Review

Huu-Hien Huynh, Amandine Bœuf, Jens Pfannkuche, Philipp Schuetz, Marc Thelen,
Gunnar Nordin, Eline van der Hagen, Patricia Kaiser, Dagmar Kesseler, Tony Badrick,
Bernard Poggi, Ulla Tiikkainen, Gareth J Davies, Anja Kessler, Mario Plebani, Joëlle Vinh
and Vincent Delatour*

Harmonization status of procalcitonin measurements: what do comparison studies and EQA schemes tell us?

Abstract: Sepsis represents a global health priority because of its high mortality and morbidity. The key to improving prognosis remains an early diagnosis to initiate appropriate antibiotic treatment. Procalcitonin (PCT) is a recognized biomarker for the early indication of bacterial infections and a valuable tool to guide and individualize antibiotic treatment. To meet the increasing demand for PCT testing, numerous PCT immunoassays have been developed and commercialized, but results have been questioned. Many comparison studies have been carried out to evaluate analytical performance and comparability of results provided by the different commercially available immunoassays for PCT, but results are conflicting. External Quality Assessment Schemes (EQAS) for PCT constitute another way to evaluate results comparability. However, when making this comparison, it must be taken into account that the variety of EQA materials consist of different matrices, the commutability of which has not yet been investigated. The present study gathers results from all published comparison studies and results from 137 EQAS surveys to describe the current state-of-the-art harmonization of PCT results. Comparison studies globally highlight a significant variability of measurement

List of collaborators: Klaus-Peter Hunfeld, Instand e.V. Reference Laboratory, Düsseldorf, Germany. Carolina Kristoffersson, External Quality Assurance in Laboratory Medicine in Sweden (Equalis), Uppsala, Sweden. Mary Annette Thomas, Weqas, Cardiff and Vale University Health Board, Cardiff, UK. Ron Meijer, Dutch Foundation for Quality Assessment in Medical Laboratories (SKML), Nijmegen, The Netherlands. Peter Graham, Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP), Sydney, Australia. Pierre-Alain Morandi, Quality Control Center Switzerland (CSCQ), Chêne-Bourg, Switzerland. Jean-Christophe Eynard, ProBioQual, Lyon, France. Laurence Chardon, ProBioQual, Lyon, France. The IFCC WG-PCT members.

*Corresponding author: Vincent Delatour, Department of Biomedical and Organic Chemistry, Laboratoire National de Métrologie et d'Essais (LNE), 1 rue Gaston Boissier, 75015, Paris, France, Phone: +(33) 1 40 43 40 75, E-mail: vincent.delatour@lne.fr
Huu-Hien Huynh, Laboratoire National de Métrologie et d'Essais, Paris, France; and Biological Mass Spectrometry and Proteomics, SMBP, ESPCI Paris, Université PSL, CNRS, Paris, France. https://orcid.org/0000-0002-7814-3189
Amandine Bœuf, Laboratoire National de Métrologie et d'Essais, Paris, France
Jens Pfannkuche, PANACECA UG, Berlin, Germany

Philipp Schuetz, University Department of Medicine, Kantonsspital Aarau, Switzerland; and Faculty of medicine, University of Basel, Basel, Switzerland
Marc Thelen, Dutch Foundation for Quality Assessment in Medical Laboratories (SKML), Nijmegen, The Netherlands
Gunnar Nordin, External Quality Assurance in Laboratory Medicine in Sweden (Equalis), Uppsala, Sweden
Eline van der Hagen, MCA laboratory, Streekziekenhuis Koningin Beatrix, Winterswijk, The Netherlands
Patricia Kaiser, Instand e.V. Reference Laboratory, Düsseldorf, Germany
Dagmar Kesseler, Quality Control Center Switzerland (CSCQ), Chêne-Bourg, Switzerland
Tony Badrick, Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP), Sydney, Australia
Bernard Poggi, ProBioQual, Lyon, France; and Laboratoire de Biochimie des Hospices Civiles de Lyon, Lyon, France
Ulla Tiikkainen, Labquality Oy, Helsinki, Finland
Gareth J Davies, Weqas, Cardiff and Vale University Health Board, Cardiff, Wales, UK
Anja Kessler, Reference Institute for Bioanalytics (RfB), Bonn, Germany
Mario Plebani, Department of Laboratory Medicine, University Hospital of Padova, Verona, Italy. https://orcid.org/0000-0002-0270-1711
Joëlle Vinh, Biological Mass Spectrometry and Proteomics, SMBP, ESPCI Paris, Université PSL, CNRS, Paris, France

results that nonetheless seem to have a moderate impact on medical decision-making. For their part, EQAS for PCT provides highly discrepant estimates of the interlaboratory CV. Due to differences in commutability of the EQA materials, the results from different peer groups could not be compared. To improve the informative value of the EQA data, the existing limitations such as non-harmonized conditions and suboptimal and/or unknown commutability of the EQA materials have to be overcome. The study highlights the need for commutable reference materials that could be used to properly evaluate result comparability and possibly standardize calibration, if necessary. Such an initiative would further improve the safe use of PCT in clinical routine.

Keywords: equivalence; external-quality-assessment; harmonization; method comparison; procalcitonin; sepsis.

Introduction

Sepsis is recognized as a global health priority, with 48.9 million reported cases in 2017 and over 11 million deaths recorded worldwide every year [1, 2]. Patient mortality is highly dependent on the delay of antimicrobial therapy administration [3, 4], early and accurate sepsis diagnosis followed by appropriate antimicrobial therapy is an essential determinant for the patient outcome [5–7] and reducing costs [8, 9]. Procalcitonin (PCT) emerged as a valuable biomarker to differentiate between infectious and non-infectious causes of systemic inflammation. In non-infectious conditions, PCT concentration in plasma or serum typically remains below 0.05 µg/L [10]. In bacterial sepsis, however, PCT concentration can increase more than 1,000 fold from the basal concentration [11, 12], whereas it remains low in the case of localized bacterial infections and viral infections. Numerous studies have demonstrated the good performance of PCT as a diagnostic marker for sepsis [13–20], and the prediction of associated mortality [21–23].

Moreover, its implication in the diagnosis of neonatal sepsis [24] and bacterial lower respiratory tract infections (LRTI) [25] has been confirmed. Studies have also proved the utility of PCT for antibiotic management and stewardship in sepsis and LRTI [26, 27], notably in decreasing the antibiotic treatment duration in ICUs [28–30]. Therefore, PCT measurement is an efficient tool to reduce the abuse of broad-spectrum antibiotics, antibiotic-related side effects, and the occurrence of antibiotic resistance [25].

To meet the high demand for PCT testing, a wide range of immunoassays has been developed employing different calibration, different detection principles and different antibodies [31]. As PCT-based diagnosis and PCT-guided antibiotic stewardship depend on specific PCT cutoffs (i.e., 0.25 µg/L for non-ICU studies and 0.5 µg/L for ICU studies as stopping rules for an antibiotic) [32], high comparability among assays is crucial. A lack of comparability may impact patients' continuity of medical care [33] and hinder the ability to aggregate and compare data from different clinical trials and epidemiological studies. Consequently, determining whether the clinical thresholds established from one trial can be applied to other trials is a central problem. Therefore, it is essential to develop analytical and clinical equivalence between results provided by the different assays. As assay design, particularly the choice and the selectivity of antibodies, can hardly be harmonized, standardization of calibration appears the best and virtually the only way to obtain comparable and accurate PCT results. In the absence of any higher-order reference method and internationally agreed to calibration material for PCT, the Brahms PCT LIA assay has been used as a surrogate reference method for Thermo Fischer Scientific and its license partners PCT assay development, hereafter referred as to the Brahms PCT assays [34, 35], and was the predicate device in the first regulatory submissions to FDA for KRYPTOR and Vidas Brahms assays. The Brahms Kryptor and Vidas Brahms assays are now used as a predicate device to evaluate the method comparison of other Brahms automated immunoassays in subsequent regulatory submissions. However, other automated PCT assays have been developed by various diagnostic companies (e.g., Maglumi, Diazyme, DiaSys, Beckman Coulter) for which there is currently no harmonization system in place. Alongside fully automated assays, a wide range of point-of-care PCT tests have also been developed and are increasingly used in critical care settings. Still, here again, there is no internationally agreed reference system of higher-order in place. Several studies have highlighted the need for harmonization of PCT results [36, 37]. However, important discrepancies can be observed across conclusions of studies conducted to evaluate the analytical performance and the comparability of results provided by the different immunoassays for PCT. In this paper, we first compiled all available comparison studies between commercially available fully-automated assays for PCT measurement. We then analyzed results from 137 external quality assessment (EQA) surveys to document the current state of the art with regards to the harmonization status of PCT measurements.

Comparison studies

We collected all available FDA 510(k) approval documents and published research papers describing the comparison of fully automated assays for PCT. Publications in languages other than English or French were included if they had a meaningful detailed abstract in English or French. Thirty-three records from the literature search were identified: 8 FDA 510(k) approval documents [38–45] and 25 research papers [36, 37, 46–68]. Details of each study are summarized in Supplementary Material, Table 1. The comparison studies reported in FDA 510(k) approval documents were performed according to CLSI EP09-A3 "Measurement Procedure Comparison and Bias Estimation Using Patient Samples". Most of the research papers did not refer to any specific guidelines for method comparison and bias evaluation. All the studies performed at least one regression analysis between results from two immunoassays. However, they didn't all estimate and discuss the bias of measurement between the PCT immunoassays apart from determining the regression analysis and the agreement of classification at clinical concentrations. As these criteria are crucial to depict the current state of the art of the correlation and agreement of classification between PCT assay methods, they were evaluated and discussed below.

Evaluation of the regression analysis between two assays

In all studies included in this evaluation, a regression plot analysis was established to evaluate the correlation between two immunoassays — most of the published studies employed Passing-Bablok or Deming regression analysis (see Supplementary Material, Table 1). The majority of published studies reported a good correlation between the different PCT assays according to the correlation coefficient above 0.85 (see Supplementary Material, Table 1). However, the degree of correlation between the two assays depends on the number of samples and the concentration range, as the correlation coefficient will tend to be higher if the concentration range is broad [69]. The correlation coefficient informs about the degree of dispersion of measurement results but not on the agreement of results. Therefore, it is preferable to use bias as an estimator of agreement. The bias can be subdivided into proportional and constant bias, and these two components can respectively be evaluated via the slope and the intercept of the regression analysis [70]. Like FDA 510(k) method comparison, the Brahms Kryptor and Vidas Brahms assays

were predominantly employed as the comparative assay across the different studies (see Supplementary Material, Table 1). The slope and the intercept of regression equations are presented in Figure 1. The FDA 510(k) approval documents show both more negligible proportional and smaller constant bias (Figure 1(A)) than in the published research papers (Figure 1(B) and (C)). In data from the FDA approval documents, the deviation from a slope of 1 was less than ±6%, whereas it rises to ±69% in some research papers. This observation could be explained by the fact that the experimental conditions strictly followed the technical protocol of FDA 510(k) approval documents. In other studies, different technical conditions (i.e., reagents, instruments, interfering sample) could be employed. The proportional bias between the Brahms Kryptor and other assays was generally lower than ±25% (Figure 1(B)). The constant positive bias of the Vidas Brahms compared to the Brahms Kryptor (>25%) was also confirmed by the manufacturer's data, in which a proportional error of up to 22% was reported [61].

Compared with the Vidas Brahms, a negative proportional bias was observed for almost all assays, with a tendency to exceed −20% (Figure 1(C)). Additionally, the evaluation of regression analysis studies employing assays other than Kryptor Brahms or Vidas Brahms as a comparative method is presented in Supplementary Material, Table 1. A proportional bias of up to 79% was reported when comparing the Liaison Brahms assay to the Elecsys Brahms (on Cobas) or Architect Brahms assay [56]. Overall, the data presented revealed a noticeable proportional bias between some assays but observations were variable between studies. The slope is a valid estimate of the discrepancies between measurement results, provided that the constant bias is negligible. A non-negligible intercept indicates that measurement results have a constant bias. In 43 of the 60 comparison studies, the intercept ranged between −0.2 and 0.2. While this constant bias may have a negligible impact on medical decision making at high concentrations of PCT, it could represent significant differences in measurement results at low clinical cutoffs (i.e. 0.1, 0.25 and 0.5 μg/L). Using the slope or intercept alone is not enough to describe the agreement between assays. We analyzed the combined influence of constant and proportional bias with a focus on the low PCT concentration. For this evaluation, the PCT concentration was set at clinically relevant concentrations (0.1, 0.25, 0.5, 1, 2, 10 μg/L) for the comparative assay and computed for the test assays using the 60 regression equations retrieved from comparison studies. The differences between the comparative assay and the test assay were then evaluated. For clinical cutoffs above 0.25 μg/L, the difference between
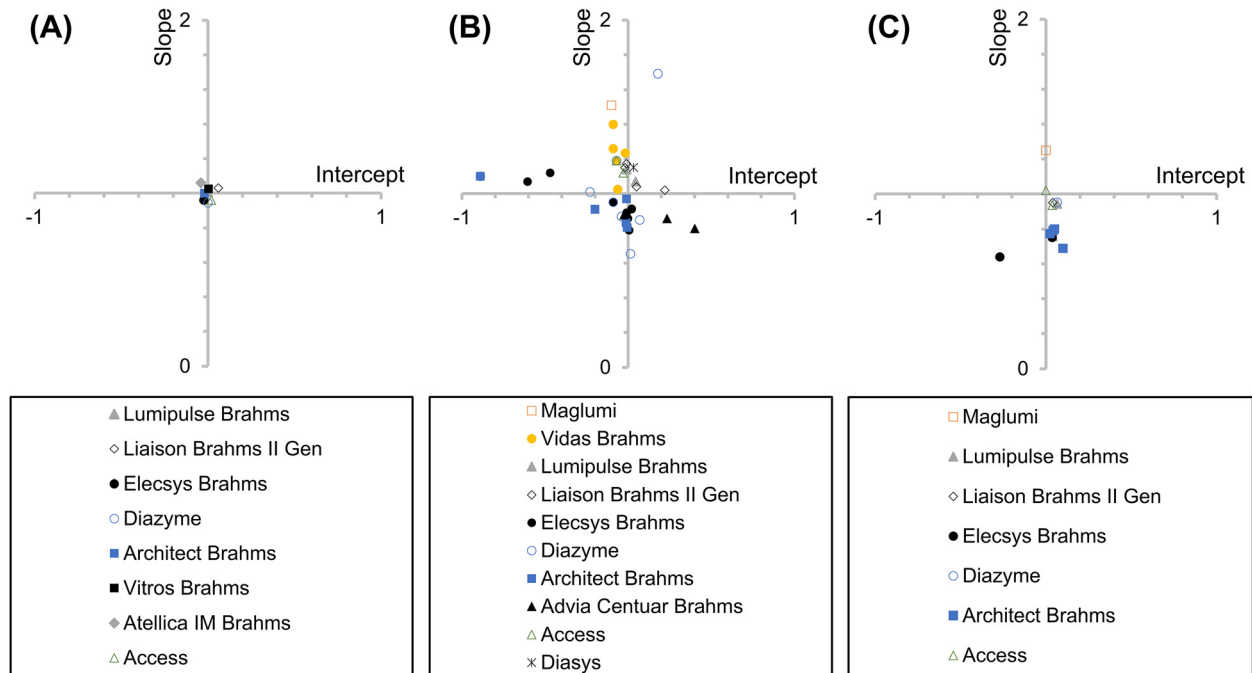
Figure 1: Slopes and intercepts of regression analyses between two PCT assays from FDA 510(k) approval documents and research documents.
(A) Brahms Kryptor (solid form) and Vidas Brahms (clear form) assay as comparative assay from FDA 510(k) approval documents. (B) Brahms Kryptor assay as comparative assay from research documents. (C) Vidas Brahms assay as comparative assay from research documents. Slope (*a*) and intercept (*b*) were obtained from the regression equation $y = a \times x + b$ established between the results measured on respective test assay as variable y and the results measured on comparative assay as variable *x*.

assays was less than 15% for 52–57% of evaluations, between 15 and 30% for 25–32% of evaluations and higher than 30% for 8–22% of assessments (see Supplementary Material, Figure 1). The discrepancy between assays is more accentuated at 0.1 µg/L: a difference between assays larger than 15% was observed for 80% of cases. However, such an evaluation of the discrepancy between assays based on the regression analysis may also be impacted by the number of samples and the concentration range used to plot the regression curve. To assess if this computing evaluation of discrepancy between assays is representative of the actual situation and is appropriate to accurately estimate assay bias at the relevant cutoffs, these results need to be completed by evaluating the experimental bias between the assays.

The bias of measurement results according to Bland-Altman plot analysis

Across selected studies, the bias of measurement results was generally evaluated using Bland - Altman analysis. In this section, we then limited our investigation to published studies based on this approach (Table 1). The different

studies reported mean biases, expressed as relative (percentage) or absolute values (concentration), between two immunoassays over a specified concentration range. However, the documented biases may vary depending on the concentration range, which differs from one studies to another one. Indeed, Dipalo et al. reported three different mean biases (2.7, 11.6, 0.6 µg/L) between the Vidas Brahms and Brahms Kryptor at concentration range 0.1–58.7, 10.0–58.7, and 0.1–10.0 µg/L, respectively [36]. This raises the difficulty to compare all studies together. A comparison of bias through different comparison studies over time for a given immunoassay should be carefully interpreted because study designs could substantially differ. Overall, one can observe that bias reported were consistently higher when comparing the Lumipulse Brahms, Vidas Brahms, Liaison Brahms, Access, Maglumi assays to the Brahms Kryptor assay. As observed for the slope, a positive bias for the Vidas Brahms assay, higher than other assays, was reported [36, 37, 62–68]. However, this bias seems to be variable across study and time, as observed in the study from Lippi et al., where the mean bias between the Vidas Brahms and Brahms Kryptor assays was only 0.2%. In addition to a direct impact on patient results, a substantial bias (positive or negative) for a given immunoassay could

be propagated and perpetuated over time if new PCT assays use this immunoassay (e.g. Vidas Brahms) as a predicate device for the approval process (FDA or CE marked). In a comparison study performed between the Liaison Brahms assay (FDA approved assays using Vidas Brahms assay) and two FDA approved assays using the Brahms Kryptor as predicate device (Elecsys Brahms and Architect Brahms assay), Eidizadeh et al. reported a positive mean bias of

Table 1: Reported mean bias between assays.

Date	Study	Test assay	Comparative assay	Mean bias, µg/L or %	Concentration range, µg/L	
					Min	Max
2017	FDA K172713 [40]	Lumipulse Brahms	Brahms Kryptor	0.185 µg/L	0.02	50
2010	Schuetz et al. [64]	Vidas Brahms	Brahms Kryptor	17.10%	0.05*	100*
2010	Hausfater et al. [62]	Vidas Brahms	Brahms Kryptor	0.108 µg/L	0.02	71.96
				0.026 µg/L	0.05	4
2011	Sanders et al. [49]	Advia Brahms	Brahms Kryptor	−0.567 µg/L	0.03	45.7
				−0.05 µg/L	0	2
2012	Lloyd et Kuyl [50]	Advia Brahms	Brahms Kryptor	0.75 µg/L	0.04	46.07
		Elecsys Brahms	Brahms Kryptor	0.08 µg/L	0.04	46.07
2014	Dipalo et al. [51]	Diazyme (Beckman)	Brahms Kryptor	0.48 µg/L	0.16	111
2015	Dipalo et al. [36]	Liaison Brahms	Brahms Kryptor	0.38 µg/L	0.1	58.7
		Vidas Brahms	Brahms Kryptor	2.7 µg/L	0.1	58.7
				11.6 µg/L	10	58.7
				0.6 µg/L	0.1	10
		Elecsys Brahms	Brahms Kryptor	0.4 µg/L	0.1	58.7
		Advia Brahms	Brahms Kryptor	−1.02 µg/L	0.1	58.7
		Diazyme (Architect)	Brahms Kryptor	−0.51 µg/L	0.1	58.7
		Diazyme (Advia)	Brahms Kryptor	−0.13 µg/L	0.1	58.7
		Diazyme (Cobas)	Brahms Kryptor	−1.01 µg/L	0.1	58.7
		Diazyme (AU5800)	Brahms Kryptor	−0.87 µg/L	0.1	58.7
2016	Du puy et al. [52]	Lumipulse Brahms	Brahms Kryptor	0.42 µg/L	0.12	50
				0.02 µg/L	0.12	1
2016	Fortunato [53]	Liaison Brahms	Brahms Kryptor	33.3%	0.02*	100*
2016	Ruzzenente et al. [37]	Lumipulse Brahms	Vidas Brahms	−3.03 µg/L	0.006	75.5
				−0.49 µg/L	0	10
2017	Leung et al. [63]	Architect Brahms	Vidas Brahms	−0.5 µg/L	0.05	24
			Elecsys Brahms	−0.1 µg/L	0.05	24
2018	Mouatani et al. [57]	Architect Brahms	Brahms Kryptor	0.572 µg/L	1	25
				0.075 µg/L	0.03	1
2018	Soh et al. [65]	Architect Brahms	Elecsys Brahms	−0.8 µg/L	0.02	100
			Liaison Brahms	−2.3 µg/L	0.02	100
			Vidas Brahms	−2.5 µg/L	0.02	100
			Brahms Kryptor	−0.6 µg/L	0.02	100
2019	Wang et al. [66]	Architect Brahms	Vidas Brahms	−0.097 µg/L	0.1	2
2019	Lippi et al. [67]	Access (Access)	Brahms Kryptor	3.8%	1*	100*
		Access (DXI)	Brahms Kryptor	2.4%	1*	100*
		Lumipulse Brahms	Brahms Kryptor	13.6%	1*	100*
		Diazyme (Cobas)	Brahms Kryptor	24.9%	1*	100*
		Elecsys Brahms	Brahms Kryptor	−14.9%	1*	100*
		Vidas Brahms	Brahms Kryptor	0.2%	1*	100*
		Liaison Brahms	Brahms Kryptor	7.2%	1*	100*
		Maglumi	Brahms Kryptor	23.7%	1*	100*
		Architect Brahms	Brahms Kryptor	−7.2%	1*	100*
2019	Katz et al. [68]	Elecsys Brahms	Vidas Brahms	−0.13 µg/L	0	2
2020	Lippi et al. [55]	Access	Brahms Kryptor	3.20%	0.02	84
2020	Dupuy et al. [61]	Diasys (Cobas)	Brahms Kryptor	1.37 µg/L	0.02	50
		Diazyme (Cobas)	Brahms Kryptor	4.69 µg/L	0.02	50

*Values reported from Bland and Altman plot.

40% between the Liaison Brahms assay and both the Elecsys Brahms and Architect Brahms assays. In contrast, the Elecsys Brahms assay did not exhibit any significant bias against the Architect Brahms assay [56]. Other studies also showed the same positive bias trend for the Liaison Brahms assay [36, 53, 65, 67].

A study of Lippi et al. based on PCT results of the same set of 176 frozen plasma samples measured on 10 PCT assays depict the best recent overview on PCT assay comparison. A bias of at least 10% was reported in 28 of 45 pairwise comparisons. Among them, the Diazyme and Maglumi assays reported the highest bias compared to other assays (up to 24.9% compared to the Brahms Kryptor (Table 1) and up to 38.6% compared to the Elecsys Brahms assay [67]. Moreover, a bias of more than ±20% was unexpectedly observed between assays belonging to the Brahms PCT group (i.e., –28.2% between Elecsys Brahms (on Cobas) and Lumipulse Brahms, –20.8% between Architect Brahms and Lumipulse Brahms, 22% between Liaison Brahms and Elecsys Brahms (on Cobas)).

Overall, most of the studies reported that substantial bias between assays exists and thus shows that harmonization between all assays is not optimal. However, it is difficult to appreciate a bias reported for a whole concentration range. The real deviation of results obtained between assays at a single point concentration cannot be retrieved. Therefore, bias between assays does not bring any information on clinical interpretation of the results. Most of the studies in which a significant bias was reported also analyzed the clinical concordance at clinical cutoff concentrations. To estimate the impact of bias on clinical decisions, we then analyzed results from studies having evaluated the agreement of classification.

Agreement of classification at clinical cutoff concentrations

The correct classification of patients is essential to ensure proper patient management all along the care pathway, particularly when changing from one immunoassay to another using the same clinical cutoffs. Therefore, agreement at different clinical cutoff concentrations was retrieved from published studies to evaluate the influence of results variability on data interpretation (see Supplementary Material, Table 2). This resulted in 327 reported agreement values. Among these, 295 (90%) were equal to or better than 90%. An agreement below 90% was mainly reported at low concentrations (0.25 and 0.5 µg/L), especially when the Diazyme assay was compared to other PCT assays. Although few authors conclude that differences in

measurement results have a moderate impact on medical decision-making [36, 67], several authors still advised that longitudinal patient monitoring should preferably be carried out using the same immunoassay [36, 67]. Added to the fact that some studies outlined a poor agreement between some assays with up to 38% of misclassification, some doubts seem to persist regarding the equivalence of results and the consistency of medical decisions when different assays are used [49, 54, 58, 61, 67, 68]. A reason for this could be that in agreement studies, patients classification is performed using all study samples, which span a high range of PCT concentration (Table 1). As a better agreement is expected to be obtained when target concentrations are far from the clinical cutoff, the agreement between assays at a specific cutoff (i.e. 0.5, 2.0, 10 µg/L) will largely depend on the distribution of samples alongside the PCT concentration range. To our knowledge, no study focused on this point, and such data couldn't be retrieved from the comparison studies. Therefore, the validity of conclusions regarding the impact on medical decision-making can be disputed. To properly evaluate the effect of differences between assays on decision making, the agreement of classification should be performed using a sufficient number of samples with concentrations close to each clinical cutoffs.

Despite differences in experimental designs, comparison studies globally highlight essential differences between assays that are reported to have a moderate impact on medical decision-making. However, comparison studies generally suffer from various limitations that put these conclusions into question. Additional data are needed to evaluate comparability PCT results at clinical decision limits and interpret differences to medically relevant performance specifications.

External quality assessment schemes

Data from 143 surveys carried out from 2014 to 2020 were supplied by 10 external quality assessment scheme (EQAS) providers from eight different countries: ANSM in France, CSCQ in Switzerland, Equalis in Sweden, Instand in Germany, Labquality in Finland, ProBioQual in France, RCPAQAP in Australia, RfB in Germany, SKML in The Netherlands and Weqas in the UK. Data from surveys with less than 10 participants and materials in which PCT concentration was lower than 0.05 µg/L were excluded from this study to enable sufficiently robust data analysis. Overall, the present study includes 137 surveys with

approximately 2,220 routine laboratory participants (see Supplementary Material, Tables 3 and 4). About 27,559 PCT measurements were performed on samples with concentrations ranging from 0.05 to 43.66 μg/L (median 2.07 μg/L). Different types of serum-based materials were used within the various EQA surveys: fresh pools without spiking with exogenous PCT for Instand, frozen without spiking with exogenous PCT for Equalis, frozen with spiking with exogenous PCT for Weqas, lyophilized without spiking with exogenous PCT for RCPAQAP, and SKML, lyophilized with spiking with exogenous PCT for ANSM, CSCQ, Labquality, ProBioQual, and RfB.

As an estimator of the harmonization level of PCT measurements, we calculated the interlaboratory coefficient of variation (CV) of results provided by all participants of a given survey. Figure 2 shows the interlaboratory CV of all 137 surveys as a function of the consensus mean concentration of PCT. All surveys together, the median interlaboratory CV was 16.1%. The interlaboratory CVs are highly heterogeneous across the different surveys. In
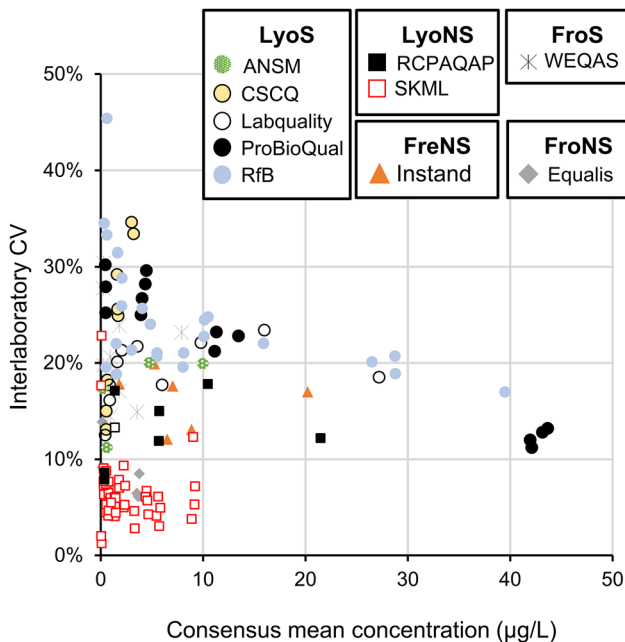
Figure 2: Interlaboratory CV obtained from the 137 EQAS on PCT carried out from 2014 to 2020 by 10 EQA providers from 7 European countries and Australia.
Consensus mean concentration represents the mean of results obtained from all immunoassay methods included in each survey. Five types of serum-based EQA materials were studied: fresh pools of samples without spiking with exogenous PCT (FreNS), frozen without spiking with exogenous PCT (FroNS), frozen with spiking with exogenous PCT (FroS), lyophilized without spiking with exogenous PCT (LyoNS) and lyophilized with spiking with exogenous PCT (LyoS).

several surveys, the interlaboratory CV was in the 5–10% range, which suggests a much better harmonization level than in other surveys in which the interlaboratory CV exceeds 25%. However, these numbers are difficult to interpret in the absence of medically relevant analytical performance specifications (APS) defining what interlaboratory CV can be deemed acceptable.

In current EQAS, the validity of individual laboratory results is most often judged using peer group consensus means as target values with predefined acceptance limits. Many EQA providers use acceptance limits of 15%, practices remain heterogeneous, and acceptance limits range between 12 and 30%. Although these numbers indicate how much results of an individual laboratory can deviate from the target, they do not provide any information on what variability across the different peer groups can be considered acceptable.

The possible origins of the discrepancies between the results of different EQAS and the identity of the main parameters contributing to the interlaboratory CV were studied as EQA programs relyed on materials with varying concentrations of PCT and various types of matrices (fresh, frozen, lyophilized, spiked with recombinant PCT or not), we investigated the effect of these parameters on the observed variability of the PCT measurement results.

Influence of the target concentrations

Two different ranges of PCT concentrations were considered: below and above 10 μg/L. Figure 2, one can notice that the dispersion of the interlaboratory CV estimates was greater for PCT concentrations lower than 10 μg/L. However, the median interlaboratory CVs were in the same range in these two concentration ranges (13.9% below 10 μg/L vs. 19.5% above 10 μg/L. This could indicate a potential impact of the target concentration on the variability of PCT measurement results. To document the interlaboratory CV close to the clinical cutoffs, we further split our analysis into narrower concentration ranges. Still, results were consistent with each other (data not shown), and the number of surveys was sometimes low in some split ranges. Therefore, the whole range was considered afterward.

Moreover, only some matrices were used at high PCT concentration (fresh non-spiked, lyophilized with or without spiking) with relatively good consistency between values of the interlaboratory CV. In contrast, all EQA materials were covered at low concentrations, and values of the interlaboratory CV were more scattered. This

confirms the need to consider the material matrix in further evaluation.

Influence of within peer group imprecision

The Non-commutability of EQA materials could compromise the ability to compare, and aggregate results from different peer groups [71], but results of all laboratories from the same peer group will all be affected the same way by matrix effects. Thus, we first evaluated results imprecision within each peer group. For each type of EQA material (fresh non-spiked, frozen or lyophilized with and without spiking) and each survey, we calculated the intra-peer group imprecision, that corresponds to the interlaboratory CV of results obtained from participants using the same analytical platform. In order to get a sufficiently robust estimate of intra-peer group imprecision, this analysis was limited to surveys including more than 20 participants per peer group. Only the following peer groups fulfilled this criterion: Elecsys Brahms (on Cobas or Elecsys), Vidas Brahms, Brahms Kryptor, Architect Brahms, Advia/Atellica Brahms, Radiometer, and Diazyme. Within these peer-groups, only fresh non-spiked, frozen non-spiked and lyophilized spiked matrices were evaluated. As presented in Supplementary Material, Figure 2, and detailed in Supplementary Material, Table 5–12 a twice higher within-peer group imprecision was observed with fresh non-spiked materials (median CV between 9.0 and 11.6%) than for frozen non-spiked and lyophilized spiked materials (median CV between 3.9 and 7.7%), except for Diazyme assay. This suggests that, for fresh non-spiked materials, within peer group imprecision largely contributed to the overall interlaboratory CV (median CV of 17.6%). On the contrary, within peer-group imprecision contributed less than between the peer-groups imprecision to the overall between-lab CV frozen non-spiked and lyophilized spiked materials. However, differences between the different peer groups could be due to an actual bias and/or matrix effects caused by non-commutability of EQA materials without discriminating these two causes. EQA organizations generally try to provide commutable material when possible, but it should be noted that this is not always a necessity depending on the purpose of EQA. In some countries, the main objective remains to compare results from labs using the same analytical platform. This is especially the case when a large number of POCT are used by labs and physicians. In this case, identifying EQA materials that are commutable for all assays can be very challenging and EQA providers may have no other choice but using materials of lower or unknown commutability. EQA schemes still provide useful information to compare results from one individual laboratory with its peers and, therefore, verify the correct implementation of assays. However, comparing results from different peer groups is limited by the use of materials of unknown commutability. This highlights the need to identify suitable matrices mimicking the most patient samples.

In search of commutable EQA materials for PCT

In an attempt to get information on the commutability of EQA materials without conducting a formal commutability study, the peer group consensus target of results provided by the most popular immunoassays was compared with the all-laboratory trimmed mean (ALTM) (see Supplementary Material, Figure 3). However, this approach showed different limitations. 5 of 10 involved EQA providers had a sufficiently large number of participants to generate robust data (>20 participants per peer group). To evaluate results harmonization, the commutability of the materials should be evaluated for all major assays. Thus, only two to six peer groups could be included per EQA survey. Using EQAS in which peer groups were large enough, two types of matrices could be represented, namely fresh non-spiked and lyophilized spiked materials, respectively. However, it was not possible to evaluate the impact of the processes of lyophilization and spiking on materials' commutability separately. Furthermore, as the market share of PCT assays dramatically varies across the different countries, the deviation of a given peer group against the ALTM will depend on the relative weight of this peer group.

In the end, it was only possible to get insights into the batch-to-batch variations in the EQA materials from a given EQA provider, but not to compare various types of EQA materials from different EQA providers. Despite differences in sample matrix and relative weight of the different peer groups, results still show that some assays consistently provide results below the ALTM and others above the ALTM. Although this could reveal actual differences between assays, no quantitative information can be obtained and interpreted in clinical impact.

Given the impossibility to discriminate between analytical bias from non-commutability bias, we analyzed results to get insights into materials characteristics that could cause differences in commutability. Spiked materials were associated with the highest median interlaboratory CV, either with frozen materials (20.2%) or lyophilized materials (21.5%) (Figure 3). In comparison, the median interlaboratory CV was only 7.1% in surveys relying on
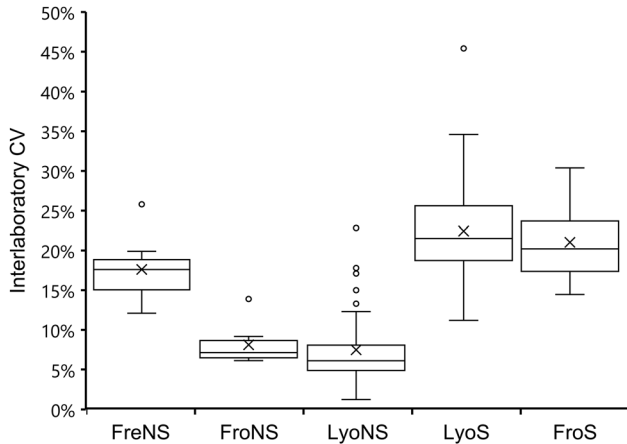
Figure 3: Box plot comparing interlaboratory CV among EQA material matrix.
Five types of serum-based EQA materials were studied: fresh pools of samples without spiking with exogenous PCT (FreNS), frozen without spiking with exogenous PCT (FroNS), frozen with spiking with exogenous PCT (FroS), lyophilized without spiking with exogenous PCT (LyoNS) and lyophilized with spiking with exogenous PCT (LyoS).

frozen non-spiked materials and 6.1% in surveys using lyophilized non-spiked materials (Figure 3). Assuming that materials non-commutability most often results in increasing differences between results assays and more significant interlaboratory CVs (79), these results suggest that spiking with recombinant PCT could represent a significant factor compromising materials commutability. The median interlaboratory CV was substantially higher in surveys using fresh non-spiked materials (17.6%) than the other non-spiked materials (7.1% for frozen and 6.1% for lyophilized).

As fresh non-spiked materials can be considered the closest matrix to fresh single donations (that are commutable by definition), fresh non-spiked materials were expected to be associated with the highest commutability level the lowest interlaboratory CV. Since the stability and the homogeneity of the fresh non-spiked materials was evaluated and demonstrated according to ISO 17043 requirements. Further studies should investigate the influence of pooling towards compromising commutability of fresh non-spiked EQA materials. Overall, the organization of studies relying on fresh materials is challenging, both from an ethical and technical point of view. It remains challenging to obtain such samples in sufficient amounts, circulate these to all EQA participants and get materials tested within the time frame of the stability approved.

Due to providing an optimized material for PCT EQAS in future, it appears desirable to identify suitable manufacturing processes leading to high commutability. Such a study is under preparation within IFCC WG-PCT with involvement from all key stakeholders. First, EQA providers will be invited to share EQA materials to cover the different types of matrices. Then, clinical partners will recruit a panel of clinical specimens that are as close as possible to patient samples (no spiking, no pooling, no lyophilization, etc.…) and cover the entire PCT concentration range. Through the involvement of all assay manufacturers, EQA materials and clinical specimens will then be measured in repeatability conditions by all available immunoassays. Statistical analysis will be performed using the difference in bias approach, and commutability acceptance criteria will be defined according to the recommendation of IFCC WG-CMT. A significant difficulty consists of sourcing many single donations in sufficient amount so that the same samples can be measured with all available immunoassays. After this work has been completed, reliable information on commutability of EQA materials will be open and make it possible to organize EQA schemes to document the true state of the art in harmonization of PCT assays. EQA can also help monitoring the clinical equivalence of assays in condition that i) clinically relevant concentrations are targeted (i.e. Close to the medical decisions limits) and ii) commutability of the EQA materials is demonstrated for all the assays which clinical equivalence is to be assessed.

Conclusions

Both published comparison studies and EQA results indicate substantial differences between results provided by the different commercially available assays for PCT. However, both strategies have limitations that make it difficult to evaluate the true current state-of-the-art in harmonization properly. In comparison studies and EQA programs, a clear identification of the involved assays is made difficult by the co-existence of multiple combinations of analyzers, reagent lots and calibrators lots. A more comprehensive description of assay characteristics appears desirable in both cases. In comparison studies, important differences in methodology were observed, such as the number of samples, material matrix, PCT concentration range and description of assays. To properly evaluate the impact of differences between assays on medical decisions making, it would be necessary to involve a sufficiently large number of samples with low PCT concentrations close to the clinically relevant cutoffs, but this information was usually not available.

EQA schemes provide valuable information regarding results consistency within a given peer group. However,

the comparison between assays is limited. The differences of PCT results are challenging to evaluate due to the absence of any information regarding materials' commutability, which hampers the ability to compare and aggregate results from different peer groups.

Although our study relies on an extensive collection of EQAS results obtained over years and presents the overall situation internationally for the first time, the variety of PCT EQAS designs, particularly the use of different materials, and their results illustrate the complexity of the situation of PCT measurements and highlight the need to conduct a definitive study to properly establish the status of harmonization of marketed PCT assays.

After this is achieved, important work will need to be done to understand the variability of results provided by the different PCT assays. Differences in results between assays can be due to the use of different calibrators and/or different antibodies [67]. It might also be caused by non-commutability of calibrators that are used in the metrological traceability chain. ISO 17511:2020 on metrological traceability in laboratory medicine stresses the need for calibrators used in the traceability chain to be commutable [72], but difficulties remain to fulfill this requirement in practice. Provided that differences between assays are mostly due to calibration-related issues rather than specificity issues, a solution could consist of standardizing calibration through commutable Certified Reference Materials or panels of patient samples value assigned with a reference method. Although a significant fraction of commercially available immunoassays provides results that are traceable to the Brahms PCT LIA assay, substantial differences between these assays have been reported in both EQA schemes and correlation studies, and even larger differences were reported with other assays. Therefore, several published correlation studies suggest that common reference calibration materials should be developed to improve agreement between PCT assays [36, 51, 56, 65, 67, 73]. To fill this gap, the IFCC working group on standardization of PCT assays initiated the production of commutable reference materials intending to evaluate the necessity, the feasibility and the potential benefit of recalibrating PCT assays [31].

IFCC WG-PCT members: Vincent Delatour, Laboratoire National de Métrologie et d'Essais, Paris, France. Amandine Boeuf, Laboratoire National de Métrologie et d'Essais, Paris, France. Qinde Liu, Chemical Metrology Division, Applied Sciences Group, Health Science Authority, Singapore. Yizhao Chen, Chemical Metrology Division, Applied Sciences Group, Health Science Authority, Singapore. Klaus Schneider, University of Applied Sciences Fresenius, Institute for Biomolecular Research, Idstein, Germany. Patricia Kaiser, Instand e.V. Reference Laboratory, Düesseldorf, Germany. Pierre Hausfater, Emergency Department, Pitie-Salpetriere University Hospital Centre, Paris, France. Philipp Schuetz, University Department of Medicine, Kantonsspital Aarau, Switzerland. Christos Tsatsanis, Laboratory of Clinical Chemistry, School of Medicine, University of Crete, Heraklion, Crete, Greece. Anne-Marie Dupuy, Department of Biochemistry, University Hospital of Montpellier, Montpellier, France. Stefano Pastori, Unit of Laboratory Medicine, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy. Jens Pfannkuche, PANACECA UG, Berlin, Germany. Chong Yuan, Diazyme Laboratories, Poway, USA. Birger Machetanz, Diazyme Laboratories, Poway, USA. Michael Rottmann, Roche Diagnostics GmbH, Penzberg, Germany. Andrew Rybin, Siemens Healthineers, New York, USA. Neil Parker, Siemens Healthineers, New York, USA. Jutta Odarjuk, Thermo Fisher Scientific, Henningsdorf, Germany. Barbara Thomas, Thermo Fisher Scientific, Henningsdorf, Germany. Michelle Solari, Beckman Coulter, Chaska, USA. Kim Paulsen, Beckman Coulter, Chaska, USA. Scott Ruetten, Abbott Diagnostics, Chicago, USA. Lauren Seaver, Abbott Diagnostics, Chicago, USA. Helene Briand, BioMérieux, Marcy l'Etoile, France. Carole Tourneur, BioMérieux, Marcy l'Etoile, France. Luca Pallavicini, DiaSorin S.p.A, Saluggia, Vercelli, Italy. Nadia Corocher, DiaSorin S.p.A, Saluggia, Vercelli, Italy. Godwin Ogbonna, Ortho Clinical Diagnostics, New York, USA. Maria Patru, Ortho Clinical Diagnostics, New York, USA. Philip Bryan, Ortho Clinical Diagnostics, New York, USA. Sameh Zion, Ortho Clinical Diagnostics, New York, USA. Matthias Grimmler, DiaSys Diagnostic Systems, Holzheim, Germany. Thomas Masetto, DiaSys Diagnostic Systems, Holzheim, Germany. Piitu Jauria, Radiometer, Turku, Finland.

Author contributions: Huu-Hien Huynh: Conceptualization, Methodology, Formal analysis, Investigation, Writing, Original draft preparation, Reviewing and Editing. Amandine Bœuf: Conceptualization, Methodology, Formal analysis, Investigation, Writing, Original draft preparation, Reviewing and Editing. Jens Pfannkuche: Conceptualization, Methodology, Writing, Original draft preparation, Reviewing and Editing. Philipp Schuetz: Conceptualization, Methodology, Writing, Original draft preparation, Reviewing and Editing. Marc Thelen: Formal analysis, Investigation, Writing, Reviewing and Editing. Gunnar Nordin: Formal analysis, Investigation, Writing, Reviewing and Editing. Eline van der Hagen: Formal analysis, Investigation, Writing, Reviewing and Editing. Patricia Kaiser: Formal analysis, Investigation, Writing, Reviewing and Editing. Dagmar Kesseler: Formal analysis, Investigation, Writing, Reviewing and Editing. Tony Badrick: Formal analysis, Investigation, Writing, Reviewing and Editing. Bernard Poggi: Formal analysis, Investigation, Writing, Reviewing and Editing. Ulla Tiikkainen: Formal analysis, Investigation, Writing, Reviewing and Editing. Gareth J Davies: Formal analysis, Investigation, Writing, Reviewing and Editing. Anja Kessler: Formal analysis, Investigation, Writing, Reviewing and Editing. Mario Plebani: Writing, Reviewing and Editing. Joëlle Vinh: Writing, Original draft preparation, Reviewing and Editing. Vincent Delatour: Conceptualization, Methodology, Writing, Original draft preparation, Supervision. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

Informed consent: Not applicable.

Ethical approval: Not applicable.

References

1. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet 2020;395:200–11.
2. Reinhart K, Daniels R, Kissoon N, Machado FR, Schachter RD, Finfer S. Recognizing sepsis as a global health priority — a WHO resolution. N Engl J Med 2017;377:414–7.
3. Weiss SL, Fitzgerald JC, Pappachan J, Wheeler D, Jaramillo-Bustamante JC, Salloo A, et al. Global epidemiology of pediatric severe sepsis: the sepsis prevalence, outcomes, and therapies study. Am J Respir Crit Care Med 2015;191:1147–57.
4. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, et al. Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med 2017;376:2235–44.
5. Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med 2006;34:1589–96.
6. Ferrer R, Artigas A, Levy MM, Blanco J, González-Díaz G, Garnacho-Montero J, et al. Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain. J Am Med Assoc 2008;299:2294–303.
7. Kiers HD, Griesdale DEG, Litchfield A, Reynolds S, Gibney RTN, Chittock D, et al. Effect of early achievement of physiologic resuscitation goals in septic patients admitted from the ward on the kidneys. J Crit Care 2010;25:563–9.
8. Schuetz P, Balk R, Briel M, Kutz A, Christ-Crain M, Stolz D, et al. Economic evaluation of procalcitonin-guided antibiotic therapy in acute respiratory infections: a US health system perspective. Clin Chem Lab Med 2015;53:583–92.
9. Mewes JC, Pulia MS, Mansour MK, Broyles MR, Nguyen HB, Steuten LM. The cost impact of PCT-guided antibiotic stewardship versus usual care for hospitalized patients with suspected sepsis or lower respiratory tract infections in the US: a health economic model analysis. PloS One 2019;14:e0214222.
10. Vijayan AL, Vanimaya, Ravindran S, Saikant R, Lakshmi S, Kartik R, et al. Procalcitonin: a promising diagnostic marker for sepsis and antibiotic therapy. J Intensive Care 2017;5:51.
11. Oberhoffer M, Stonans I, Russwurm S, Stonane E, Vogelsang H, Junker U, et al. Procalcitonin expression in human peripheral blood mononuclear cells and its modulation by lipopolysaccharides and sepsis-related cytokines in vitro. J Lab Clin Med 1999;134:49–55.
12. Müller B, Becker KL. Procalcitonin: how a hormone became a marker and mediator of sepsis. Swiss Med Wkly 2001;131: 595–602.
13. Jeong S, Park Y, Cho Y, Kim HS. Diagnostic utilities of procalcitonin and C-reactive protein for the prediction of bacteremia determined by blood culture. Clin Chim Acta 2012; 413:1731–6.
14. Vincent J-L, Beumier M. Diagnostic and prognostic markers in sepsis. Expert Rev Anti Infect Ther 2013;11:265–75.
15. Kim SJ, Hwang SO, Kim YW, Lee JH, Cha K-C. Procalcitonin as a diagnostic marker for sepsis/septic shock in the emergency department; a study based on sepsis-3 definition. Am J Emerg Med 2019;37:272–6.
16. Wanner GA, Keel M, Steckholzer U, Beier W, Stocker R, Ertel W. Relationship between procalcitonin plasma levels and severity of injury, sepsis, organ failure, and mortality in injured patients. Crit Care Med 2000;28:950–7.
17. AlRawahi AN, AlHinai FA, Doig CJ, Ball CG, Dixon E, Xiao Z, et al. The prognostic value of serum procalcitonin measurements in critically injured patients: a systematic review. Crit Care 2019;23:390.
18. Cabral L, Afreixo V, Almeida L, Paiva JA. The use of procalcitonin (PCT) for diagnosis of sepsis in burn patients: a meta-analysis. PloS One 2016;11:e0168475.
19. Ren H, Li Y, Han C, Hu H. Serum procalcitonin as a diagnostic biomarker for sepsis in burned patients: a meta-analysis. Burns 2015;41:502–9.

20. Tan M, Lu Y, Jiang H, Zhang L. The diagnostic accuracy of procalcitonin and C-reactive protein for sepsis: a systematic review and meta-analysis. J Cell Biochem 2019;120:5852–9.

21. Sager R, Wirz Y, Amin D, Amin A, Hausfater P, Huber A, et al. Are admission procalcitonin levels universal mortality predictors across different medical emergency patient populations? Results from the multi-national, prospective, observational TRIAGE study. Clin Chem Lab Med 2017;55:1873–80.

22. Schuetz P, Birkhahn R, Sherwin R, Jones AE, Singer A, Kline JA, et al. Serial procalcitonin predicts mortality in severe sepsis patients: results from the Multicenter Procalcitonin Monitoring Sepsis (MOSES) study. Crit Care Med 2017;45:781–9.

23. Liu D, Su L, Han G, Yan P, Xie L. Prognostic value of procalcitonin in adult patients with sepsis: a systematic review and meta-analysis. PloS One 2015;10:e0129450.

24. Vouloumanou EK, Plessa E, Karageorgopoulos DE, Mantadakis E, Falagas ME. Serum procalcitonin as a diagnostic marker for neonatal sepsis: a systematic review and meta-analysis. Intensive Care Med 2011;37:747–62.

25. Schuetz P, Müller B, Christ-Crain M, Stolz D, Tamm M, Bouadma L, et al. Procalcitonin to initiate or discontinue antibiotics in acute respiratory tract infections. Cochrane Database Syst Rev 2012;9: CD007498.

26. Schuetz P, Bretscher C, Bernasconi L, Mueller B. Overview of procalcitonin assays and procalcitonin-guided protocols for the management of patients with infections and sepsis. Expert Rev Mol Diagn 2017;17:593–601.

27. Sager R, Kutz A, Mueller B, Schuetz P. Procalcitonin-guided diagnosis and antibiotic stewardship revisited. BMC Med 2017;15:15.

28. Bouadma L, Luyt C-E, Tubach F, Cracco C, Alvarez A, Schwebel C, et al. Use of procalcitonin to reduce patients' exposure to antibiotics in intensive care units (PRORATA trial): a multicentre randomized controlled trial. Lancet 2010;375:463–74.

29. de Jong E, van Oers JA, Beishuizen A, Vos P, Vermeijden WJ, Haas LE, et al. Efficacy and safety of procalcitonin guidance in reducing the duration of antibiotic treatment in critically ill patients: a randomized, controlled, open-label trial. Lancet Infect Dis 2016; 16:819–27.

30. Bloos F, Trips E, Nierhaus A, Briegel J, Heyland DK, Jaschinski U, et al. Effect of sodium selenite administration and procalcitonin-guided therapy on mortality in patients with severe sepsis or septic shock: a randomized clinical trial. JAMA Intern Med 2016; 176:1266–76.

31. Huynh H-H, Bœuf A, Vinh J, Delatour V, Delatour V, Bœuf A, et al. Evaluation of the necessity and the feasibility of the standardization of procalcitonin measurements: activities of IFCC WG-PCT with involvement of all stakeholders. Clin Chim Acta 2021;515:111–21.

32. Schuetz P, Beishuizen A, Broyles M, Ferrer R, Gavazzi G, Gluck EH, et al. Procalcitonin (PCT)-guided antibiotic stewardship: an international experts consensus on optimized clinical use. Clin Chem Lab Med 2019;57:1308–18.

33. Vesper HW, Myers GL, Miller WG. Current practices and challenges in the standardization and harmonization of clinical laboratory tests. Am J Clin Nutr 2016;104(3 Suppl):907S–12S.

34. Roche. Elecsys® BRAHMS procalcitonin (PCT). Available from: https://diagnostics.roche.com/ch/fr/products/params/elecsys-brahms-procalcitonin-pct.html [Accessed 26 Nov 2020].

35. National Institute for health and Care Excellence. Procalcitonin testing for diagnosing and monitoring sepsis (ADVIA Centaur BRAHMS PCT assay, BRAHMS PCT Sensitive Kryptor assay, Elecsys BRAHMS PCT assay, LIAISON BRAHMS PCT assay and VIDAS BRAHMS PCT assay)-2015. Available from: https://www.nice.org.uk/guidance/dg18 [Accessed 26 Nov 2020].

36. Dipalo M, Guido L, Micca G, Pittalis S, Locatelli M, Motta A, et al. Multicenter comparison of automated procalcitonin immunoassays. Pract Lab Med 2015;2:22–8.

37. Ruzzenente O, Salvagno GL, Gelati M, Lippi G. Analytical evaluation of the novel Lumipulse G BRAHMS procalcitonin immunoassay. Pract Lab Med 2016;6:8–13.

38. Food and Drug Administration. 510(k) premarket notification K170652 Architect Brahms PCT. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K170652 [Accessed 15 Jun 2020].

39. Food and Drug Administration. 510(k) premarket notification K173683 Liaison Brahms PCT II gen. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K173683 [Accessed 15 Jun 2020].

40. Food and Drug Administration. 510(k) premarket notification K172713 Lumipulse G Brahms PCT. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K172713 [Accessed 15 Jun 2020].

41. Food and Drug Administration. 510(k) premarket notification K173927 Elecsys Brahms PCT. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K173927 [Accessed 15 Jun 2020].

42. Food and Drug Administration. 510(k) premarket notification K181002 atellica IM Brahms PCT. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K181002 [Accessed 15 Jun 2020].

43. Food and Drug Administration. 510(k) premarket notification K162297 Diazyme procalcitonin assay. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K162297 [Accessed 15 Jun 2020].

44. Food and Drug Administration. 510(k) premarket notification K200236 Vitros Brahms PCT. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K200236 [Accessed 15 Jun 2020].

45. Food and Drug Administration. 510(k) premarket notification K192271 access PCT. Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf19/K192271.pdf [Accessed 25 Now 2020].

46. Lagabrielle J-F, Tachet A, Boin V. Dosage de la procalcitonine: comparaison des résultats obtenus sur Kryptor® (Brahms) et Vidas® (BioMérieux). Immuno-Anal Biol Specialisee 2008;23: 245–50.

47. De Wolf HK, Gunnewiek JK, Berk Y, van den Ouweland J, de Metz M. Comparison of a new procalcitonin assay from Roche with the established method on the Brahms Kryptor. Clin Chem 2009;55: 1043–4.

48. Prieto B, Alvarez FV. Lack of transferability of results between procalcitonin assays. Clin Chem 2009;55:2226–7.

49. Sanders R-J, Schoorl M, Dekker E, Snijders D, Boersma W, Boekel E. Evaluation of a new procalcitonin assay for the Siemens ADVIA Centaur with the established method on the B.R.A.H.M.S Kryptor. Clin Lab 2011;57:415–20.

50. Lloyd MM, Kuyl JM. Comparison of three methods for procalcitonin analysis. Med Technol SA 2012;26:48–52.

51. Dipalo M, Buonocore R, Gnocchi C, Picanza A, Aloe R, Lippi G. Analytical evaluation of Diazyme procalcitonin (PCT) latex-

enhanced immunoturbidimetric assay on Beckman Coulter AU5800. Clin Chem Lab Med 2015;53:593–7.

52. Dupuy AM, Bargnoux AS, Andreeva A, Zins C, Kuster N, Badiou S, et al. Analytical performances of a novel point-of-care procalcitonin assay. Pract Lab Med 2020;18:e00145.

53. Fortunato A. A new sensitive automated assay for procalcitonin detection: LIAISON® BRAHMS PCT® II GEN. Pract Lab Med 2016; 6:1–7.

54. Ceriotti F, Marino I, Motta A, Carobene A. Analytical evaluation of the performances of Diazyme and BRAHMS procalcitonin applied to Roche Cobas in comparison with BRAHMS PCT-sensitive Kryptor. Clin Chem Lab Med 2017;56:162–9.

55. Lippi G, Salvagno GL, Gelati M, Pucci M, Demonte D, Faggian D, et al. Analytical evaluation of the new Beckman Coulter access procalcitonin (PCT) chemiluminescent immunoassay. Diagnostics 2020;10:128.

56. Eidizadeh A, Asif AR, von Ahsen N, Binder L, Schnelle M. Differences in procalcitonin measurements between three BRAHMS-partnered immunoassays (Liaison, Elecsys and Architect). Clin Chem Lab Med 2019;57:e207–10.

57. Mouatani AE, Rotaru I, Beaudeux J-L, Hennequin C. Evaluation of procalcitonin assay on the Abbott Architect i2000 SR® analyzer. Ann Biol Clin 2018;76:308–12.

58. Chambliss AB, Hayden J, Colby JM. Evaluation of procalcitonin immunoassay concordance near clinical decision points. Clin Chem Lab Med 2019;57:1414–21.

59. Dipalo M, Gnocchi C, Avanzini P, Musa R, Di Pietro M, Aloe R. Comparison of procalcitonin assays on KRYPTOR and LIAISON® XL analyzers. Diagnostics 2019;9:94.

60. Gruzdys V, Cahoon K, Pearson L, Lehman CM. Method verification shows a negative bias between 2 procalcitonin methods at medical decision concentrations. J Appl Lab Med 2019;4:69–77.

61. Dupuy AM, Bargnoux AS, Larcher R, Merindol A, Masetto T, Badiou S, et al. Bioanalytical performance of a new particle-enhanced method for measuring procalcitonin. Diagnostics 2020;10:461.

62. Hausfater P, Brochet C, Freund Y, Charles V, Bernard M. Procalcitonin measurement in routine emergency medicine practice: comparison between two immunoassays. Clin Chem Lab Med 2010;48:501–4.

63. Leung A, Hau CI, Lai RWM. Comparison of a new automated Abbott Architect procalcitonin assay with biomerieux Vidas and Roche Elecsys methods. J Clinc Lab Med 2017;2:117–20.

64. Schuetz P, Christ-Crain M, Huber AR, Müller B. Long-term stability of procalcitonin in frozen samples and comparison of Kryptor and VIDAS automated immunoassays. Clin Biochem 2010;43:341–4.

65. Soh A, Binder L, Clough M, Hernandez MH, Lefèvre G, Mostert K, et al. Comparison of the novel ARCHITECT procalcitonin assay with established procalcitonin assay systems. Pract Lab Med 2018;12:e00110.

66. Wang D, Caddell B, Nolte FS, Babic N. Comparison of the Abbott Architect BRAHMS and the biomérieux Vidas BRAHMS procalcitonin assays. J Appl Lab Med 2019;3:580–6.

67. Lippi G, Salvagno GL, Gelati M, Pucci M, Lo Cascio C, Demonte D, et al. Two-center comparison of 10 fully-automated commercial procalcitonin (PCT) immunoassays. Clin Chem Lab Med 2019;58: 77–84.

68. Katz SE, Sartori LF, Szeles A, McHenry R, Stanford JE, Xu M, et al. Agreement between two procalcitonin assays in hospitalized children. Infect Dis Ther 2019;8:463–8.

69. Stralen KJV, Jager KJ, Zoccali C, Dekker FW. Agreement between methods. Kidney Int 2008;74:1116–20.

70. Bilić-Zulle L. Comparison of methods: passing and Bablok regression. Biochem Med 2011;21:49–52.

71. Delatour V, Clouet-Foraison N, Jaisson S, Kaiser P, Gillery P. Trueness assessment of HbA1c routine assays: are processed EQA materials up to the job? Clin Chem Lab Med 2019;57:1623–31.

72. ISO. ISO 17511:2020 in vitro diagnostic medical devices – requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. Available from: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/99/69984.html [Accessed 25 Oct 2021].

73. Wojtalewicz N, Schellenberg I, Hunfeld K-P. Evaluation of INSTAND e.V.'s external quality assessment for C-reactive protein and procalcitonin. PloS One 2019;14:e0221426.