R, Nienhaus A. Point-of-care COVID-19 antigen testing in German emergency rooms - a cost-benefit analysis. Pulmonology 2022;28:164-72.

Supplementary File 1	Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 3-4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 4
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 5-6
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 5-6
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 5-6
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 5-6

Section and Topic	Item #	Checklist item	Location where item is reported
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 5-6
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 5-6
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 5-6
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	N/A
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 6
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Page 5-6
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 5-6
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 6

Section and Topic	Item #	Checklist item	Location where item is reported
Synthesis methods	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page 6
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Page 6
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 7
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Page 7
Study characteristics	17	Cite each included study and present its characteristics.	Page 7 – Tables 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	N/A
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/ credible interval), ideally using structured tables or plots.	Page 7,8 - Tables 1 & 2

Section and Topic	Item #	Checklist item	Location where item is reported
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	N/A
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Page 7,8 – Tables 1 & 2 – Figures 1-4
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Page 7,8
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Page 7,8
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 9,10
	23b	Discuss any limitations of the evidence included in the review.	Page 10,11
	23c	Discuss any limitations of the review processes used.	Page 9-11
	23d	Discuss implications of the results for practice, policy, and future research.	Page 10-11
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
Registration and protocol	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	N/A
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 12
Competing interests	26	Declare any competing interests of review authors.	Page 12
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Upon request to corr. author

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71..
For more information, visit: http://www.prisma-statement.org/.

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Development of control material for exhaled breath-alcohol testing and its application

Krittin Chumsawat[1,2,3], Somsak Fongsupa[2,3], Sudawadee Kongkhum[2,3], Pramote Sriwanitchrak[2,3], Narisa K. Bordeerat[1,2,3]

[1] *Graduate Program in Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathumthani, Thailand*

[2] *Department of Medical Technology, Faculty of Allied Health Sciences, Thammasat University, Pathumthani, Thailand*

[3] *Research Unit in Developing of Quality Control Materials for Medical Laboratory, Thammasat University, Pathumthani, Thailand*

ARTICLE INFO

Corresponding author:
Narisa K. Bordeerat
Department of Medical Technology
Faculty of Allied Health Sciences
Thammasat University
Pathumthani, 12120
Thailand
Phone: (668)1-912-2694
E-mail: narisa.k@allied.tu.ac.th

Key words:
breath alcohol testing,
control material, uncertainty

ABSTRACT

Background

Breath analyser tests are used worldwide to obtain proof of alcohol intoxication and often used in the conviction of traffic violators. These tests are conducted to quickly and painlessly determine the existing concentration of alcohol in arterial blood by measuring the amount of ethanol in exhaled breath, which can be identified with an electrochemical sensor.

At present, the calibration and maintenance of analysers used for these tests are typically performed regularly but lack quality control. Consequently, test results may not be accurate because of calibration deterioration.

The aim of this study was to develop and evaluate the uncertainty of control materials used in breath-alcohol testing at the Bangkok Metropolitan Police Station.

Material and methods

Ethyl alcohol (99.99%; Certified Reference Material grade) diluted at three different concentrations was kept under design conditions. The concentrations were 28, 67, and 134 mg/dL, determined by performing headspace gas chromatography, and the uncertainty was set as ±1.3925, ±2.8736, and ±1.8231 mg/dL (±4.97%, ±4.29%, and ±2.72% for the concentrations, respectively), as per ISO Guide 35:2017.

Results

The total error percentages of the developed control materials were 4.97%, 4.29%, and 2.72% for concentrations of 28, 67, and 134 mg/dL, respectively. Each concentration of the materials was tested by using measurements from 70 breath-alcohol analysers belonging to the Bangkok Metropolitan Police Station.

Conclusion

These control materials are applicable to quality assurance and standards tests and may help to ensure the accuracy of breath-alcohol testing in the future.

❖❖❖❖❖

1. INTRODUCTION

According to the World Health Organization's 'Global Status Report on Road Safety', road accidents cause approximately 1.35 million fatalities per year, making it the 8th leading cause of mortality for all ages [1]. Driving under the influence of alcohol (DUI) is an important contributing factor in these accidents [2]. The amount of alcohol consumed is directly proportional to the risk and severity of accidents. In Thailand, despite legislation against DUI, the Ministry of Transport reported that out of 316 DUI accidents per year, 61 were fatal and 293 required medical attention [3],[4].

Additionally, Thai Law stipulates that the blood-alcohol level of drivers above the age of 20 years must not exceed 50 mg/dL. Drivers who are below this age, drivers with a temporary licence, licenced drivers who have received any other type of licence, and drivers whose licence has been revoked or whose application is on hold must have no more than 20 mg/dL alcohol concentration in their blood [5], as measured by police officers using an breath-alcohol analyser at the scene of the accident or when the driver is suspected of DUI.

A breath-alcohol analyser is a device that measures the alcohol in exhaled breath using colorimetric [6], semiconductor [7], or infrared absorption [8]. Detection methods requiring electrochemical-sensor-based devices are the most popular because of their portability, short analytical time, accuracy, good sensitivity, and adequate specificity [9],[10],[11]. However, these methods involve the risk of errors common to all medical laboratory equipment, such as insufficient biological sampling and traceability issues. These errors can be detected through quality control materials [12],[13], [14].

Because electrochemical-sensor-based breath analysers are used routinely by traffic police officers, these instruments are calibrated every six months by an external organisation. However, internal quality control is not regularly performed for this kind of routine task, creating doubt about the reliability of the results in the event of a lawsuit. Furthermore, calibration is expensive. The aim of this study was to develop control materials for breath-alcohol analysers. We evaluated the measurement uncertainty of our control materials according to ISO Guide 35:2017 [15] and compared it to the allowable total error (TEa) specified by

CLIA2019 [16]. Finally, we applied the materials to the breath-alcohol analysers used in the Bangkok Metropolitan Police Station. These materials have the potential to make quality control more accessible to all police stations, thereby improving standardised and reliable results.

2. MATERIALS AND METHODS

2.1 Development of control material

Three concentrations of the control material—low, medium, and high—were prepared using 3.267, 8.171, and 16.340 mL of 99.99% ethyl alcohol (HPLC grade, DAEJUNG, Republic of Korea) with 10 L of distilled water each. The samples were mixed by applying inversion and divided into 20 plastic bottles containing 500 mL each. The bottles were sealed with parafilm, an aluminium sheet, and finally a plastic screw cap. All materials were stored at a temperature of 25 ±2 ℃ and a humidity of 50% ±5% for three months.

2.2 Uncertainty of measurement (MU)

The MU of our control materials was determined as per ISO Guide 35:2017 [15],[17] by using headspace gas chromatography (HSGC; SHIMAZU GC-2010, Japan). These experiments were performed in the toxicology laboratory of the Institute of Forensic Medicine, Police General Hospital, Bangkok, Thailand. The allowable total error; TEa outlined in the CLIA2019 criteria (20%) was used to determine performance characteristics and uncertainty values. Furthermore, the HSGC method was using as the appropriate choice for the measurement procedure.

2.2.1 Homogeneity studies

Standard uncertainties were assessed as bottle-to-bottle heterogeneity (SU_{bb}) on day zero for each control substance concentration. The minimum number of units was then calculated. The bottles were sampled using a simple randomized strategy. Outlier and trend analyses were also conducted. The uncertainties between units were analysed using one-way ANOVA software for Excel.

2.2.2 Characterization study

The standard uncertainty owing to the characterisation study (SU_{char}) was assessed for each control material concentration. The average result was used as the assigned value for each concentration.

2.2.3 Stability study

Standard uncertainty resulting from long-term instability (SU_{lts}) was assessed over a 3-month period with storage conditions of 25 ± 2℃ and 50% ± 5% humidity, and no transportation conditions.

A classic stability study was also conducted. Two bottles of the control material were sampled at six time points: 0, 7, and 14 d, and 1, 2, and 3 months for each concentration. The resulting trends were analysed, and the SU_{lts} at each concentration was evaluated using a t-test.

2.2.4 Expanded uncertainty

The expanded uncertainty (U_x) was calculated from SU_{bb}, SU_{char}, and SU_{lts} with a 95% confidence interval (coverage factor k = 2). The equation is:

$$U_x = k\sqrt{SU_{bb}^2 + SU_{char}^2 + SU_{lts}^2}.$$

2.3 Application in breath-alcohol analysers from Bangkok Metropolitan Police Station

Consent and questionnaire surveys were sent to Bangkok Metropolitan Police Stations.

Our control materials were tested with 70 electrochemical breath-alcohol analysers (SD-400 Touch, Lion, UK) by using a wet-bath simulator. Quality was evaluated by using |%BIAS| from

the HSGC-assigned value to assess accuracy and %CV for precision, and the total error (TE) was calculated.

3. RESULTS

3.1 MU of developed control materials

The HSGC procedure was evaluated by comparing the repeatability standard deviation (S_r), the number of observations of each of the 10 aliquots (n_{al}), and the target uncertainty (u_{trg}), calculated using 20% TEa for each concentration. The results showed that the HSGC procedure produced good precision for all concentrations of the control material (Table 1).

3.1.1 Homogeneity study

The minimum number of control materials was 3, or 10% of the batch. This study chose 10 bottles: sample numbers 2, 3, 6, 7, 9, 11, 13, 16, 18, 20. The results showed alcohol concentrations of 28.00% (SD = 0.87%), 66.74% (SD = 0.61%), and 134.17% (SD = 1.31%), and are summarized in Table 2. None of the data showed trends or outliers (P > 0.05). The developed control materials were homogeneous—p = 0.3736, 0.9013, and 0.0672 for 28, 67, and 134 mg%, respectively—and SU_{bb} reported 0.2810, 0.3402, and 0.9042 mg%, respectively. These data are shown in Table 3.

3.1.2 Characterization study

The SU_{char} of the control materials was evaluated without the unweighted mean or laboratory uncertainties, by referring to the SDM results with assigned values of 28, 67, and 134 mg% for low, medium, and high concentrations, respectively (Table 2). The SU_{char} showed a minimum of 67 mg% (SU_{char} = ±0.1032 mg%) and a maximum of 134 mg% (SU_{char} = ±0.3595 mg%), as outlined in Table 4.

3.1.3 Stability study

Storage-controlled materials remained acceptable with no significant change for any concentration over 3 months (p > 0.05). The minimum and maximum SU_{lts} values were 28 mg% (SU_{lts} = 0.6030 mg%) and 134 mg% (SU_{lts} = 1.5417 mg%) (Table 5), respectively.

3.1.4 Expanded uncertainty

The U_x of developed control materials were calculated with a 95% CI (coverage factor = 2). Results showed that U_x = 1.3925, 2.8736, and 1.8231 mg% for concentrations of 28, 67, and 134 mg%, respectively (Table 6).

3.2 Application in breath-alcohol analysers from Bangkok Metropolitan Police Station

The Bangkok Metropolitan Police Station routinely uses breath-alcohol analysers. Our control materials were tested on 70 instruments which were grouped according to the time after the latest calibration: < 2 months (1), 2–4 months (2), and > 4 months (3) (Table 7). The results from applying the developed control materials showed precision and %CV which were minimum for group (1)—67 mg% (CV = 2.90%)—and maximum for group (3)—28 mg% (CV = 14.24%)—illustrated in Figure 1(a). The accuracy is shown as |%BIAS| which was also at its minimum in group (1) at 28 mg% (|BIAS|= 4.23%) and at its maximum in group (3), 28 mg% (|BIAS|=12.70%). This is summarized in Figure 1(b). The TE was also calculated for each analyser; minimum TE was found in group (1) at 134 mg% (TE=8.60%) and the maximum was found in group (3) at28 mg% (TE=26.94%), outlined in Figure 1(c). Notably, the calculated TE showed that only the SD-400Touch instruments in groups (1) and (2) met the 20% CLIA2019 TEa standard.

Table 1 Evaluation of HSGC procedure with 20% TEa (CLIA2019)

Concentration	u_{trg}	s_r	$s_r / \sqrt{(n^{al})}$ (A)	$u_{trg} / 3$ (B)	Conclusion
Low	5.60	0.87	0.02	1.87	Good precision[1]
Medium	13.4	0.61	0.02	4.47	
High	26.8	1.31	0.03	8.93	

[1] *criterion (A) < (B) conclusion indicates 'good precision'*

Abbreviations: $\mathbf{u}_{trg}$ *represents the target uncertainty;* $\mathbf{s}_r$ *is the standard deviation; and* $\mathbf{n}_{al}$ *denotes the unit for measurement.*

Table 2 Alcohol concentration in control materials analysed by HSGC at day 1

	Low concentration (mg%)			Medium concentration (mg%)			High concentration (mg%)		
Bottle no.	1st	2nd	average	1st	2nd	average	1st	2nd	average
2	28.12	27.88	28.00	66.06	67.79	66.93	135.49	134.43	134.96
3	27.47	28.05	27.76	67.05	67.42	67.24	133.59	134.49	134.04
6	27.81	28.07	27.94	65.75	66.29	66.02	133.03	133.92	133.48
7	27.91	28.22	28.07	66.25	66.70	66.48	133.26	134.85	134.06
9	27.74	28.74	28.24	66.47	66.80	66.64	133.37	135.23	134.30
11	28.14	30.53	29.34	67.71	66.19	66.95	132.37	135.06	133.72
13	27.97	28.98	28.48	66.72	66.91	66.82	134.02	135.05	134.54
16	27.20	27.23	27.22	65.72	67.57	66.65	133.89	134.82	134.36
18	26.87	29.20	28.04	67.29	66.40	66.85	134.20	133.16	133.68
20	27.01	26.94	26.98	66.72	66.96	66.84	134.64	134.50	134.57
Average			28.00			66.74			134.17
SD			0.87			0.61			1.31
Assign Value			28			67			134

Table 3 Standard uncertainty from inhomogeneity (SU_{bb})

Concentration	Source of variation	SS	df	MS	F	p-value	SU_{bb} (mg%)	U_{bb} of conc. (%)
Low	Between groups Within groups Total	7.5976 6.8623 14.4599	9 10 19	0.8442 0.6862 (u_{bb})	1.2301	0.3736	***±0.2810***	***1.00***
Medium	Between groups Within groups Total	1.9173 5.1758 7.0931	9 10 19	0.2130 0.5176 (u'_{bb})	0.4416	0.9013	***±0.3402***	***0.51***
High	Between groups Within groups Total	23.2574 9.4883 32.7457	9 10 19	2.5842 0.9488 (u_{bb})	2.7235	0.0672	***±0.9042***	***0.67***

Abbreviations: SS represents the sum of squares; df denotes the degrees of freedom; MS indicates the mean squares; F is the F ratio; ***SU_{bb}*** *is the standard uncertainty due to inhomogeneity;* ***U_{bb}*** *is the uncertainty due to inhomogeneity.*

Table 4 Standard uncertainty due to characterization (SU_{char})

Concentration	Average (mg%)	SD (mg%)	SDM (mg%)	SU_{char} (mg%)	U_{char} of conc. (%)
Low	28.00	0.65	0.2054	0.2054	0.73
Medium	66.74	0.33	0.1032	0.1032	0.15
High	134.17	1.14	0.3595	0.3595	0.27

Abbreviations: SD is standard deviation; SDM is the standard deviation mean; and SU_{char} is the standard uncertainty due to characterisation; U_{char}, is the uncertainty due to characterisation.

Table 5 Standard uncertainty due to long-term instability (SU_{lts})

Duration	Concentration (mg%)		
	Low	Medium	High
0 day	28.00	66.74	134.17
7 days	28.14	68.42	135.40
14 days	28.15	66.58	133.07
1 month	27.39	64.95	131.87
2 months	27.47	65.73	132.94
3 months	28.68	66.47	134.00
SU_{lts} (mg%)	0.6030	1.3921	1.5417
U_{lts} of conc. (%)	2.15	2.08	1.15

Abbreviations: SU_{lts} is the standard uncertainty due to long-term instability; U_{lts} is the uncertainty due to long-term instability.

Table 6 Expended uncertainty of developed control materials (U_x).

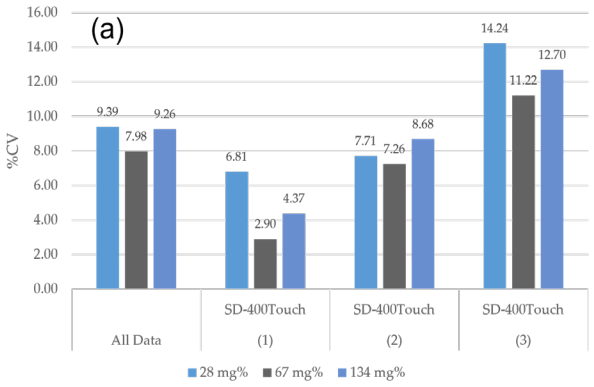
Source of *SU*	Concentration (mg%)		
	Low	Medium	High
SU_{bb}	0.2810	0.3402	0.9042
SU_{char}	0.2054	0.1032	0.3595
SU_{lts}	0.6030	1.3921	1.5417
Combined uncertainty (U_c)	0.6962	1.4368	1.8231
U_x (mg%)	1.3925	2.8736	1.8231
U_x of conc. (%)	4.97	4.29	2.72

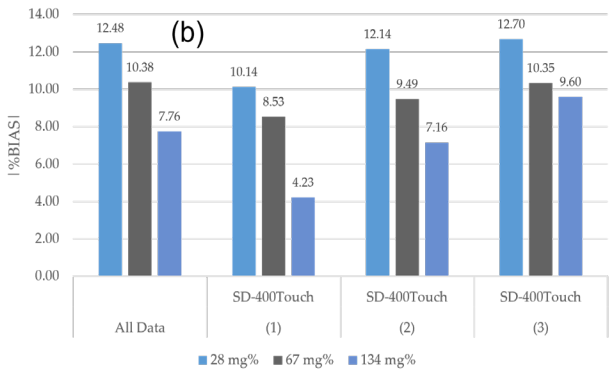
Abbreviations: SU_{bb} is the standard uncertainty due to inhomogeneity; SU_{char} is the standard uncertainty due to characterisation; SU_{lts} is the standard uncertainty due to long-term instability; U_c is the combined uncertainty; U_x is the expended uncertainty.

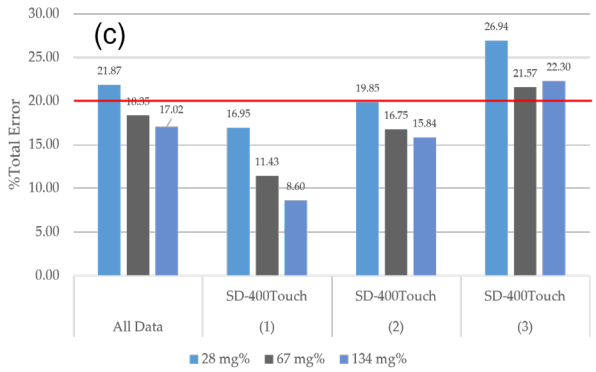
Table 7 Breath-alcohol analysers grouped according to time after latest calibration. This structuring is applied to the data in Figure 1

Model	Time after latest calibration/analyser (n)			sum
	< 2 months (1)	2–4 months (2)	> 4 months (3)	
SD-400Touch	30	25	15	70

Figure 1 Results from the application of control materials to breath-alcohol analysers, grouped according to time after latest calibration: (a) %CV in each group and concentration; (b) |%BIAS| in each group and concentration; (c) %TE in each group and concentration







4. DISCUSSION

Quality control plays an important role in the reporting of reliable results from medical laboratories. For this reason, quality control is critical for preventing inaccuracy or imprecision in the results of all tests—even breath-alcohol testing performed by police officers on motorists at sideroads. Ensuring that breath-alcohol analysers report values correctly is challenging. This is because the reference materials are expensive, and calibration is scheduled in six month intervals.

According to ISO/IEC 17043:2010, the 'general requirements for proficiency testing' describe qualifications in procedures and reference materials used for determining result quality [15], [17], [19]. The MU of the reference materials was determined as per ISO Guide 35:2017 specifying three causes of uncertainty: inhomogeneity, characterisation, and long-term instability, which may be evaluated to expand the uncertainty [15],[17].

In this study, the reference material was used as a control to determine the quality status of each breath-alcohol analyser. We developed the control material by diluting ethyl alcohol 99.99% (Certified Reference Material grade) in deionised water in three concentrations and then sealing the samples with parafilm, aluminium sheet, plastic screw caps and placing them in storage at 25±2 ℃ and 50 ± 5% humidity for 3 months. This accessible procedure could substitute for the more expensive reference materials currently in use. The %TE for each concentration was found to be 4.97%, 4.29% and 2.72% in control material concentrations of 28, 67, and 134 mg% respectively. Our study did not differ from other studies in which TE = 4.72%, 4.72%, and 4.27% was found in alcohol reference material concentrations of 46.6, 50.8, and 56.3 mg% [20].

The results of applying the developed control materials in 70 police-issued breath-alcohol analysers revealed the TE to be acceptable only when the most recent calibration was performed less than four months ago, according to the 20% TEa standard outlined in CLIA2019. The further the instruments are removed from their latest calibration date, past the four-month mark, increased the imprecision of the analysers. Additionally, multiple confounding factors were found, including the service life of each instrument and the experience of the users.

A limitation of this study is that our control materials were applied only in the Bangkok area and must still be tested for commutability. The confounding factors (service life, user competency) also remain unexplored.

5. CONCLUSIONS

Our control materials were developed by employing a designed protocol and storage conditions that met the criteria of CLIA2019. We determined the TEa limit as per the specifications of ISO Guide 35:2017.

These materials could be used to routinely evaluate the quality of breath-alcohol analysers for more reliable results.

❖❖❖❖❖

Funding

This research has received financial support from the Thammasat University Research Unit in Developing of Quality Control Materials for Medical Laboratory.

Credit authorship contribution statement

Narisa Kengtrong Bordeerat: Conceptualization, Data curation, Methodology, Resources, Formal analysis, Investigation.

Krittin Chumsawat: Investigation, Visualization, Data curation, Methodology, Resources, Formal analysis Writing – original draft.

Somsak Fongsupa: Conceptualization, Formal analysis, Methodology.

Sudawadee Kongkhum: Visualization, Writing – review and editing.

Pramote Sriwanitchrak: Writing – review and editing.

Acknowledgments

We thank the Bangkok Metropolitan Police Station for allowing us to use the data and Breath-alcohol analyzer for this research, and Language advisor from Faculty of allied Health Sciences for improving the use of English in the manuscript.

Funding for open access charge: University of Thammasat University / Thailand.

Authors' disclosures declarations

No known conflicts of interest are associated with this publication.

❖❖❖❖❖

REFERENCES

1. Organization, W.H., Global status report on alcohol and health 2018. 2019: World Health Organization.

2. Fell, J.C. and R.B. Voas, The effectiveness of reducing illegal blood alcohol concentration (BAC) limits for driving: evidence for lowering the limit to. 05 BAC. Journal of safety research, 2006. 37(3): p. 233-243.

3. Accidental on road data. Office of the Permanent Secretary 2022; Available from: https://datagov.mot.go.th/dataset/roadaccident/resource/960813aa-f410-4356-bd5a-10f63c8e4ce0?inner_span=True.

4. Royal Gazette, The Land Traffic Act Update. 1979; Available from: http://www.thailawforum.com/land-traffic-act/

5. Ministerial regulations, The Land Traffic Act Update. 2017.

6. Jung, Y., et al., Smartphone-based colorimetric analysis for detection of saliva alcohol concentration. Applied optics, 2015. 54(31): p. 9183-9189.

7. Wang, X., S. Yee, and P. Carey, An integrated array of multiple thin-film metal oxide sensors for quantification of individual components in organic vapor mixtures. Sensors and Actuators B: Chemical, 1993. 13(1-3): p. 458-461.

8. Pérez-Ponce, A., S. Garrigues, and M. de La Guardia, Vapour generation–fourier transform infrared direct determination of ethanol in alcoholic beverages. Analyst, 1996. 121(7): p. 923-928.

9. Modjtahedi, A., A. Amirfazli, and S. Farhad, Low catalyst loaded ethanol gas fuel cell sensor. Sensors and Actuators B: Chemical, 2016. 234: p. 70-79.

10. Rahman, M.R., et al., The application of power-generating fuel cell electrode materials and monitoring methods to breath alcohol sensors. Sensors and Actuators B: Chemical, 2016. 228: p. 448-457.

11. Allan, J.T., H.L. Geoffrey, and E.B. Easton, The effect of the gas diffusion layer on the performance of fuel cell catalyst layers in ethanol sensors. Sensors and Actuators B: Chemical, 2018. 254: p. 120-132.

12. Gullberg, R.G., Estimating the measurement uncertainty in forensic breath-alcohol analysis. Accreditation and quality assurance, 2006. 11(11): p. 562-568.

13. Dubowski, K.M., Quality assurance in breath-alcohol analysis. Journal of analytical toxicology, 1994. 18(6): p. 306-311.

14. Gullberg, R., Methodology and quality assurance in forensic breath alcohol analysis. Forensic science review, 2000. 12(1-2): p. 49-68.

15. Guide, I., 35 (2017) Reference materials—guidance for characterization and assessment of homogeneity and stability. International Organization for Standardization (ISO), Geneva, 2017. 114.

16. Westgard, J.O. 2019 : CLIA Proposed Changes to PT Acceptable Limits. Available from: https://www.westgard.com/2019-clia-changes.htm.

17. Botha, A., et al., Outline for the revision of ISO Guide 35. Accreditation and Quality Assurance, 2013. 18(2): p. 115-118.

18. Bartels, R., et al., A Perspective on a Quality Management System for AI/ML-Based Clinical Decision Support in Hospital Care. Frontiers in Digital Health, 2022: p. 119.

19. Trapmann, S., et al., The new International Standard ISO 17034: general requirements for the competence of reference material producers. Accreditation and Quality Assurance, 2017. 22(6): p. 381-387.

20. Bunsoong, J. and T. Bunsoong, Preparation of standard alcohol for quality control testing in alcohol breath testing. BULLETIN OF THE DEPARTMENT OF MEDICAL SCIENCES, 2020. 62(2): p. 96-105.

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Age-stratified lithium therapeutic ranges for older adults with bipolar disorder – from awareness to an action plan

Angela W. S. Fung[1,2], Kenneth I. Shulman[3,4,9], Danijela Konforte[5], Hilde Vandenberghe[6], Julia Stemp[7], Victoria R. Yuan[6], Paul M. Yip[6,8,9], Lei Fu[6,8,9]

[1] *Department of Pathology and Laboratory Medicine, St. Paul's Hospital, Vancouver, BC*
[2] *Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC*
[3] *Department of Psychiatry, Sunnybrook Health Sciences Centre, Toronto, ON*
[4] *Department of Psychiatry, University of Toronto, Toronto, ON*
[5] *LifeLabs Medical Laboratories, Toronto, ON*
[6] *Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON*
[7] *Institute for Quality Management in Healthcare, Toronto, ON*
[8] *Precision Diagnostics & Therapeutics Program (Lab Medicine), Sunnybrook Health Sciences Centre, Toronto, ON*
[9] *Sunnybrook Research Institute, Toronto, ON*

ARTICLE INFO

Co-corresponding authors:
Dr. Paul M. Yip & Dr. Lei Fu
Precision Diagnostics & Therapeutics Program (Laboratory Medicine)
Sunnybrook Health Sciences Centre
2075 Bayview Ave
Toronto, Ontario, M4N 3M5
Canada
Phone (PMY): 416-480-6100 ext. 1594
Phone (LF): 416-480-6100 ext. 89571
E-mail: paul.yip@sunnybrook.ca
E-mail: Lei.Fu@sunnybrook.ca

Key words:
lithium, therapeutic drug monitoring, elderly, therapeutic ranges, bipolar disorder

ABSTRACT

Lithium is the first-line treatment for maintenance therapy in bipolar disorder. It is an effective mood stabilizer agent, and may have potential benefits in neuroprotection and reducing the risk of suicide. Toxicity has been a concern in recent decades, particularly in older adults (≥60 years). In 2019, the Older Adults Task Force within the International Society for Bipolar Disorder (ISBD) published recommendations for age-stratified lithium therapeutic ranges for therapy of Older Age Bipolar Disorder (OABD), namely 0.4 – 0.8 mmol/L for ages 60 to 79 and 0.4 – 0.7 mmol/L for ages 80 and above. Clinical laboratory practice surveys in Canada indicated that adoption and implementation of the proposed ranges has been limited to date. In this article, we describe the approach and

steps taken to evaluate and implement recommended lithium therapeutic ranges in Ontario and other provinces in Canada for laboratory quality improvement. Sources of variation in lithium reporting practices are discussed and shared here to highlight potential barriers to implementation. The overall goal of this article is to bring attention across the global laboratory community that lower lithium therapeutic target ranges in older patients are crucial for patient safety in OABD.

❖❖❖❖❖

Abbreviations

CV, coefficient of variation;

IQMH, Institute for Quality Management in Healthcare;

ISBD, International Society for Bipolar Disorders;

L-DOPA, levodopa;

OABD, old age bipolar disorder;

RCPAQAP, The Royal College of Pathologists of Australasia Quality Assurance Programs;

SD, standard deviation;

SE, standard error.

❖❖❖❖❖

INTRODUCTION

Older age bipolar disorder (OABD) is defined as bipolar disorder in individuals aged 60 and over, and it represents approximately 25% of all bipolar disorder (BD) worldwide [1]. This group includes individuals with both early and late onset BD. With the growing older population, the proportion of OABD is projected to be over 50% by 2030 [2]. Lithium carbonate remains the first-line treatment in the maintenance of OABD due to its effectiveness in both phases of the illness, including depression and mania/hypomania [3,4]. In addition to mood stabilization, it may also have additional benefits in reducing the risk of suicide [5,6] and have neuroprotective properties for the prevention of dementia [7]. Yet, lithium toxicity has been a concern in recent decades, especially in older adults where the laboratory community needs to highlight and thereby reverse the decline of prescribing practice [4,8].

In older adults, special considerations regarding the use of lithium include increased risk of toxicity associated with normal and pathological decreases in renal function, medical co-morbidities, and drug-drug interactions with commonly used medications such as diuretics, ACE inhibitors and nonsteroidal anti-inflammatories [9–11]. It has been reported that 33% of OABD patients are prescribed these common medications, which may increase the serum lithium level by up to 50% [8]. Moreover, lithium toxicity is often misdiagnosed and attributed to other common conditions in older adults, including gastrointestinal symptoms (diarrhea), urological disorder (polyuria), impaired cognition (dementia) and neurologic symptoms similar to parkinsonism (tremor and rigidity) [12]. If not recognized as toxicity secondary to lithium use, this can result in a “prescribing cascade” whereby inappropriate and unnecessary drugs are additionally prescribed for perceived new disorders [13]. For example, parkinsonism secondary to lithium therapy can result in unnecessary treatment with L-DOPA, while impaired cognition may be interpreted as dementia and managed inappropriately with cognitive enhancers [12].

To date, there is only one randomized controlled trial that specifically addressed pharmacological treatment using lithium carbonate in older adults with bipolar disorders - the GERI BD study (Acute Pharmacotherapy in Late-Life Mania) [14]. Recent clinical practice guidelines generally recommend a lithium target maintenance therapeutic range of 0.6 to 0.8 mmol/L, without considering the age of the patient, the

phase of their illness, or medical comorbidities [3]. There is also a lack of specific recommendations for OABD in international clinical practice guidelines [3,15].

Considering the lack of systematic evidence and direction from clinical practice guidelines for use of lithium in older adults, the International Society of Bipolar Disorder (ISBD) established an Older Age Task Force comprised of international experts with real-world knowledge and experience in OABD. The group has published a report as well as a Delphi consensus survey aimed to provide specific direction for lithium and its maintenance use in OABD [1,16]. In brief, the ISBD task force on OABD recommended that lithium remains the preferred choice for maintenance treatment of OABD [16]. Second line choices include: valproate, lamotrigine, quetiapine and olanzapine. It is recommended that serum lithium levels be monitored 5 to 7 days after a dose adjustment, three to six months thereafter, as clinically necessary and if co-medications were initiated or adjusted while receiving lithium therapy [16]. Monitoring of target serum lithium levels generally relies on trough levels as the efficacy of lithium are dose-dependent and correlates well with trough levels. Trough levels are typically collected just before the next dose. In clinical practice, lithium is mostly prescribed as lithium carbonate and may be administered in divided doses, so lithium trough levels are routinely measured 12 hours following the previous dose.

The ISBD task force on OABD also provided specific recommendation on reporting separate lithium level therapeutic ranges for older adults [16]. Serum lithium target therapeutic ranges were recommended for ages 60 to 79 in the range of 0.4 to 0.8 mmol/L, and for those 80 and over in the range of 0.4 to 0.7 mmol/L [16]. The most common therapeutic range reported by laboratories was in the range of 0.6 to 1.2 mmol/L and without specific age dependent stratification [16]. The lack of age stratification may pose risks in missing lithium toxicity in older adults. Providing narrower and lower therapeutic ranges in older patients would help to increase sensitivity to adverse side effects, particularly neurotoxicity.

Given the vulnerability to toxicity and the tendency for lithium to be underutilized in this population, requests were made to the clinical laboratory community to update and provide narrower and lower therapeutic ranges for lithium in older adults [4,17]. In this article, we share our approach aiming to determine the feasibility of implementing the ISBD OABD Task Force recommended standardized therapeutic ranges in Canada for laboratory quality improvement by: a) determining the association of serum lithium concentration with age via retrospective laboratory data review, b) evaluating method agreement between common lithium methods via reviewing proficiency testing survey reports, and c) determine the current practice of clinical laboratories in the reporting of lithium levels in Canada through two voluntary surveys of clinical laboratories conducted in 2017 and 2022. Sources of variation in lithium reporting practices are discussed and shared here to highlight potential barriers to implementation.

FEASIBILITY ASSESSMENT FOR ADOPTION OF STANDARDIZED THERAPEUTIC RANGES FOR SERUM LITHIUM

To evaluate the feasibility of adopting and implementing the ISBD OABD Task Force recommendation for standardized age-stratified therapeutic ranges for lithium, we retrospectively reviewed a Toronto hospital serum lithium data for association with age, and we reviewed results from two external quality assurance providers for lithium method performance agreement.

Serum lithium laboratory results from April 1, 2020 to March 31, 2022 (n = 504) were extracted from theSunnybrook Health Sciences Centre Laboratory Information System (Toronto, Ontario, Canada). Patients from emergency department, critical care and maternal care units were excluded, and the final analysis included clinically stable patients undergoing treatment with lithium. This retrospective study has been registered with the Sunnybrook Research Ethics Board as a quality improvement project.

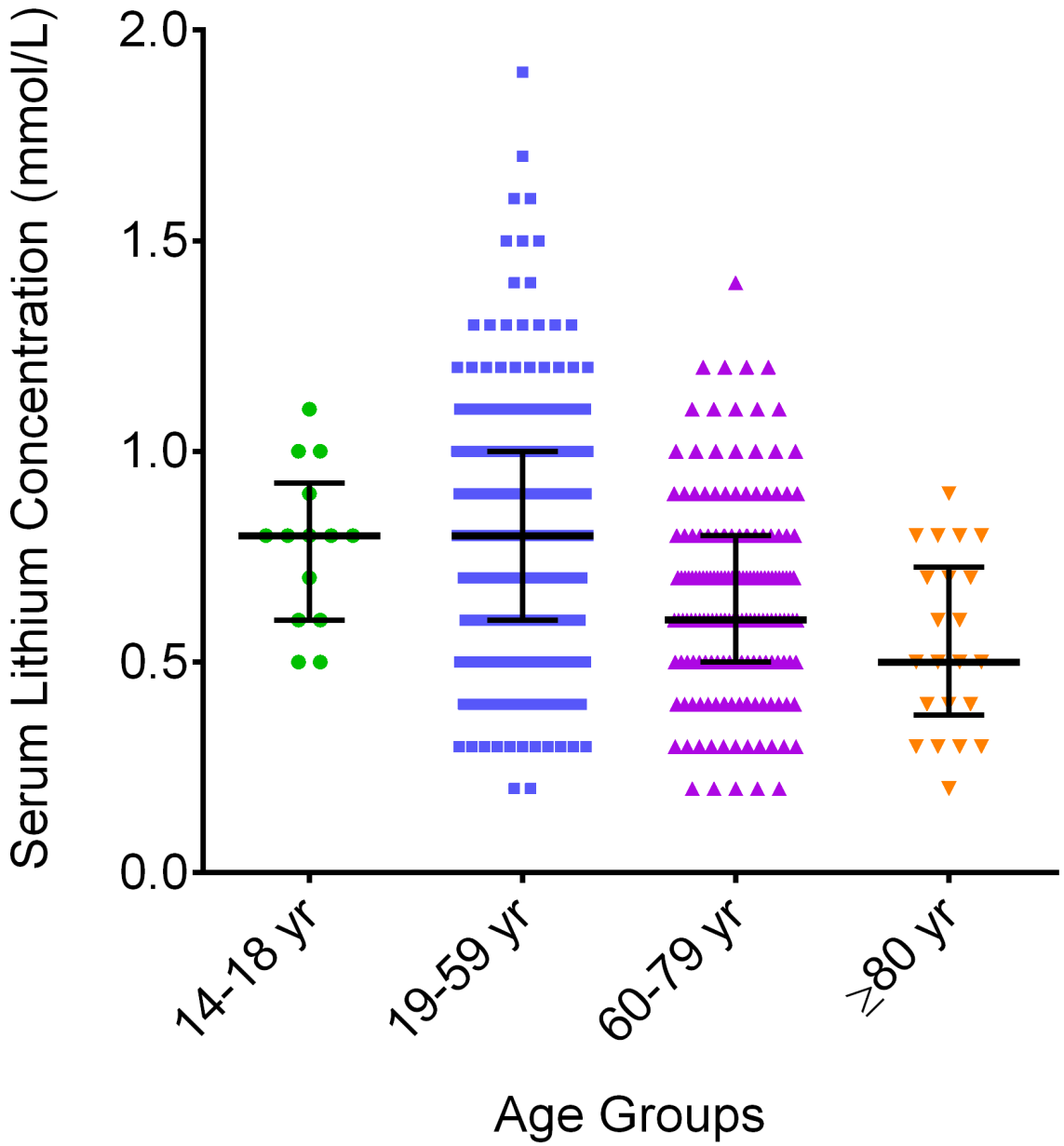
One-way ANOVA was used to compare mean serum lithium concentration between four age groups: ages 14 to 18, 19 to 59, 60 to 79, and ≥80 years old. Statistical analyses were performed by IBM SPSS Statistics V. 28.0.1.1 software. The analyses confirm that there is a significant difference in mean lithium concentration between the age groups of 19 to 59 years old and 60 to 79 years old (mean 0.80 *vs.* 0.65 mmol/L, $p < 0.001$), and between the age groups of 19 to 59 years old and ≥80 years old (mean 0.80 *vs.* 0.55 mmol/L, $p < 0.001$). Therefore, the serum lithium concentration in sera of older adults (60 to 79 years old and ≥80 years old) is significantly lower than in that of younger adults (19 to 59 years old). Figure 1 illustrates the distribution of serum lithium concentration for each age groups. This retrospective review of laboratory data provides supporting evidence that patients ≥60 years old should have a lower therapeutic target serum lithium level compared to younger adults. Prior to local implementation in other laboratory testing sites, similar results were observed from community laboratory data (i.e., Ontario and British Columbia) as well as in other provinces (data not shown).

To evaluate agreement between commonly used lithium methods for the use of standardized therapeutic target ranges, we reviewed proficiency testing survey reports between September 2020 to September 2022 from the Institute for Quality Management in Healthcare (IQMH) based in Toronto, Canada and The Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP) based in St Leonards, Australia. Both programs offer an ISO 17043:2010 accredited proficiency testing program to clinical laboratories.

The aggregate analytical performance is summarized in Table 1 and includes a total of seven surveys and 21 samples covering a range of lithium concentrations. The surveys included eight different instrument groups (Abbott Architect/Alinity c, Beckman Coulter AU, Beckman Coulter Unicel DxC, Ortho Vitros, Roche cobas c/Integra 400, Siemens Advia/Atellica, Siemens Dimension and Siemens Vista) from five major manufacturers, and reported data from 311 RCPAQAP and 86 IQMH clinical laboratory participants. All the methods are based on the colorimetric method principle. The all-methods' mean, standard deviation and coefficient of variation ranges were summarized for four categories of lithium ranges: a) <0.4 mmol/L, b) 0.4 - 0.8 mmol/L, c) 0.8 - 1.5 mmol/L, and d) >1.5 mmol/L, which represent major clinical decision limits.

The variation between lithium methods is minimal, with a range of all-methods' standard deviation of 0.04 - 0.06 mmol/L for concentrations ≤1.5 mmol/L, and <0.12 mmol/L for concentrations >1.5 mmol/L. A practice-oriented quality specification for lithium was proposed with a desirable imprecision of 5.2%, bias of 2.1% and total error allowable of 10.7% [18]. Overall, this indicates that there is acceptable and sufficient agreement between commonly used colorimetric lithium methods, and demonstrates feasibility to use narrow, age-stratified, and standardized therapeutic target ranges for serum lithium (i.e., 0.4 to 0.8 mmol/L for ages 60 to 79, and 0.4 to 0.7 mmol/L for ages ≥80).

Figure 1 **Serum lithium concentration distribution stratified by age groups**



Serum lithium concentrations (mmol/L) were plotted stratified by patient age groups (14 to 18 years, 19 to 59 years, 60 to 79 years, ≥80 years). The horizontal line and error bars represent the median and interquartile range for serum lithium concentration for each age group, respectively. This figure was generated by GraphPad Prism 5 software.

Table 1 Summary of analytical performance of colorimetric methods for serum lithium obtained from IQMH (Canada) and RCPAQAP (Australia) proficiency testing surveys between September 2020 to September 2022

Target Lithium Concentration (mmol/L)	Number of surveys	Number of survey samples included	Range of all-methods' mean# (mmol/L)	Range of all-methods' SD# (mmol/L)	Range of all-methods' CV# (%)
< 0.4	2	2	0.29 – 0.35	0.04 – 0.05	12.6 – 14.3
0.4 – 0.8	3	4	0.59 – 0.73	0.04 – 0.05	6.0 – 7.1
0.8 – 1.5	4	4	1.03 – 1.43	0.05 – 0.06	3.6 – 5.0
>1.5	7	11	1.71 – 2.70	0.05 – 0.12	3.1 – 4.5

 # All-methods mean, standard deviation (SD) and coefficient of variation (CV) presented are summarized from a total of 7 surveys and 21 samples from the Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP) and Institute of Quality Management in Health Care (IQMH) proficiency testing surveys, and reported from 311 participating clinical laboratories in RCPAQAP and 86 IQMH Proficiency Testing Program. This data represents lithium measurements from eight instrument groups (Abbott Architect/Alinity c, Beckman Coulter AU, Beckman Coulter Unicel DxC, Ortho Vitros, Roche cobas c/Integra 400, Siemens Advia/Atellica, Siemens Dimension, and Siemens Vista) based on the colorimetric method principle.

SOURCES OF VARIATION IN LITHIUM THERAPEUTIC RANGES

Based on the proficiency testing survey results, we further explored whether there are other sources of variation in lithium therapeutic range such as method traceability, use of alternate method principles, and the use of outdated reference sources (e.g., manufacturer instructions for use (IFU), textbooks, publications, or clinical practice guidelines). There are currently 5 registered reference methods and 8 registered reference materials for lithium in serum or plasma in the Joint Committee for Traceability in Laboratory Medicine (JCTLM) Database [19]. Current routine commercially available lithium methods may be broadly categorized, from most common to least common as colorimetric, ion selective electrode, and atomic absorption spectrophotometry. Common colorimetric methods from 5 major manufacturers were further reviewed for traceability and were traceable to at least four different NIST standards (i.e., SRM956, SRM3129, SRM924, SRM 909). Review of recent 2022 proficiency testing survey reports from College of American Pathologists (CAP) based in the USA showed that although most methods are generally agreeable, there are some rarer methods, such as direct ion selective electrode, can have a bias of up to +0.3 mmol/L when compared to the all-methods' means. These biases may be present due to method specific differences and interferences, or the initial versions of a commercial assay released at a time when reference methods and/or materials are not available, or if the method's calibration traceability has not

been updated. Thus prior to adoption and implementation of the recommended standardized ISBD OABD therapeutic ranges, we continue to recommend a review of local, site-specific laboratory and clinical data.

Therapeutic ranges for lithium published from clinical practice guidelines, manufacturer IFUs, and textbooks were also reviewed. Reviews of recent clinical practice guidelines showed that target range varies, with the lower limit ranging from 0.4 to 0.6 mmol/L and upper limit ranging from 0.6 to 1.2 mmol/L [3,15]. Manufacturer IFUs of lithium assays from the five major vendors were reviewed, and the lithium therapeutic ranges and their reference source are summarized in Table 2. The main cited sources from these IFUs are based on the Tietz Textbook of Clinical Chemistry and Molecular Diagnostics [20,21], and the Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics [22,23]. The Tietz textbooks and Bakerman's interpretive laboratory reference account as the sources of the most used therapeutic ranges noted in the practice surveys [21–28]. Interestingly, newly commercially available analyzers, such as the Siemens Atellica or the Abbott Alinity, did not provide an update to their lithium IFU and continued to cite the original reference source published with their predecessors [29,30]. Additionally, some of the textbook editions did not reference the original source of their recommended ranges. For example, the latest edition of the Tietz Textbook provided an updated recommended range of 0.5 to 1.0 mmol/L for all age groups, however this modification did not provide a new reference source and continued to reference an older edition of the textbook [25,26].

Table 2 Common lithium methods therapeutic range referenced in manufacturers instructions for use

Source of therapeutic range from Manufacturer Instruction for Use (IFU) or Textbook	Therapeutic range for lithium level (mmol/L)	Reference cited
Abbott Architect [32]	1.0 – 1.2	Tietz Textbook 4th ed. [21]
Abbott Alinity [29]	1.0 – 1.2	Tietz Textbook 4th ed. [21]
Beckman AU [33]	1.0 – 1.2 (trough) 0.6 (minimum effective)	Tietz Fundamentals 6th ed. [23]
Beckman Synchron and DxC [34]	1.0 – 1.2 (trough) 0.6 (minimum effective)	Tietz Fundamentals 6th ed. [23]
Siemens Advia [35]	1.0 – 1.2	Tietz Textbook 2nd ed. [20]

Siemens Atellica [30]	1.0 – 1.2	Tietz Textbook 2nd ed. [20]
Siemens Dimension Vista [36]	0.6 – 1.2	Tietz Fundamentals 6th ed. [23]
Ortho Vitros [37]	0.6 – 1.2	Tietz Fundamentals 5th ed. [22]
Roche cobas [38]	0.6 – 1.2	Tietz Fundamentals 5th ed. [22]
Roche Direct ISE [39]	0.6 – 1.2	Tietz Clinical Guide 3rd ed. [28]
Tietz Textbook 6th ed. [26]	0.5 – 1.0	Tietz Textbook 5th ed. [25]
Bakerman's ABC's Interpretive Laboratory Data 5th ed. [24]	0.5 – 1.2 (acute mania) 0.5 – 1.0 (sustained prophylactic)	Practice Guideline 2002 Am J Psych [40]

For comparison, the ISBD task force on OABD has made specific recommendations on reporting for older adults with target therapeutic ranges for ages 60 to 79 in the range of 0.4 to 0.8 mmol/L, and for those 80 and over in the range of 0.4 to 0.7 mmol/L [17].

SURVEY OF LITHIUM REPORTING PRACTICE IN CANADA

Two surveys querying the Canadian clinical laboratories about their serum lithium reporting practices were conducted in 2017 and subsequently in 2022. Both surveys were conducted a few months prior to educational sessions on the safe and effective use of lithium in OABD presented at national and international clinical laboratory conferences (e.g., Canadian Society of Clinical Chemists Annual Scientific Meeting and International Association of Therapeutic Drug Monitoring and Clinical Toxicology Congress) [31].

The first survey was administered by IQMH with most laboratories representing the provinces of Ontario and Newfoundland. In May 2017, a 5-question voluntary survey about laboratory practices related to serum lithium collection instructions, reporting of therapeutic ranges and toxic levels were included with the IQMH DRUG proficiency testing survey and sent to clinical laboratories who subscribed to the program. Qualitative comments were received from a total of 85 laboratories that perform lithium testing, with a distribution of 77 (91%) hospital laboratories and 8 (9%) community laboratories. Of the 85 laboratories, 3 (4%) laboratories did not provide their lithium therapeutic ranges, 7 (8%) laboratories did not provide their lithium toxic alert concentration threshold, 2 (2%) laboratories provided pediatric (<18 years old) specific ranges, and 4 (5%) laboratories provided geriatric (≥65 years old) specific ranges. The surveyed lower therapeutic limit varied from 0.0 to 0.8 mmol/L, and the upper limit varied from 1.1 to 1.5 mmol/L, for all age groups (Figure 2). Like the Delphi survey results conducted by ISBD OABD Task Force, the majority (89%) of laboratories surveyed provided a single lithium therapeutic range for all age groups, and the most common range (62%) reported was 0.6 to 1.2 mmol/L [16]. For the surveyed

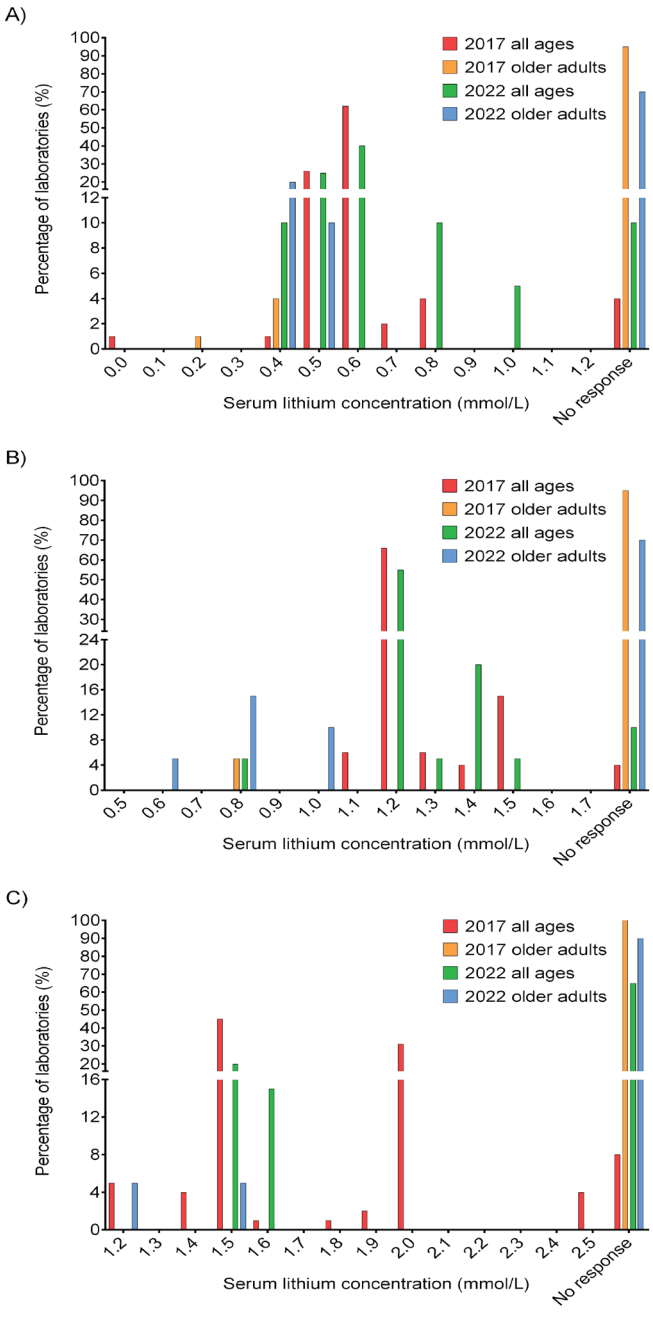
geriatric therapeutic ranges (≥65 years old), the lower limit of the therapeutic range varied from 0.2 to 0.4 mmol/L, and the upper limit was 0.8 mmol/L for all four sites. In terms of toxic alert levels, the upper threshold varied from 1.2 to 2.5 mmol/L and the most common toxic upper threshold is 1.5 mmol/L (45%) for all ages. No laboratory reported a separate toxic range for patients ≥65 years old. The therapeutic ranges for lithium are based on trough ranges. Forty-four of the 85 laboratories (53%) indicated they had collection instructions for measurement of lithium trough levels with 41 (91%) laboratories used the 12 hours post-dose instructions and 3 (9%) laboratories used the immediately prior to next dose instruction.

In May 2022, a second survey including 5 questions was circulated to members of the Canadian Society of Clinical Chemists (CSCC) via the CSCC electronic mailing list to determine whether the relevant clinical recommendations and educational efforts made an impact in raising awareness about the need for age-stratified therapeutic ranges for lithium in older adults. Qualitative responses were received from a total of 20 laboratories performing lithium testing from British Columbia, Manitoba, Ontario, and Quebec in Canada as well as Minnesota in the USA. Most of the responses (95%) were from hospital laboratories. Not every respondent answered all questions, but all responses received were included in the final survey report. Of the 20 laboratories who responded, 14 (66%) laboratories provided a single therapeutic range, and the most common therapeutic range used was 0.6 to 1.2 mmol/L (30%), followed by 0.4 to 1.4 mmol/L (10%), 0.5 to 1.2 mmol/L (10%), and 0.5 to 1.3 mmol/L (10%). Six (29%) laboratories provided a separate lithium therapeutic range for older adults, where the age limit varied between 60 to 65 years of age and over (Figure 2). Two of those six laboratories additionally provided a separate lithium therapeutic range for ages ≥80. For the geriatric population (age 60 and over) therapeutic lower limit varied from 0.4 to 0.6 mmol/L, and the upper limit varied from 0.6 to 1.0 mmol/L. The most common surveyed toxic alert limit was ≥1.5 mmol/L for all ages (35%). Toxic alert for older adults was reported by two laboratories with upper thresholds of 1.1 and 1.4 mmol/L. Eight (38%) laboratories provided collection instructions, and two (10%) laboratories provided interpretative comments regarding toxicity concentrations. The practice surveys in 2017 and 2022 suggest a slow but increasing adoption and implementation of age-specific therapeutic ranges for lithium (Figure 3).

DISCUSSION AND CONCLUSION

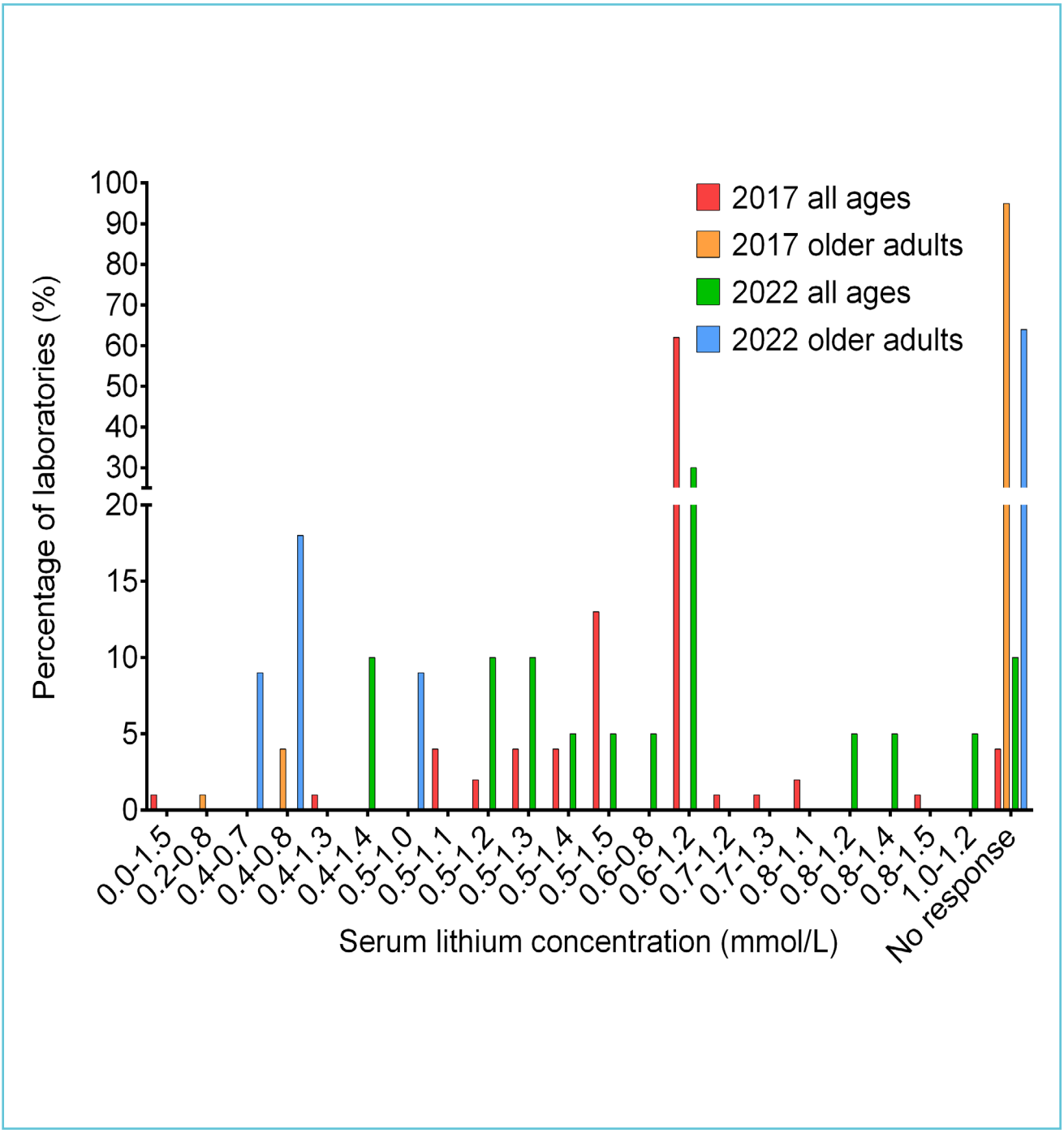
The Older Adults Task Force within the International Society for Bipolar Disorder (ISBD) has recommended age-stratified lithium therapeutic ranges for older adults with bipolar disorder (i.e., for ages 60 to 79 in the range of 0.4 to 0.8 mmol/L, and for ages ≥80 in the range of 0.4 to 0.7 mmol/L [17]). Here, we determine the feasibility of using standardized therapeutic ranges for lithium. Our analysis demonstrates that there is an association of lower serum lithium concentration with increased age, and there is generally good agreement between commonly used colorimetric lithium methods. Additional assessment of feasibility is required if using uncommon methods (i.e., ion selective electrode). Together, these data generally align the ISBD OABD Task Force recommendation with laboratory evidence. Interestingly, reporting practice surveys in Canada indicated that there is significant variability in the reporting of serum lithium therapeutic ranges with some laboratories reporting upper limit >1.2 mmol/L and up to 1.5 mmol/L. Review of potential sources of variation in therapeutic ranges shows that the upper therapeutic limit referenced from a variety of sources is generally not greater than 1.2

Figure 2 **Survey results for serum lithium reporting practices: variation in the (A) lower limit; (B) upper limit of therapeutic range; (C) toxic limit reported by laboratories**



Red denotes 2017 survey for all ages, orange denotes 2017 survey for older adults, green denotes 2022 survey for all ages, and blue denotes 2022 survey for older adults. For all age groups, majority of laboratories report a lower limit of 0.6 mmol/L and an upper limit of 1.2 mmol/L for therapeutic range, and ≥1.5 mmol/L for toxic limit. Considerable variability exists for both the therapeutic and toxic limits. The definition of older adult was variable and ranged between 60 to 65 years old. For older adults, majority of laboratories currently do not report age-stratified therapeutic ranges or toxic limits. This figure was generated by GraphPad Prism 5 software.

Figure 3 **Survey results for serum lithium reporting practices: variation in the therapeutic range reported by laboratories**



Red denotes 2017 survey for all ages, orange denotes 2017 survey for older adults, green denotes 2022 survey for all ages, and blue denotes 2022 survey for older adults. Majority of laboratories report a therapeutic range of 0.6 – 1.2 mmol/L for all ages, and a lack of age-stratified ranges defined for older adults. There is an increase in adoption of the ISBD recommended ranges for OABD from 2017 (orange) to 2022 (blue) in the therapeutic ranges of 0.4-0.7 and 0.4-0.8 mmol/L (i.e., for ages 60 to 79 in the range of 0.4 to 0.8 mmol/L, and for ≥80 years old in the range of 0.4 to 0.7 mmol/L [17].) This figure was generated by GraphPad Prism 5 software.

mmol/L, which suggests that reporting upper limits >1.2 mmol/L is an outdated practice. An upper therapeutic limit of >1.2 mmol/L can put older adults at a risk of lithium toxicity without being recognized by clinicians who may consider this level to be within the normal range. Although the Delphi survey did not make a specific recommendation on toxic alert concentrations, it has been suggested that 1.5 mmol/L is a practical, clinically-based toxicity alert for older adults. Together this highlights an important need for clinical laboratories to periodically review reference and therapeutic ranges and update obsolete ranges where clinically necessary. In addition, there is also a need to engage text book authors, editors, and manufacturers to review validity of their published lithium therapeutic ranges, and to include ISBD OABD recommendation through collaboration with national and international clinical chemistry and toxicology societies such as the Canadian Society of Clinical Chemists (CSCC), American Association for Clinical Chemistry (AACC), International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), and International Association of Therapeutic Drug Monitoring and Clinical Toxicology (IATDMCT).

In conclusion, the lack of age-stratified lithium therapeutic ranges may put older adults at risk of developing lithium toxicity as some potentially toxic results may be disregarded as "within the therapeutic range". Adoption and implementation of clinically appropriate, age-stratified therapeutic ranges for OABD have been slowly but steadily increasing in Canada. A small group of clinical laboratories in Ontario have championed the implementation of revised therapeutic lithium ranges for OABD and are now sharing their experience with interested laboratories in other provinces. Champion leaders have also been identified in provincial laboratory groups in British Columbia and Alberta to drive implementation across Canada where applicable. While we transform awareness of the ISBD OABD Task Force recommendations into action in Canada, we hope data presented in this article will help raise awareness and promote the safe and effective use of lithium in patients with OABD globally.

❖❖❖❖❖

Acknowledgements

We thank the members and administrative support from the Institute for Quality Management in Healthcare Chemistry Scientific Committee, and the Canadian Society of Clinical Chemists (CSCC) in facilitating the clinical laboratory practice surveys in Canada.

Conflict disclosures

None to disclose.

Author contributions

A.F., P.Y. K.S, and L.F. devised the project, main conceptual ideas, and outline. L.F., P.Y., and J.S., planned and carried out the voluntary practice surveys. A.F., V.Y., D.K., H.V., J.S., L.F. carried out data analysis and prepare data table and figures. A.F., K.S., D.K., V.Y., P.Y., and L.F. wrote the manuscript. All authors provided critical feedback and contributed to the final manuscript.

Ethics

As confirmed by the Sunnybrook Health Sciences Centre Research Ethics Board (REB), the retrospective study on association of serum lithium concentration and patient age groups did not require REB approval as it was deemed to be a quality improvement project and not human subject research.

❖❖❖❖❖

REFERENCES

1. M. Sajatovic, S.A. Strejilevich, A.G. Gildengers, A. Dols, R.K. al Jurdi, B.P. Forester, L.V. Kessing, J. Beyer, F. Manes,

S. Rej, A.R. Rosa, S.N. Schouws, S.-Y. Tsai, R.C. Young, K.I. Shulman, A report on older-age bipolar disorder from the International Society for Bipolar Disorders Task Force., Bipolar Disord. 17 (2015) 689–704. https://doi.org/10.1111/bdi.12331.

2. D. v. Jeste, G.S. Alexopoulos, S.J. Bartels, J.L. Cummings, J.J. Gallo, G.L. Gottlieb, M.C. Halpain, B.W. Palmer, T.L. Patterson, C.F. Reynolds, B.D. Lebowitz, Consensus statement on the upcoming crisis in geriatric mental health: research agenda for the next 2 decades, Arch Gen Psychiatry. 56 (1999) 848–853. https://doi.org/10.1001/ARCHPSYC.56.9.848.

3. G.S. Malhi, D. Gessler, T. Outhred, The use of lithium for the treatment of bipolar disorder: Recommendations from clinical practice guidelines, J Affect Disord. 217 (2017) 266–280. https://doi.org/10.1016/J.JAD.2017.03.052.

4. J. Fotso Soh, S. Klil-Drori, S. Rej, Using Lithium in Older Age Bipolar Disorder: Special Considerations, Drugs Aging. 36 (2019) 147–154. https://doi.org/10.1007/S40266-018-0628-1.

5. A. Cipriani, K. Hawton, S. Stockton, J.R. Geddes, Lithium in the prevention of suicide in mood disorders: updated systematic review and meta-analysis, BMJ. 346 (2013). https://doi.org/10.1136/BMJ.F3646.

6. C. Fitzgerald, R.H.B. Christensen, J. Simons, P.K. Andersen, M.E. Benros, M. Nordentoft, A. Erlangsen, K. Hawton, Effectiveness of medical treatment for bipolar disorder regarding suicide, self-harm and psychiatric hospital admission: Between- and within-individual study on Danish national data, British Journal of Psychiatry. 221 (2022) 692–700. https://doi.org/10.1192/bjp.2022.54.

7. L.V. Kessing, J.L. Forman, P.K. Andersen, Does lithium protect against dementia?, Bipolar Disord. 12 (2010) 87–94. https://doi.org/10.1111/J.1399-5618.2009.00788.X.

8. S. Rej, N. Herrmann, K. Shulman, H.D. Fischer, K. Fung, A. Gruneir, Current psychotropic medication prescribing patterns in late-life bipolar disorder, Int J Geriatr Psychiatry. 32 (2017) 1459–1465. https://doi.org/10.1002/GPS.4635.

9. S. Rej, K. Shulman, N. Herrmann, Z. Harel, H.D. Fischer, K. Fung, A. Gruneir, Prevalence and correlates of renal disease in older lithium users: a population-based study, Am J Geriatr Psychiatry. 22 (2014) 1075–1082. https://doi.org/10.1016/J.JAGP.2014.01.015.

10. D.N. Juurlink, M.M. Mamdani, A. Kopp, P.A. Rochon, K.I. Shulman, D.A. Redelmeier, Drug-induced lithium toxicity in the elderly: a population-based study, J Am Geriatr Soc. 52 (2004) 794–798. https://doi.org/10.1111/J.1532-5415.2004.52221.X.

11. M. Sun, N. Herrmann, K.I. Shulman, Lithium Toxicity in Older Adults: a Systematic Review of Case Reports, Clin Drug Investig. 38 (2018) 201–209. https://doi.org/10.1007/S40261-017-0598-9.

12. C. Marras, N. Herrmann, H.D. Fischer, K. Fung, A. Gruneir, P.A. Rochon, S. Rej, S. Vigod, D. Seitz, K.I. Shulman, Lithium Use in Older Adults is Associated with Increased Prescribing of Parkinson Medications, Am J Geriatr Psychiatry. 24 (2016) 301–309. https://doi.org/10.1016/J.JAGP.2015.11.004.

13. P.A. Rochon, J.H. Gurwitz, The prescribing cascade revisited, Lancet. 389 (2017) 1778–1780. https://doi.org/10.1016/S0140-6736(17)31188-1.

14. R.C. Young, B.H. Mulsant, M. Sajatovic, A.G. Gildengers, L. Gyulai, R.K. al Jurdi, J. Beyer, J. Evans, S. Banerjee, R. Greenberg, P. Marino, M.E. Kunik, P. Chen, M. Barrett, H.C. Schulberg, M.L. Bruce, C.I.F. Reynolds, G.S. Alexopoulos, GERI-BD: A Randomized Double-Blind Controlled Trial of Lithium and Divalproex in the Treatment of Mania in Older Patients With Bipolar Disorder, Am J Psychiatry. 174 (2017) 1086–1093. https://doi.org/10.1176/APPI.AJP.2017.15050657.

15. A. Dols, L.V. Kessing, S.A. Strejilevich, S. Rej, S.Y. Tsai, A.G. Gildengers, O.P. Almeida, K.I. Shulman, M. Sajatovic, Do current national and international guidelines have specific recommendations for older adults with bipolar disorder? A brief report, Int J Geriatr Psychiatry. 31 (2016) 1295–1300. https://doi.org/10.1002/GPS.4534.

16. K.I. Shulman, O.P. Almeida, N. Herrmann, A. Schaffer, S.A. Strejilevich, C. Paternoster, S. Amodeo, A. Dols, M. Sajatovic, Delphi survey of maintenance lithium treatment in older adults with bipolar disorder: An ISBD task force report, Bipolar Disord. 21 (2019) 117–123. https://doi.org/10.1111/BDI.12714.

17. K.I. Shulman, M. Sajatovic, A. Dols, Laboratories should provide a separate therapeutic range for serum lithium levels in maintenance treatment of older adults with bipolar disorder (OABD), Bipolar Disord. 21 (2019) 190–191. https://doi.org/10.1111/BDI.12769.

18. S. Pauwels, P. Vermeersch, K. Desmet, F. Vanstapel, Practice-oriented quality specifications for therapeutic drug monitoring, Clin Chem. 60 (2014) 559–560. https://doi.org/10.1373/clinchem.2013.217125.

19. Joint Committee for Traceability in Laboratory Medicine (JCTLM) Database: higher-order reference materials, methods and services, (2023). https://www.jctlmdb.org/#/app/home (accessed March 17, 2023).

20. C.A. Burtis, E.R. Ashwood, eds., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 2nd ed., WB Saunders Co., Philadelphia, PA, 1994.

21. C.A. Burtis, E.R. Ashwood, D.E. Bruns, eds., Tietz Textbook of Clinical Chemistry and Molecular diagnostics, 4th ed., Saunders Elsevier, St. Louis, MO, 2006.

22. C.A. Burtis, E.R. Ashwood, eds., Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics, 5th ed., WB Saunders, Philadelphia, 2001.

23. C.A. Burtis, E.R. Ashwood, D.E. Bruns, eds., Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics, 6th ed, Saunders Elsevier, St. Louis, MO, 2008.

24. S. Bakerman, Bakerman's ABC's of Interpretive Laboratory Data, 5th ed., Interpretive Laboratoray Data Inc., Scottsdale, 2014.

25. W. Roberts, McMillan GA, eds., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 5th ed., Elsevier, St. Louis, 2013.

26. N. Rifai, A.R. Horvath, C.T. Wittwer, eds., Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 6th ed., Elsevier, St. Louis, MO, 2018.

27. C.A. Burtis, D.E. Bruns, eds., Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics, 7th ed, Elsevier, St. Louis, MO, 2018.

28. N.W. Tietz, P.R. Finley, eds., Tietz Clinical Guide to Laboratory Tests, 3rd ed., WB Saunders Co., United Kingdom, 1995.

29. Abbott Alinity c System Lithium REF 08P5320 Rev Jan 2018 Instructions for Use, (2018).

30. Siemens Healthcare Diagnostics Atellica CH analyzer Lithium REF 11532401 V2.0, (2020).

31. H. Vandenberghe, D. Konforte, J. Stemp, L. Fu, K.I. Shulman, P.M. Yip, How are laboratories in Ontario reporting lithium for therapeutic drug monitoring?, Clin Biochem. 61 (2018) 50–71. https://doi.org/10.1016/j.clinbiochem.2018.09.005.

32. Abbott Laboratories Architect cSystem Lithium REF 8L25-01 V2.0, (2009).

33. Beckman Coulter AU Chemistry analyzers Lithium REF OSRT056 Instructions for Use, (n.d.).

34. Beckman Coulter Synchron and DxC systems Lithium REF JL840897 Instruction for Use, (2013).

35. Siemens Healthcare Diagnostics Advia 1800 and 2400 Chemistry Systems Lithium REF 00141230 Instructions for Use Rev. J, (2020).

36. Siemens Healthcare Diagnostics Dimension and Dimension Vista System Lithium REF K4150 Rev E, (2019).

37. Ortho Clinical Diagnostics Vitros Lithium REF C-372 Instructions for Use V3.0, 2012.

38. Roche Diagnostics cobas Lithium REF 08057974190 Instruction for Use V4.0, (2021).

39. Roche Diagnostics cobas Lithium ISE REF 03003523001 V10.0, (2021).

40. R.M.A. Hirschfeld, C.L. Bowden, M.J. Gitlin, P.E. Keck, R.H. Perlis, T. Suppes, M.E. Thase, K.D. Wagner, Practice Guideline for the Treatment of Patients With Bipolar Disorder (Revision), Focus (Madison). 1 (2003) 64–110. https://doi.org/10.1176/foc.1.1.64.

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Predictive value of the platelet times neutrophil-to-lymphocyte ratio (SII index) for COVID-19 in-hospital mortality

Santiago J. Ballaz[1], Martha Fors[2]

[1] *Medical School, Universidad Espíritu Santo, Samborondón, Ecuador*

[2] *Medical School, Universidad de las Américas-UDLA, Quito, Ecuador*

ARTICLE INFO

Corresponding author:
Santiago J. Ballaz, PhD
Medical School
Universidad Espíritu Santo
Vía Samborondón
0901952 Samborondón
Ecuador
Phone: +593 0992393691
E-mail: sballazg@gmail.com

Key words:
biomarkers, COVID-19, mortality, ROC curve, severity of illness index

ORCIDs
0000-0001-5878-8679 (Santiago J. Ballaz)
0000-0002-0844-199X (Martha Fors)

ABSTRACT

Objective

The performance of the platelet times neutrophil-to-lymphocyte ratio, namely systemic immune inflammation (SII) index, is an inflammatory index that shows controversial results as a predicting indicator of the poor outcomes of COVID-19. In this study, this indicator was analyzed in 3280 patients admitted at a COVID-19 reference hospital in Quito (Ecuador).

Methods

The Receiver Operating Characteristic (ROC) curve analysis was conducted on SII values upon admission to identify the most appropriate cut-off values in discriminating COVID-19 severity and in-hospital mortality.

Results

SII was higher in both severe patients and in those who finally died (cut-off points of 757.3 and 808.5

respectively). However, the AUC-ROC analysis (0.60-0.67) demonstrated a modest discriminating performance of SII for COVID-19 severity (61.2% sensitivity and 61.5% specificity), which sensibly improved for COVID-19 mortality (AUC-ROC: 0.73-0.83, sensitivity: 80.6% specificity; 63.6%).

Conclusion

SII index may well be an indicator of inflammatory conditions secondary to COVID-19 leading to a higher mortality, rather than a predictor of severe forms of the disease.

❖❖❖❖❖

Abbreviations

AUC-ROC, Area Under the Curve-Receiver Operating Characteristic;

NLR, Neutrophil-to-Lymphocyte Ratio;

SII, Platelet by Neutrophil-to-Lymphocyte Ratio.

❖❖❖❖❖

1. INTRODUCTION

During SARS-CoV-2 infection, the activation of hyper-inflammation, cytokine storm, coagulopathies, and disseminated intravascular coagulation are the most prevalent pathobiological processes leading to severe acute respiratory distress syndrome (ARDS) and multi-organ failure. This is the rationale of including platelet counts in the development of new blood-cell based composed inflammation indices for the forecasting of COVID-19 patients. The systemic immune inflammation index (SII= (platelet counts × neutrophil counts)/lymphocyte counts), a composed blood parameter originally set to assist with the diagnosis, progress, and risk stratification of inflammatory diseases [1], can also predict COVID-19 severity and mortality [2-9]. Despite the evidence, doubts exist regarding the prognostic performance of SII [10,11], which may fall short as a biomarker of poor clinical outcomes, especially in hospitalized COVID-19 patients who present comorbidities [12,13]. The use of patients at different stages of the disease, or even the disparate sizes of the samples used (from 119 to 1800 patients), would account for the controversy. In this study, it was investigated the association between SII values and COVID-19 severity and in-hospital mortality using the records of a large sample of patients from a COVID-19 reference hospital of Northern Ecuador. During the epidemic, Ecuador had the highest death toll in all South America.

2. METHODS

This observational retrospective study included 3280 consecutive patients over 18 years old, who were admitted at the IESS Hospital Quito Sur in Quito (Ecuador) from March 13 to June 17, 2020 with COVID-19 (CDC 2019-Novel Coronavirus Real-Time RT-PCR Diagnostic Panel in upper and lower respiratory specimens) and showing COVID-19-like symptoms. Blood samples for routine laboratory tests were drawn upon admission (after a 40-min interval of average) and subsequently run in a Sysmex XN-550™ Hematology Analyzer (Sysmex America Inc., Japan). In addition, an arterial blood gas exam was also conducted on room air under a controlled oxygen environment at the time of admission using a RAPIDPoint® 500 blood gas system (Siemens Healthcare GmbH; Germany). Patient categorization was conducted in agreement with the NIH guidelines (https://www.covid19treatmentguidelines.nih.gov/overview/clinical-spectrum/) for the severity of COVID-19 pneumonia. The "Severe group" included a total of 635 patients who had blood hypoxemia (PaO_2 < 60 mmHg; SpO_2 < 94%). The rest (2645 cases) were classified as "Non-Severe" (PaO_2 ≥ 60 mmHg and SpO_2 ≥ 94%.). Discrepant cases showing values of PaO_2 < 60 mmHg and SpO_2 ≥ 94% were not included in the analysis.

Non normally distributed variables were determined by the Shapiro Wilk test and expressed as medians and interquartile ranges (IRs). The Mood test was chosen to compare the medians of NLR and SII (annotated as P*N/L) across COVID-19 severity and in-hospital mortality/survival groups. Odds ratios (ORs) were calculated using a simple logistic regression as well as a multiple logistic regression analysis when convenient to estimate the likelihood of severity and mortality on a multiple variable basis of confounding factors such as age, sex and the SII index. The Receiver Operating Characteristic curve (AUC-ROC) analysis was conducted to evaluate the discriminative performance of the SII and NLR indices as well as to identify the optimal cut-off points of both the maximum sensitivity and specificity (Youden index). Concerning the performance criterion, AUC values were interpreted as follows: 0.5-0.6 (failed), 0.6-0.7 (worthless), 0.7-0.8 (poor), 0.8-0.9 (good), > 0.9 (excellent) [14]. Alpha value was set at 0.05. Written informed consent was waived due to the use of secondary data obtained from anonymized patients (Public Health Ministerial order of December 31st, 2014). The STROBE (Strengthening The Reporting of OBservational Studies in Epidemiology) guidelines were followed in reporting this study.

3. RESULTS

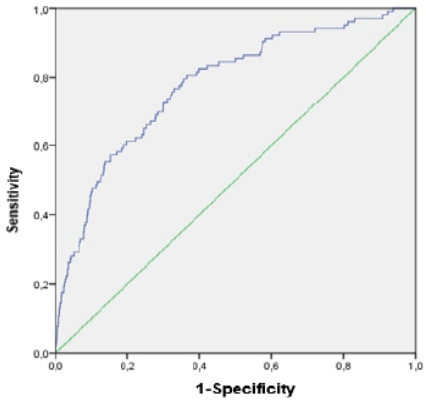
The sample included similar numbers of men (1643) and women (1637). There were 635 Severe cases (19.4%) and 2645 non-Severe cases (80.6%). The median age was as follows (years (SD)): Non-Severe 42.1 (22.3), Severe 42.9 (21.7), Survivor 42.2 (22.2), non-Survivor 45.5 (22.3). Mortality rate (3.1%, two thirds of whom were men) in the Severe group was higher than in the non-severe group (4.9% and 2.7% respectively). The median of SII and NLR were significantly higher in both the Non-Survival and the Severe groups (Table 1).

In the AUC-ROC analysis (Figure 1), the optimal cut-off points for severity were 757.3 (SII) and 2.28 (NLR), while for mortality were 808.5 (SII) and 3.81 (NLR). The AUCs of the SII index were modest (0.66 for severity and 0.77 for mortality) compared to the NLR ratio (AUC values of 0.57 and 0.84 for severity and mortality respectively). The sensitivity and specificity of the SII index for severity were 61.2% and 61.5% respectively. As to mortality, the sensitivity of SII increased up to 80.6%, while its specificity barely changed (63.6%). For the association analysis, the patients were divided in two groups according to the SII cut-off points so that 58.8% of the patients (n=1929) were below the cut-off for severity and 62.2% of the patients (n=2041) for mortality. After considering potential confounders such as age, gender and SII, the logistic regression analysis revealed that those patients with levels of SII above the cut-off were more likely to either have severe COVID-19 (adjusted OR [95%CI]: 2.233(1.787-2.790)) or to die during hospitalization (adjusted OR [95%CI]: 5.690 (3.438-9.416)). The SII index was independently associated with mortality in patients hospitalized for COVID-19.

4. DISCUSSION

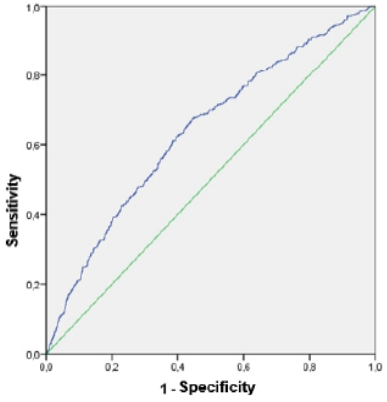
Despite the fact that COVID-19 mortality risk is undoubtedly associated with a high SII value [2], our ROC analysis demonstrated a modest discriminating power for SII in predicting COVID-19 severity upon admission compared to other studies [4-6,8,9,13], where the sensitivity and specificity as well as the AUC values were higher than in this study. Different categorization criteria for COVID-19 severity (blood hypoxemia versus ICU admission and invasive mechanical ventilation) suggests that SII cannot discriminate well disease severity at early stages of the disease as the NLR parameter does [15]. However, the discriminating power for SII notably improved when considering mortality

Figure 1 ROC curves of SII (top) and NLR (bottom) for mortality (left) and severity (right)



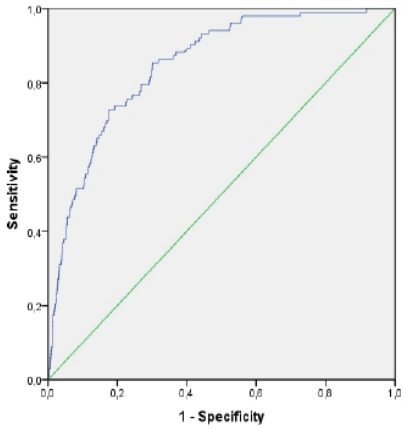
Mortality SII AUC (IC95%): 0.77 (0.73-0.83)

Cutoff 805. Sensitivity 80.6. Specificity 60.6 p < 0.001



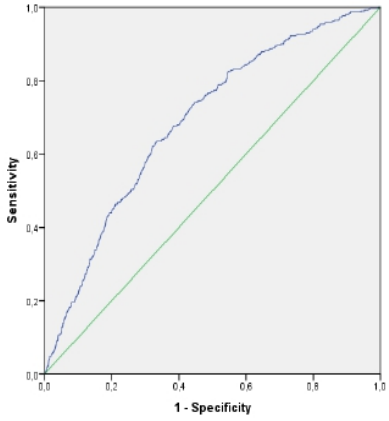
Severity SII AUC (IC95%): 0.63 (0.60-0.67)

Cutoff 757.34. Sensitivity 61.2 Specificity 61.5 p < 0.001



Mortality NLR AUC (IC95%): 0.84 (0-81-0.88)

Cutoff 3.81. Sensitivity 83.1 Specificity 70.0 p < 0.001



Severity NLR AUC (IC95%): 0.57 (0.55-0.60)

Cutoff 2.28. Sensitivity 66.0 Specificity 48.0 p < 0.001

Table 1 Medians of the SII index and NLR ratio in COVID-19 patients grouped according to mortality and severity

SII (P*N/L)	Median (IR)	Median (IR)	p-value*
Mortality	Yes n=103	No n=3177	< 0.001
	1766.7 (884.3-3339.0)	615.0 (379.7-1082.9)	
Severity	Severe n=635	Non-Severe n=2645	
	684.0 (421.7-1293.7)	609.0 (376.8-1090.4)	< 0.001
NLR			
Mortality	Yes n=103	No n=3177	< 0.001
	9.0 (4.7-15.0)	2.4 (1.5-4.3)	
Severity	Severe n=635	Non-Severe n=2645	
	2.9 (1.8-5.5)	2.4 (1.5-4.3)	< 0.001

**Mood test (median test).*

showing AUC values (0.73-0.83), which were in line with others [5,6-8]. The SII cut-off value for mortality (835) was in line with previous reports [2,3,12]. The sensitivity of SII for mortality (80.6%) was even higher to what was reported in other studies [2,7]. This improvement of the discriminating power of SII for mortality may partially be related to the ongoing increment of platelets counts, which along with an excessively high values of NLR often shown by COVID-19 patients, might provoke coagulopathies in late stages of the illness [16]. Although the analysis of records from a single institution may be considered a limitation, this study showed that the SII index at the time of admission was only associated with COVID-19 in-hospital mortality. Whereas a complex interaction between inflammation and hemostasis may be the reason for the modest performance of SII in the prediction of severe COVID-19 [11], the index may well be an indicator of inflammatory conditions secondary to COVID-19 disease, which could ultimately precipitate death [5,12,13,16].

❖❖❖❖❖

Acknowledgments

We thank Francisco Mora (IESS Quito Sur Hospital) who worked with us in the database management.

Compliance with ethical standards

Written informed consent for participation was not required for this study due to the use of secondary data in accordance with the national legislation and the institutional requirements. We followed the STROBE guideline to report this study. The authors declare they had no access to identifying patient information when analyzing the data.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests

The authors have no competing interests to declare.

Authors´ contributions statement

Both authors have contributed sufficiently and equally to this research work.

❖❖❖❖❖

REFERENCES

1. Putzu, C., Cortinovis, D.L., Colonese, F. Canova S, Carru C, Zinellu A, Paliogiannis P (2018) Blood cell count indexes as predictors of outcomes in advanced non-small-cell lung cancer patients treated with Nivolumab. Cancer Immunol Immunother 67:1349–1353. https://doi.org/10.1007/s00262-018-2182-4

2. Fois AG, Paliogiannis P, Scano V, Cau S, Babudieri S, Perra R, Ruzzittu G, Zinellu E, Pirina P, Carru C, Arru LB, Fancellu A, Mondoni M, Mangoni AA, Zinellu A. (2020) The systemic inflammation index on admission predicts in-hospital mortality in COVID-19 patients. Molecules 25:5725. https://doi.org/10.3390/molecules25235725

3. Rokni M, Ahmadikia K, Asghari S, Mashaei S, Hasanali F (2020) Comparison of clinical, para-clinical and laboratory findings in survived and deceased patients with COVID-19: diagnostic role of inflammatory indications in determining the severity of illness. BMC Infect Dis; 20:869. doi: 10.1186/s12879-020-05540-3

4. Gujar RK, Meena A, Chouhan SS, Likhar KS (2021) Hematological profiles of COVID-19 patients at the Ratlam district, Madhya Pradesh State, India. Bioinformation 17:686-690. doi: 10.6026/97320630017686

5. Moisa E, Corneci D, Negoita S, Filimon CR, Serbu A, Negutu MI, Grintescu IM (2021) Dynamic changes of the Neutrophil-to-Lymphocyte Ratio, Systemic Inflammation Index, and derived Neutrophil-to-Lymphocyte Ratio independently predict invasive mechanical ventilation need and death in critically ill COVID-19 patients. Biomedicines 9:1656. doi: 10.3390/biomedicines9111656

6. Halmaciu I, Arbănași EM, Kaller R, Mureșan AV, Arbănași EM, Bacalbasa N, Suciu BA, Cojocaru II, Runcan AI, Grosu F, Vunvulea V, Russu E (2022) Chest CT Severity Score and Systemic Inflammatory biomarkers as predictors of the need for invasive mechanical ventilation and of COVID-19 patients' mortality. Diagnostics (Basel) 12:2089. doi: 10.3390/diagnostics12092089

7. Ghobadi H, Mohammadshahi J, Javaheri N, Fouladi N, Mirzazadeh Y, Aslani MR (2022) Role of leukocytes and systemic inflammation indexes (NLR, PLR, MLP, dNLR, NLPR, AISI, SIR-I, and SII) on admission predicts in-hospital mortality in non-elderly and elderly COVID-19 patients. Front Med (Lausanne) 9:916453. doi: 10.3389/fmed.2022.916453

8. Mallappa S, Khatri A, Bn G, Kulkarni P (2022) A retrospective analysis of the importance of biochemical and hematological parameters for mortality prediction in COVID-19 cases. Cureus 14:e30129. doi: 10.7759/cureus.30129

9. Xia W, Tan Y, Hu S, Li C, Jiang T (2022) Predictive value of Systemic Immune-Inflammation index and neutrophil-to-lymphocyte ratio in patients with severe COVID-19. Clin Appl Thromb Hemost 28: 10760296221111391. doi: 10.1177/10760296221111391

10. Citu C, Gorun F, Motoc A, Sas I, Gorun OM, Burlea B, Tuta-Sas I, Tomescu L, Neamtu R, Malita D, Citu IM (2022) The predictive role of NLR, d-NLR, MLR, and SIRI in COVID-19 mortality. Diagnostics (Basel). 12:122. doi: 10.3390/diagnostics12010122

11. Nooh HA, Abdellateif MS, Refaat L, Kandeel EZ, Bayoumi A, Samra M, Khafagy M (2021) The role of inflammatory indices in the outcome of COVID-19 cancer patients. Med Oncol 39:6. doi: 10.1007/s12032-021-01605-8

12. Sevinc C, Demirci R, Timur O (2021) Predicting hospital mortality in COVID-19 hemodialysis patients with developed scores. Semin Dial 34:347-359. https://doi.org/10.1111/sdi.13004

13. Hamad DA, Aly MM, Abdelhameid MA, et al. (2022) Combined blood indexes of systemic inflammation

as a mirror to admission to intensive care unit in COVID-19 patients: a multicentric study. J Epidemiol Glob Health 12:64–73. doi:10.1007/s44197-021-00021-5

14. Polo TCF, Miot HA (2020) Use of ROC curves in clinical and experimental studies. J Vasc Bras 19:e20200186. doi: 10.1590/1677-5449.200186

15. Liu J, Liu Y, Xiang P, Pu L, Xiong H et al. (2020) Neutrophil-to-lymphocyte ratio predicts critical illness patients with 2019 coronavirus disease in the early stage. J Trans Med 18:206. https://doi.org/10.1186/s12967-020-02374-0

16. Wool GD, Miller JL (2021) The impact of COVID-19 disease on platelets and coagulation. Pathobiology 88:15-27. doi: 10.1159/000512007

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Utility of anti-GM-CSF antibodies in the diagnosis of pulmonary alveolar proteinosis: a case report

Antonio Sierra-Rivera, Jorge Ferriz-Vivancos, Marta Fandos-Sánchez, Pilar Teresa Timoneda-Timoneda, Goitzane Marcaida-Benito

Department of Clinical Biochemistry, Hospital General Universitario Valencia, Valencia, Spain

ARTICLE INFO

Corresponding author:
Antonio Sierra-Rivera
Department of Clinical Biochemistry
Hospital General Universitario
Av. de les Tres Creus, 2, 46014
Valencia
Spain
E-mail: asierrarivera@gmail.com

Key words:
pulmonary alveolar proteinosis, anti-GM-CSF, pulmonary surfactant, bronchoalveolar lavage

ORCID ID: 0000-0001-7824-7239

ABSTRACT

Introduction

Pulmonary alveolar proteinosis (PAP) is a disease characterized by the accumulation of lipoproteinaceous material in the alveoli as a consequence of deficient processing of pulmonary surfactant. It is classified into primary, secondary, and congenital forms. Primary PAP (autoimmune origin) is characterized by the presence of antibodies against granulocyte-macrophage colony-stimulating factor (GM-CSF), while secondary PAP is due to multiple causes such as exposure to certain environmental substances. We present a case of a patient with probable mixed PAP, primary and secondary, due to exposure at the patient's workplace.

Case presentation

A 35-year-old male patient attends the outpatient clinic of pulmonology due to symptoms of exertional

dyspnea for one year. Pulmonary function tests are performed, and the chest X-ray reveals diffuse bilateral lung involvement with a ground-glass pattern. Incision and excision lung biopsy show findings compatible with predominant PAP in the left lower lobe (LLL). Additionally, a positive anti-GM-CSF antibody result is obtained. The patient is treated with bronchoalveolar lavage (BAL) and nebulized sargramostim. The patient shows satisfactory progress.

Discussion

The clinical, analytical, radiological, and histological manifestations were compatible with the diagnosis of autoimmune PAP, and there was suspicion of secondary PAP due to exposure to rock wool. The role of the laboratory, in this case, was essential for the diagnostic confirmation of PAP by performing the determination of anti-GM-CSF antibodies.

❖❖❖❖❖

INTRODUCTION

Pulmonary alveolar proteinosis (PAP) is a rare diffuse interstitial disease characterized by excessive accumulation of lipoproteinaceous material derived from pulmonary surfactant in the alveolar spaces and terminal bronchioles, which can range from asymptomatic to severe respiratory failure (1).

In all its forms, the physiopathological substrate lies in the accumulation of surfactant in the alveolar spaces due to deficient activity in its processing by macrophages. GM-CSF is necessary for the final differentiation and maturation of alveolar macrophages (2).

PAP is classified into three types: primary (autoimmune), secondary, and congenital. The mechanisms that lead to macrophage dysfunction differ in each clinical form.

The primary or idiopathic variant (PAPi) is the most common form, accounting for up to 90% of cases. This clinical form is characterized by the presence of neutralizing IgG antibodies against granulocyte-macrophage colony-stimulating factor (anti-GM-CSF), which blocks the bioactivity of GM-CSF in vivo (3). These antibodies against GM-CSF affect the terminal differentiation of macrophages and, therefore, prevent the growth of these cells that are responsible for eliminating surfactant in the lungs (4).

Secondary forms are associated with hematological disorders or diseases (myelodysplastic syndromes, monoclonal gammopathies, leukemias, lymphomas), non-hematological neoplastic diseases, immunodeficiencies, chronic inflammatory syndromes, infections (Mycobacterium tuberculosis, Nocardia, Pneumocystis jirovecii), or exposure to various environmental substances such as silica, aluminum, titanium, or some fertilizers (5).

In the case of congenital PAP, the problem is associated with recessive anomalies in the gene that codes for the α (CSF2RA) and β (CSF2RB) chains of the GM-CSF receptor (6). It can also be secondary to mutation of GM-CSF or mutations in genes that code for surfactant proteins B (SP-B), C (SP-C) (7) and mutation of the ATP-Binding Cassette transporter A3 (ABCA3) (8).

Next, we describe a case of mixed PAP, of both primary and possibly secondary origin, that occurred in our hospital.

CASE REPORT

A 35-year-old male with a history of smoking since he was 16 years old (8-10 cigarettes/day) and employed for about 2 years in a fireproofing company processing paint and rock wool (material with high silica content). The patient reports working with machines that grind rock wool and project dust, so exposure to this material is constant. He also reports previous episodes

of rhinorrhea with the expulsion of black soot for two days. He has been experiencing exertional dyspnea for a year and a weight loss of 6 kg during this period. Subsequently, due to exacerbation of dyspnea and respiratory difficulty, he was referred to external pulmonology consultations. He has no other relevant medical history, does not have pets, nor does he engage in related activities.

On physical examination, he presented with a blood pressure of 121/79 mmHg, heart rate of 84 bpm, arterial blood gas with an O2 saturation of 92%, and an inspired oxygen fraction (FiO2) of 0.21. Additionally, exertional dyspnea was identified without cough, expectoration, fever, or thermal sensation, and no orthopnea. The patient reports occasional apneas with some asphyctic awakenings. The tests to evaluate the respiratory function, including forced vital capacity (FVC), forced expiratory volume in the first second (FEV1), and FEV1/FVC ratio were normal. On laboratory analysis, increased levels of lactate dehydrogenase (LDH) of 530 mg/dL (RI: 210-425 mg/dL) and angiotensin converting enzyme (ACE) of 75 U/L (RI: 20-70 U/L) were observed. The chest X-ray and high-resolution computed tomography (HRCT) revealed diffuse bilateral lung involvement with a ground-glass pattern and smooth thickening of the interlobular septa, adopting a paving stone pattern with a tendency to consolidation in the posterior region of the middle third of the left inferior lobe Figures 1 and 2.

The rest of the blood and urine tests were normal, including complete blood count, coagulation, antinuclear antibodies (ANA), and antineutrophil cytoplasmic antibodies (ANCA), complete biochemistry with liver, renal, and bone metabolism functions. Moreover, microbiological results were negative for Gram staining, fungal cultures, and PCR for *Adenovirus (A, B, C, D, and E), Parainfluenza virus (1, 2, 3, and 4), Rhinovirus (A, B, and C), Influenza A and B viruses, Metapneumovirus, Aspergillus, Pneumocystis jirovecii, and Mycobacteria.*

With these findings, a differential diagnosis was raised between: alveolar proteinosis (primary or secondary to rock wool inhalation) as the first possibility, or less likely sarcoidosis (no lymph node or pleural involvement), lipoid pneumonia (no areas of fatty density observed, and the involvement was extensive), infectious origin (bacterial, viral, and P. jiroveci cultures were negative), or bronchioalveolar carcinoma.

After discussing the case with the immunology service, determination of anti-GM-CSF antibodies was requested due to suspicion of primary PAP, and the results were positive with a figure of 8.2 U/mL (RI: <5 U/mL). The determination of anti-GM-CSF antibodies was performed using a ClinMax™ Human GM-CSF Quantitative ELISA kit, which is a standard sandwich immunoassay designed to quantify GM-CSF present in complex biological matrices such as human serum, plasma, and buffer solution. Additionally, the patient underwent a transbronchial biopsy in the left inferior lobe, and the samples were sent to the Pathology Department. The histological studies of the lung biopsy confirmed the diagnosis of PAP.

Based on the clinical, analytical, radiological, and anatomopathological findings, the patient was diagnosed with primary pulmonary alveolar proteinosis (PAP). The secondary origin could not be demonstrated due to the absence of previous sera to determine anti-GM-CSF concentrations. Combined treatment was decided upon, consisting of bronchoalveolar lavage (BAL) techniques along with nebulized sargramostim administration.

DISCUSSION

Pulmonary alveolar proteinosis (PAP) was formerly called Rosen Castleman Liebow syndrome,

Figure 1 Chest X-ray (1.A and 1.B) showing bilateral and diffuse lung involvement

Figure 1.A

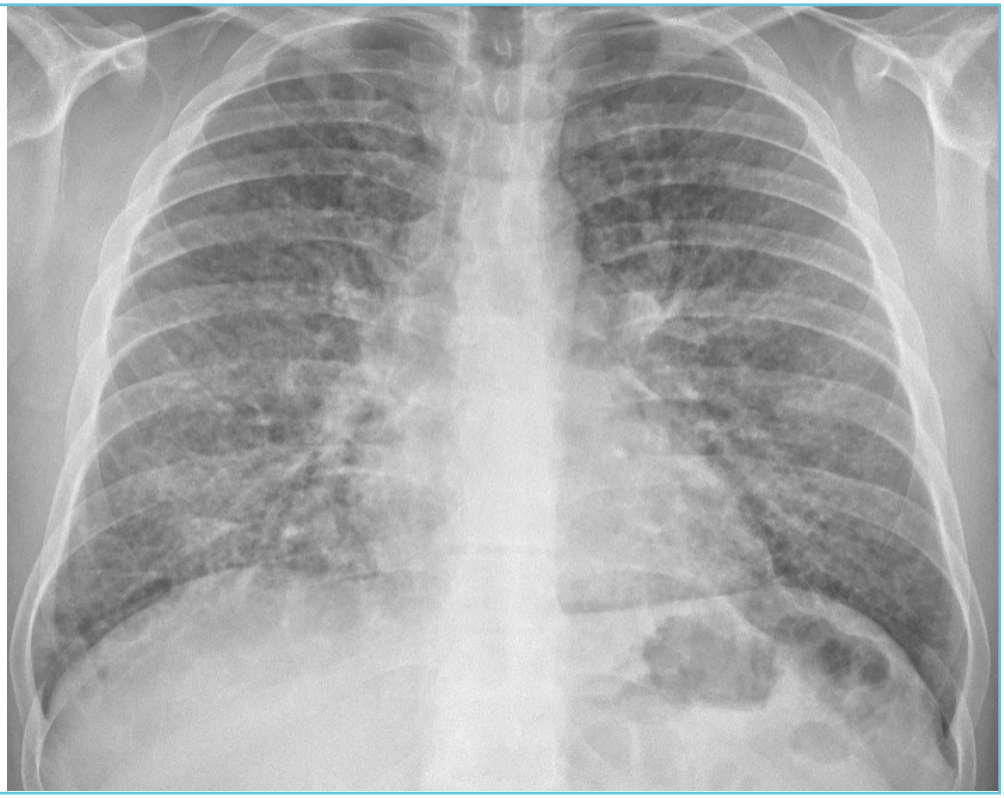


Figure 1.B

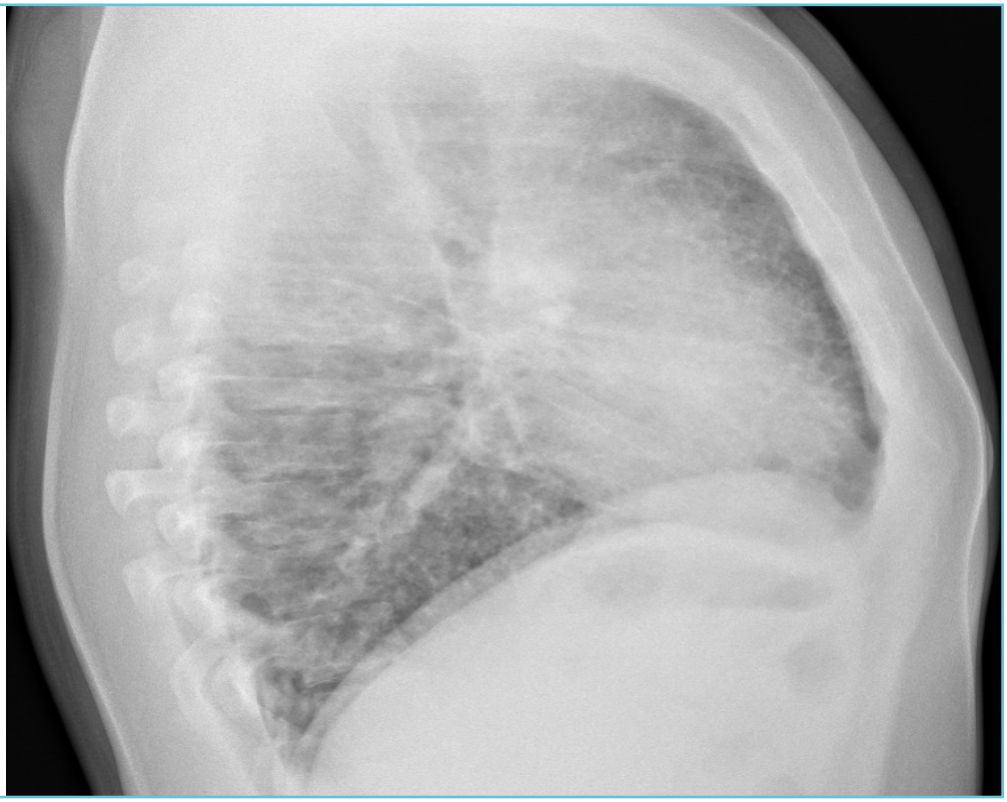


Figure 2 **High-resolution computed tomography (HRCT) of the chest (2.A and 2.B) showing bilateral and diffuse lung involvement with ground-glass opacities of a patchy appearance, with smooth thickening of the interlobular septa, adopting a reticular pattern known as "crazy paving"**

Figure 2.A

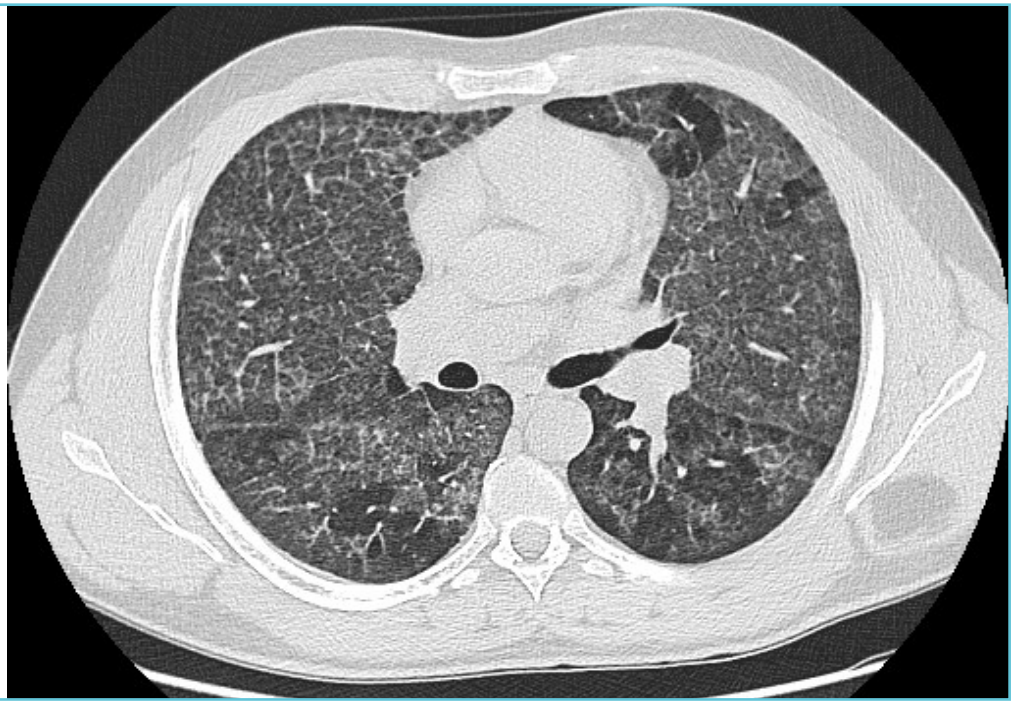
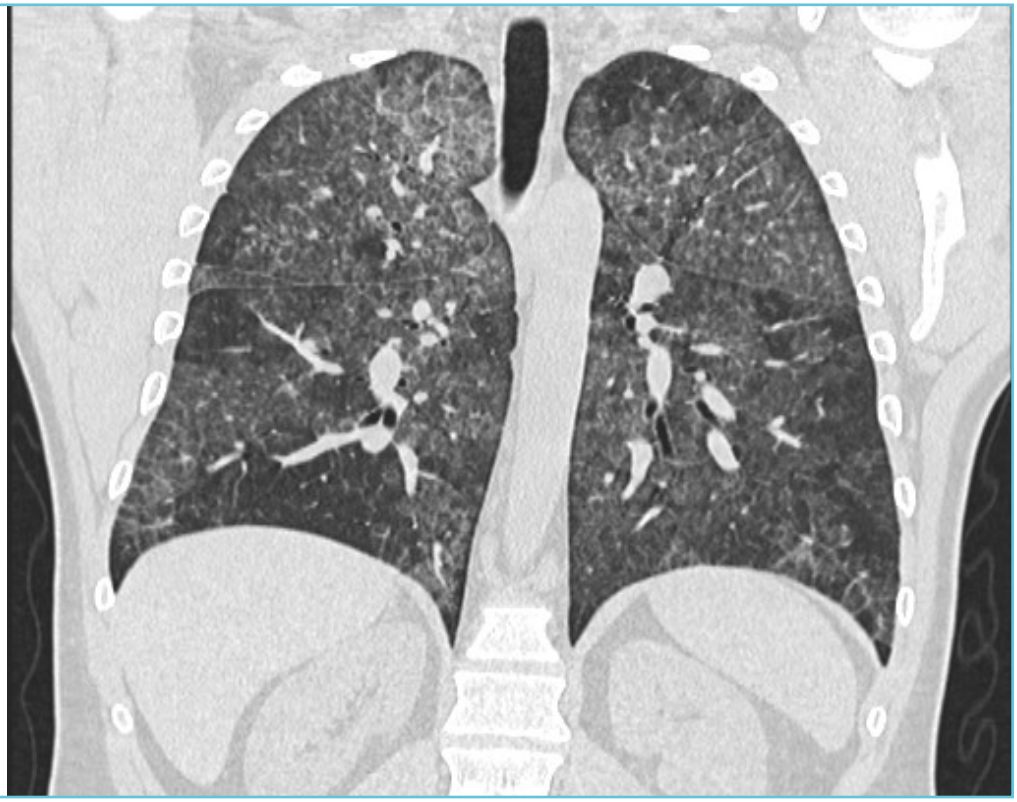


Figure 2.B



in honor of the authors of a publication of a series of 27 cases in 1958 that described pulmonary infiltrates of proteinaceous material and PAS-positive respiratory specimens (9).

This disease has an estimated incidence of 0.2-0.4 cases per million people per year and a worldwide prevalence of nearly seven cases per million people (10). It usually presents between the third and sixth decades of life.

It is a heterogeneous entity, characterized by productive or dry cough, dyspnea, fever, fatigue, and chest pain. Additionally, it can be associated with polycythemia, hypergammaglobulinemia, elevated LDH, elevated tumor markers (CEA, CA 19.9), as well as the presence of serum antibodies (anti-GM-CSF), both in serum and in BAL in the case of primary PAP (1). Its clinical course is variable and ranges from spontaneous resolution to death from infections or progressive respiratory failure (11).

The diagnosis includes clinical evaluation, respiratory function tests, HRCT, determination of anti-GM-CSF autoantibody levels, and genetic testing.

Regarding treatment, BAL is currently the gold standard, although the therapeutic approach will depend on the diagnosis and severity of the disease (12).

The detection of high levels of type IgG anti-GM-CSF antibodies by latex agglutination or ELISA techniques in peripheral blood and BAL (13) is currently accepted as a useful tool in the diagnosis of PAPi, with a sensitivity of 100% and a specificity of 98%. Levels above 5 U/mL are consistent with the diagnosis, even in asymptomatic phases. Its knowledge has also guided the development of new treatment strategies for PAPi, such as the administration of this cytokine exogenously, as an alternative or complementary therapy to BAL.

The applicability of anti-GM-CSF in monitoring and as a marker of treatment response is still under discussion. A possible correlation between anti-GM-CSF antibody titers and the extent of the disease has been sought, but the studies carried out have yielded contradictory results (14).

On the other hand, it has been seen that the inhalation of industrial dust, especially silica, titanium, and aluminum, induces the appearance of autoantibodies associated with rheumatoid arthritis, systemic lupus erythematosus, scleroderma, and glomerulonephritis. In the case of PAP, it could be considered that occupational inhalation of dust induced the appearance of anti-GM-CSF antibodies, but such an association is not clear, as there are Japanese studies of patients exposed to industrial dust who develop PAP but without the presence of anti-GM-CSF antibodies (15).

In conclusion, PAP is a rare disease that often poses diagnostic difficulties and, in many cases, requires confirmation through lung biopsy to obtain a definitive diagnosis. With the available scientific evidence, the determination of anti-GM-CSF antibodies has proven to be a useful tool in the diagnosis of PAPi, and its involvement in the long-term monitoring of anti-GM-CSF serum levels could clarify its usefulness as an early detector of recurrences and would allow for individualized values and cutoff points preceding the establishment of clinical symptoms, although studies need to be continued to corroborate this.

TAKE-HOME MESSAGES

- Importance of clinical laboratory in determining anti-GM-CSF as a confirmatory method for primary PAP.
- Different types of PAP lead to the accumulation of pulmonary surfactant and dysfunction of macrophages.

- PAP is a rare disease that can present with an asymptomatic clinical picture or progress to respiratory failure.
- Importance of using BAL as the treatment of choice in patients with PAP.

❖❖❖❖❖

Compliance with ethical standards

Conflict of interest: The authors have declared that no conflict of interest exists.

Author contributions: All authors were involved in the diagnosis and management of this case.

Funding: The authors have not declared a specific grant for this research from any funding agency or source in the public, commercial or not-for-profit sectors.

❖❖❖❖❖

REFERENCES

1. Trapnell BC, Whitsett JA, Nakata K. Pulmonary alveolar proteinosis. N Engl J Med. 2003;349(26):2527-2539. doi:10.1056/NEJMra023226

2. Carey B, Trapnell BC. The molecular basis of pulmonary alveolar proteinosis. Clin Immunol. 2010;135(2):223-235. doi:10.1016/j.clim.2010.02.017

3. Jouneau S, Kerjouan M, Briens E, et al. Pulmonary alveolar proteinosis. Rev Mal Respir. 2014;31(10):975-991. doi:10.1016/j.rmr.2014.08.009

4. McCarthy C, Avetisyan R, Carey BC, Chalk C, Trapnell BC. Prevalence and healthcare burden of pulmonary alveolar proteinosis. Orphanet J Rare Dis. 2018;13(1):129. doi:10.1186/s13023-018-0846-y

5. Ben-Dov I, Segel MJ. Autoimmune pulmonary alveolar proteinosis: clinical course and diagnostic criteria. Autoimmun Rev. 2014;13(4-5):513-517. doi:10.1016/j.autrev.2014.01.046

6. Suzuki T, Maranda B, Sakagami T, et al. Hereditary pulmonary alveolar proteinosis caused by recessive CSF2RB mutations. Eur Respir J. 2011;37(1):201-204. doi:10.1183/09031936.00090610

7. Nogee LM, Dunbar AE 3rd, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene is associated with familial interstitial lung disease. N Engl J Med. 2001;344(8):573-579. doi:10.1056/NEJM200102223440805

8. Deutsch GH, Young LR, Deterding RR, et al. Diffuse lung disease in young children: application of a novel classification scheme. Am J Respir Crit Care Med. 2007;176(11):1120-1128. doi:10.1164/rccm.200703-393OC

9. Rosen SH, Castleman B, Liebow AA. Pulmonary alveolar proteinosis. N Engl J Med. 1958 Jun 5;258(23):1123-42. doi: 10.1056/NEJM195806052582301.

10. Salvaterra E, Campo I. Pulmonary alveolar proteinosis: from classification to therapy. Breathe (Sheff). 2020;16(2):200018. doi:10.1183/20734735.0018-2020

11. Riley L, Wang T, Ataya A. Pulmonary Alveolar Proteinosis (PAP). Am J Respir Crit Care Med. 2019;200(8):P16-P17. doi:10.1164/rccm.2008P16

12. Talamoni H, Pisapia N, Cortes L, Arellano M, Wainstein E, Svetliza G, et al. Therapeutic bilateral lung lavage in an adolescent with pulmonary alveolar proteinosis. Case report. Arch Argent Pediatr. 2020;118(3):e324-28. doi: 10.5546/aap.2020.e324.

13. Uchida K, Nakata K, Carey B, et al. Standardized serum GM-CSF autoantibody testing for the routine clinical diagnosis of autoimmune pulmonary alveolar proteinosis. J Immunol Methods. 2014;402(1-2):57-70. doi 10.1016/j.jim.2013.11.011

14. Nei T, Urano S, Itoh Y, et al. Light chain (κ/λ) ratio of GM-CSF autoantibodies is associated with disease severity in autoimmune pulmonary alveolar proteinosis. Clin Immunol. 2013;149(3):357-364. doi: 10.1016/j.clim.2013.10.002

15. Borie R, Danel C, Debray MP, et al. Pulmonary alveolar proteinosis. Eur Respir Rev. 2011;20(120):98-107. doi: 10.1183/09059180.00001311

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Autoantibodies to intrinsic factor can jeopardize pernicious anemia diagnosis: a case report

Lucía Fraile, Ana Sopena, Carlos E. Chávez, Maria Font-Font, Aureli Esquerda

Servicio de Análisis Clínicos, Hospital Universitari Arnau de Vilanova, Lleida, Spain

ARTICLE INFO

Corresponding author:
Lucía Fraile García
Servicio de Análisis Clínicos
Hospital Universitari Arnau de Vilanova
Avenida Rovira Roure 80
25198. Lleida
Spain
Phone: +34 691754564
E-mail: lucia.cuellar96@gmail.com

Key words:
vitamin B12, cobalamin, pernicious anemia, interference

ABSTRACT

Vitamin B12 deficiency may cause neurological and hematological alterations. Its assessment should be easy considering that the access to its measurement is available in majority of the clinical laboratories. The presence of technical interference when measuring vitamin B12 can lead to an erroneous or a more difficult diagnosis of conditions as pernicious anemia. We report a case in which an interference in the evaluation of vitamin B12 concentration led to the realization of invasive tests and almost a misdiagnosis of a patient who actually had pernicious anemia. Professionals need to be aware of these interferences when we assess outcomes.

INTRODUCTION

Vitamin B12, or cobalamin, is an essential co-factor that takes part in methylation reactions related to DNA and cell metabolism, such as conversion of methylmalonyl-CoA (MMA) to succinyl-CoA and synthesis of methionine. It is a water-soluble vitamin, which is mostly present in foods of animal origin. Clinical signs of cobalamin deficiency are associated with neurological and hematological symptoms including paresthesia, ataxia, weakness, anemia and pancytopenia (1).

One of the causes of cobalamin deficiency is linked to its malabsorption. Once the cobalamin is ingested, it is absorbed in the ileum thanks to the intrinsic factor, which is bound to it and is secreted in the gastric juice by the parietal cells. Intrinsic factor is required for absorption of vitamin B12. There are different causes that can explain this malabsorption, like gastrectomy, surgical resections, and a variety of bacterial or inflammatory diseases affecting the small intestine. However, the most common cause is a defect in the secretion of intrinsic factor due to the presence of autoantibodies, either because of antibodies to parietal cells or to the intrinsic factor itself, resulting in an inadequate vitamin B12 absorption from foods. This condition is called pernicious anemia. The prevalence of pernicious anemia is 0.1% in the general population and 2-3% in individuals over the age of 65 (female:male ratio ~ 2:1) (2).

Vitamin B12 could be measured by different methods, based on competitive binding immunoenzymatic assays (Figure 1). In our laboratory, the determination of Vitamin B12 was done by Access Vitamin B12 assay (Beckman Coulter) (Figure 1A). It has an initial denaturation step that inactivates intrinsic factor blocking antibodies that can be present in the patient's serum. Approximately 50% of patients with pernicious anemia have these anti-intrinsic factor antibodies. The vitamin B12 present in the sample binds to the intrinsic factor-alkaline phosphatase conjugate, preventing the conjugate from binding to the solid phase anti-intrinsic factor. The chemiluminescent substrate is added to the vessel and light generated by the reaction is measured with a luminometer. The amount of light is the result of the reaction produced by the non-vitamin B12 binded conjugate bound to the solid phase, and is inversely proportional to the concentration of vitamin B12 concentration in the sample.

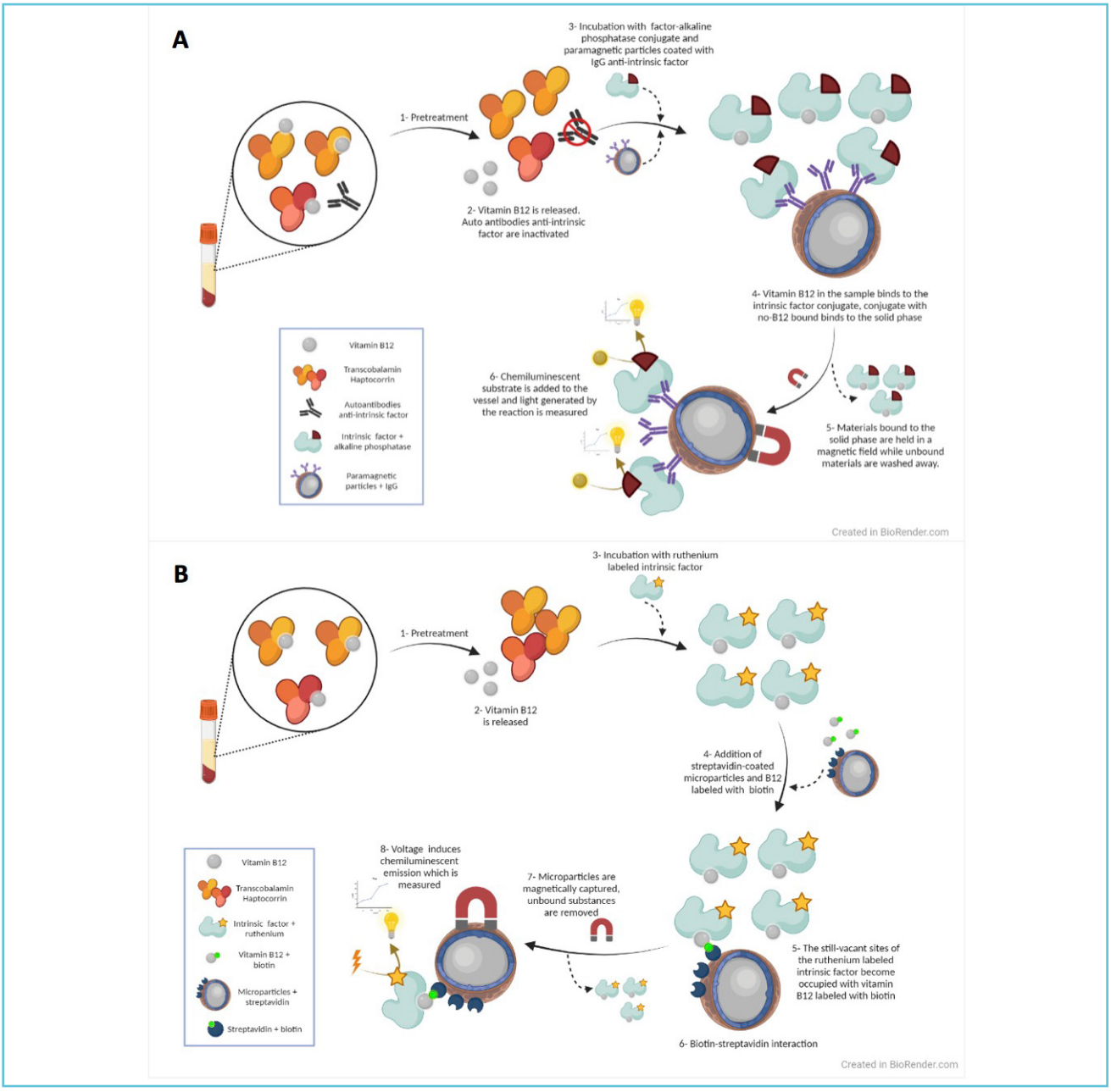
Even though the demonstration of vitamin B12 deficiency is one of the main aspect for the diagnosis, it has been proved that measuring serum vitamin B12 level alone is not sufficient to diagnose the deficiency and its sensitivity is questionable. This is caused by an analytical interference in the vitamin B12 measurement that has been reported in various publications (3–5). Therefore, the evaluation of MMA and homocysteine levels, as functional biomarkers of vitamin B12 deficiency, is recommended to evaluate the diagnosis (6).

We report a case in which the diagnosis of pernicious anemia was complicated due to an erroneous result in the measurement of vitamin B12 that made the diagnosis more difficult and led to indication of more diagnostic tests and a possible misdiagnosis.

CLINICAL - DIAGNOSTIC CASE

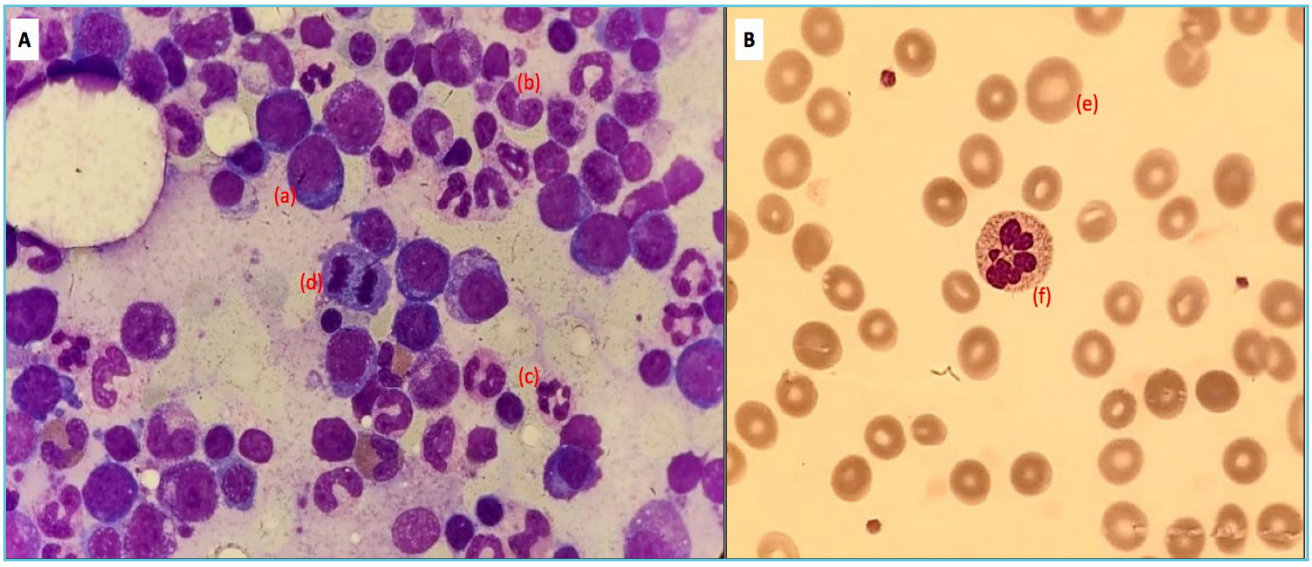
An 86-year-old woman with antecedents of arterial hypertension, type-II diabetes mellitus and paroxysmal atrial fibrillation, was referred for a hematological evaluation due to macrocytic anemia and leukopenia. The analytical results obtained were: leukocytes 3.71 x 10^9/L [reference range (RR): 4.8-10.8], red blood cells 2.6 x 10^{12}/L [RR: 4-5.4], hemoglobin 107 g/L

Figure 1 **Comparing different methods for vitamin B12 measurement:**
A) Access Vitamin B12 assay by Beckman Coulter®. It represents a chemiluminescent immunoassay based on the binding of the conjugate to solid phase anti-intrinsic factor. The light production is inversely proportional to the concentration of vitamin B12 in the sample.
B) Vitamin B12 assay by cobas Roche®. It represents an electrochemiluminescence immunoassay based on biotin-streptavidin binding. The light production is inversely proportional to the concentration of vitamin B12 in the sample.



[RR: 120-160], MCV 116 fL [RR: 80-100], MCH 41.2 pg [RR: 26-34], vitamin B12 269 pmol/L [RR: 133-675] and folic acid >53 nmol/L [RR: 7-41.5]. The most probable diagnosis would be megaloblastic anemia as a result of cobalamin or folic acid deficiency, but since no such deficiency was observed, these results suggested a low-risk myelodysplastic syndrome, and a check-up was performed after three months. These new results showed a clinical deterioration with significant asthenia: leukocytes 3.8 x 10^9/L [RR: 4.8-10.8], hemoglobin 95 g/L [RR: 120-160], MCV 128 fL [RR: 80-100], and low reticulocytes. A bone-marrow aspirate and a biopsy were performed because of a suspicion of a bone-marrow insufficiency. The bone-marrow study did not reveal any dysplastic features in any hematological series, but manifested megaloblastic features (Figure 2A): red blood cells and their progenitors were macrocytic and we could observe several mitosis and vacuoles, Howell-Jolly bodies, some binucleated cells and 0.89 % of blasts. Cytomorphology of peripheral blood (Figure 2B) also suggested a megaloblastic anemia (macroovalocytes and hypersegmented neutrophils) but with the discordance of non-decreased values of vitamin B12 or folic acid. In addition, the patient had a polyclonal increase of immunoglobulins and, in order to establish a clear diagnosis, the laboratory specialists started to consider a technical interference, initially due to this high titer of immunoglobulins. Therefore, serum immunoglobulins were precipitated by adding polyethylene glycol (PEG) in a 1:1 dilution with serum, after which vitamin B12 was measured again, showing a lower level below the reference value (Table 1). With this high suspicion of a megaloblastic anemia and a very possible interference, a measurement of autoantibodies to intrinsic factor and parietal cells, MMA, homocysteine and the evaluation of vitamin B12 was

Figure 2 A) Bone marrow aspirate with megaloblastic features: macrocytic red blood cells and its progenitors (a), giant band neutrophils (b), hypersegmented neutrophils (c) and mitosis (d). 1000x, May-Grünwald-Giemsa stain.
B) Peripheral blood suggesting a megaloblastic anemia: macroovalocytes (e) and a hypersegmented neutrophil (f). 1000x, May-Grünwald-Giemsa stain.



performed in another clinical laboratory with a different analytical method (electrochemiluminescence immunoassay "ECLIA" at Cobas e of Roche®) (Table 1, Table 2).

These new results confirmed vitamin B12 deficiency, which was not evident with our analytical method, because of the presence of antibodies to the intrinsic factor. The patient was treated with cianocobalamine (Optovite B12®) and she was reevaluated revealing clinical and analytical improvement.

DISCUSSION

The deficiency of vitamin B12 is relatively common and it may result in a variety of clinical symptoms (1). In case of a severe deficiency, it can present as bone marrow suppression resulting in anemia, neutropenia and/or thrombocytopenia. Its demonstration should be easy,

Table 1 Test results for vitamin B12 evaluation by different techniques and/or pretreatments

Equipment	Vitamin B12 result
UniCel DxI 800 Beckman Coulter ®	269 pmol/L
UniCel DxI 800 Beckman Coulter ® (1:1 dilution with PEG)	<37 pmol/L
cobas e 601 Roche ®	37 pmol/L

Table 2 Test results for the diagnosis of pernicious anemia*

	Result	Reference range
Vitamin B12	37 pmol/L	133 - 675
Homocysteine	100.84 µmol/L	4.3 - 11.1
Methyl-malonic acid (MMA)	13.7 µmol/L	0.08 - 0.56
Antibodies to intrinsic factor	>600 U/mL	0 - 7
Antibodies to parietal cells	negative	

** Vitamin B12 measured by an electrochemiluminescence immunoassay in a cobas 6 601 Roche®, homocysteine measured by immunoturbidimetry, MMA measured by mass spectrometry, antibodies to intrinsic factor measured by fluoroenzyme immunoassay and antibodies to parietal cells determined by indirect immunofluorescence. Reference ranges are provided.*

because it is a routine parameter that can be measured in the majority of clinical laboratories, but this may not always be the case.

In very rare cases, during the pretreatment and measurement of vitamin B12, certain samples may not be inactivated because of the heterogeneity or extremely high titer of the anti-intrinsic factor antibodies. Such interfering autoantibodies may give erroneous results, as we have seen in the reported case. This situation has been described in other cases or studies and associated with different analyzers (5). This fact is most probably due to an ineffective inactivation of interfering anti-intrinsic factor auto-antibodies and its binding to the conjugate, although it can also be produced by other immunoglobulins.

This interference can have an important impact for both diagnosis and treatment. We report a case in which the suspicion of a pernicious anemia that cannot be confirmed has revealed a significant interference. It causes falsely high or normal levels of vitamin B12 in a patient with an actual deficiency. The deficiency could not be proved and the diagnosis was oriented as a possible myelodysplastic syndrome, even requiring a bone marrow study to confirm the pathology. We think that it is important for laboratory and clinical specialists to be aware of this interference or technical limitation. Patients should be given a further assessment if they are suspected of having these auto-antibodies or if the results of vitamin B12 conflict with other clinical or laboratory findings. This must be taken into consideration and confirmed by additional tests prior to performing invasive procedures such as bone marrow aspiration or initiating treatment. We suggest evaluating the presence of anti-intrinsic factor antibodies, given the possibility of this technical interference. As we discussed above, it is necessary to use metabolic tests such as the measurement of homocysteine and/or MMA in patients strongly suspected to have a pernicious anemia without a low cobalamin level (6). It is said that these parameters should always accompany cobalamin concentration, or the performance of the evaluation of vitamin B12 with other techniques, like mass spectrometry, although it is not available in the majority of the laboratories yet.

TAKE HOME MESSAGES / LEARNING POINTS

- Laboratory specialists play a key role in the detection of technical interferences during the evaluation of some parameters, especially in immunoassays.
- A misdiagnosis of a pernicious anemia could raise suspicion of a hematological disease , which may lead to invasive diagnostic procedures and treatment that are unnecessary for the patient.
- The measurement of homocysteine and MMA, together with vitamin B12 concentration, is highly recommended if we have a suspicion of pernicious anemia.

❖❖❖❖❖

Conflicts of interest

None of the authors have any conflict of interest to report.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Contributorship

LF researched the literature, wrote and edited the manuscript, AS reviewed and corrected the manuscript. All authors reviewed and approved the final version of the manuscript.

REFERENCES

1. Hunt A, Harrington D, Robinson S. Vitamin B12 deficiency. BMJ. 2014 Sep 4;349:g5226.

2. Esposito G, Dottori L, Pivetta G, Ligato I, Dilaghi E, Lahner E. Pernicious Anemia: The Hematological Presentation of a Multifaceted Disorder Caused by Cobalamin Deficiency. Nutrients. 2022 Apr 17;14(8):1672.

3. Yang DT, Cook RJ. Spurious Elevations of Vitamin B12 with Pernicious Anemia. N Engl J Med. 2012 May 3; 366(18):1742–3.

4. van Rossum AP, Vlasveld LT, Castel A. Falsely elevated cobalamin concentration in multiple assays in a patient with pernicious anemia: a case study. Clin Chem Lab Med. 2013 Sep;51(9):e217-9.

5. Ralph C, Pal AY. Failures of Cobalamin Assays in Pernicious Anemia. N Engl J Med. 2012 Jul 26;367(4): 385–6.

6. İspir E, Serdar MA, Ozgurtas T, et al. Comparison of four automated serum vitamin B12 assays. Clin Chem Lab Med. 2015 Jul;53(8):1205-13.

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

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.

Produced by: Insoft Digital Web Solutions

Published by: IFCC International Federation of Clinical Chemistry and Laboratory Medicine

www.ifcc.org