

## Delta Checks in the clinical laboratory

Edward W Randell & Sedef Yenice

To cite this article: Edward W Randell & Sedef Yenice (2019): Delta Checks in the clinical laboratory, *Critical Reviews in Clinical Laboratory Sciences*

To link to this article: <https://doi.org/10.1080/10408363.2018.1540536>



Published online: 11 Jan 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)

---



## Delta Checks in the clinical laboratory

Edward W Randell<sup>a,b</sup> and Sedef Yenice<sup>c</sup>

<sup>a</sup>Discipline of Laboratory Medicine, Memorial University; Eastern Health Authority, St. John's, NL, Canada; <sup>b</sup>Faculty of Medicine, Memorial University; Eastern Health Authority, St. John's, NL, Canada; <sup>c</sup>Department of Core Laboratory Medicine, Gayrettepe Florence Nightingale Hospital, Istanbul, Turkey

### ABSTRACT

International standards and practice guidelines recommend the use of delta check alerts for laboratory test result interpretation and quality control. The value of contemporary applications of simple univariate delta checks determined as an absolute change, percentage change, or rate of change to recognize specimen misidentification or other laboratory errors has not received much study. This review addresses these three modes of calculation, but in line with the majority of published work, most attention is focused on the identification of specimen misidentification errors. Investigation of delta check alerts are time-consuming and the yield of identified errors is usually small compared to the number of delta check alerts; however, measured analytes with low indices of individuality frequently perform better. While multivariate approaches to delta checks suggest improved usefulness over simple univariate delta check strategies, some of these are complex and not easily applied in contemporary laboratory information systems and middleware. Nevertheless, a simple application of delta checks may hold value in identifying clinically significant changes in several clinical situations: for acute kidney injury using changes in serum creatinine, for risk of osmotic demyelination syndrome using rapid acute changes in serum sodium levels, or for early triage of chest pain patients using high sensitivity troponin assays. A careful and highly selective approach to identifying delta check analytes, calculation modes, and thresholds before putting them into practice is warranted; then follow-up with careful monitoring of performance and balancing true positives, false negatives, and false positives among delta check alerts is needed.

**Abbreviations and glossary:** AD $\Delta$ : absolute difference delta; ARD $\Delta$ : absolute rate difference delta; AKI: acute kidney injury; ALT: alanine aminotransferase; ALP: alkaline phosphatase; AMI: acute myocardial infarction; AST: aspartate aminotransferase; BUN: blood urea nitrogen; CAP: College of American Pathologists; CBC: complete blood count; CK: creatine kinase; CCD: composite CBC delta; CLSI: Clinical Laboratory Standards Institute; CV: coefficient of variation; CV<sub>A</sub>: analytical variability (imprecision); CV<sub>I</sub>: intra-individual variability; CV<sub>G</sub>: inter-individual variability; HIL: hemolysis icterus lipemia; HMR: hemoglobin, MCH, red blood cell distribution width; LIS: laboratory information system; LDC: logical delta check; MCH: mean corpuscular hemoglobin; MCV: mean corpuscular volume; ODS: osmotic demyelination syndrome; PD $\Delta$ : percentage difference delta; PPV: positive predictive value; PRD $\Delta$ : percentage rate difference delta; QC: quality control; RCV: reference change value; TAT: turnaround time; WBC: white blood cells

### ARTICLE HISTORY

Received 13 July 2018  
Accepted 21 October 2018  
Published online 10 January 2019

### KEYWORDS

Delta Checks; error identification; index of individuality; patient based quality control; reference change value; specimen misidentification

## Introduction

Monitoring clinical change in patients is a common reason for ordering laboratory tests. Changes in laboratory test results over sequential and consecutive samples can occur due to factors that impact the pre-analytical, analytical, and post-analytical phases of laboratory testing, and because of biological variability, pathophysiological factors, and errors. The need for quality assurance procedures in addition to routine quality control (QC) arises because most laboratory errors occur during the pre-analytical phase [1], and almost 50% of these are related to

specimen collection and labeling errors [2]. The impact of laboratory errors on patient care involves quality, safety, and financial costs. Strategies have been developed to address clinically significant errors by using patient data to assist with error detection [3,4]. Delta checks and checks for absurd and unusual test results or other indicators for specimen integrity and result accuracy are used to augment the laboratory's ability to detect errors, particularly those originating at the pre-analytical phase. Delta checks, as they are applied here, are defined as comparison of the difference between measurements of

an analyte (or combinations of analytes) on two separate samples from the same patient to predefined thresholds that represent the limits of acceptable change.

Delta checks are applied to assess specimen integrity and detect error prior to the release of results into the patient record. Several delta check techniques have evolved through different calculation modes, action thresholds, and other criteria based on their purpose. Beyond the totally nonsensical result, it is an enormous challenge for the laboratory to verify the validity of any single laboratory test result in isolation. The situation differs when there is more than one test result from serial samples which are collected at different time points from the same patient. The delta check alert can identify a testing problem that occurred either in the previous or the current specimen, prompting corrective action concerning the previously reported result or the result currently held for reporting. If established thresholds are exceeded, repeat measurement of both the previous and the current specimens can be done, and/or an investigation can be conducted to determine if the change is due to true clinical change or due to error, especially mislabeling of the original sample or an aliquot or patient misidentification error (Figure 1; Table 1). For the purpose of this review, misidentified specimens refer to all cases where a specimen and/or its results are assigned to the wrong patient, and the error is referred to as a specimen misidentification error.

Study of delta check performance in real-world situations is expensive and labor intensive [5]; hence only a few studies have taken this approach. More common is the use of error simulation studies in which laboratory data is downloaded and manipulated to simulate errors to assess the performance of error detection strategies. The most studied simulated errors are those related to misidentified specimens. As with any test, the delta check carries with it the potential for true positive and true negative, and false-positive and false negative results. When used in the context of error detection, a true positive delta check identifies a real error, while a true negative indicates no significant error or change in the patient status as supported by test results. A false-positive can occur with a significant change in test results for reasons other than an error. A false negative indicates a situation where the error has occurred but the degree of change is less than the delta check threshold and no alert is generated. False-positive delta checks are common in hospital in-patients, particularly in the critically ill [6], and carry with them increased frequency of test ordering, unnecessary follow-up, delays in reporting, and wasted effort in investigating. False negatives

indicate missed opportunities for intervention or potential delay in identifying significant change occurring in patients or for corrective action in case of error. When a specimen misidentification error has occurred, the effectiveness of a delta check to detect the error and its clinical significance is related to the magnitude of differences between different individuals [7] compared to the individual.

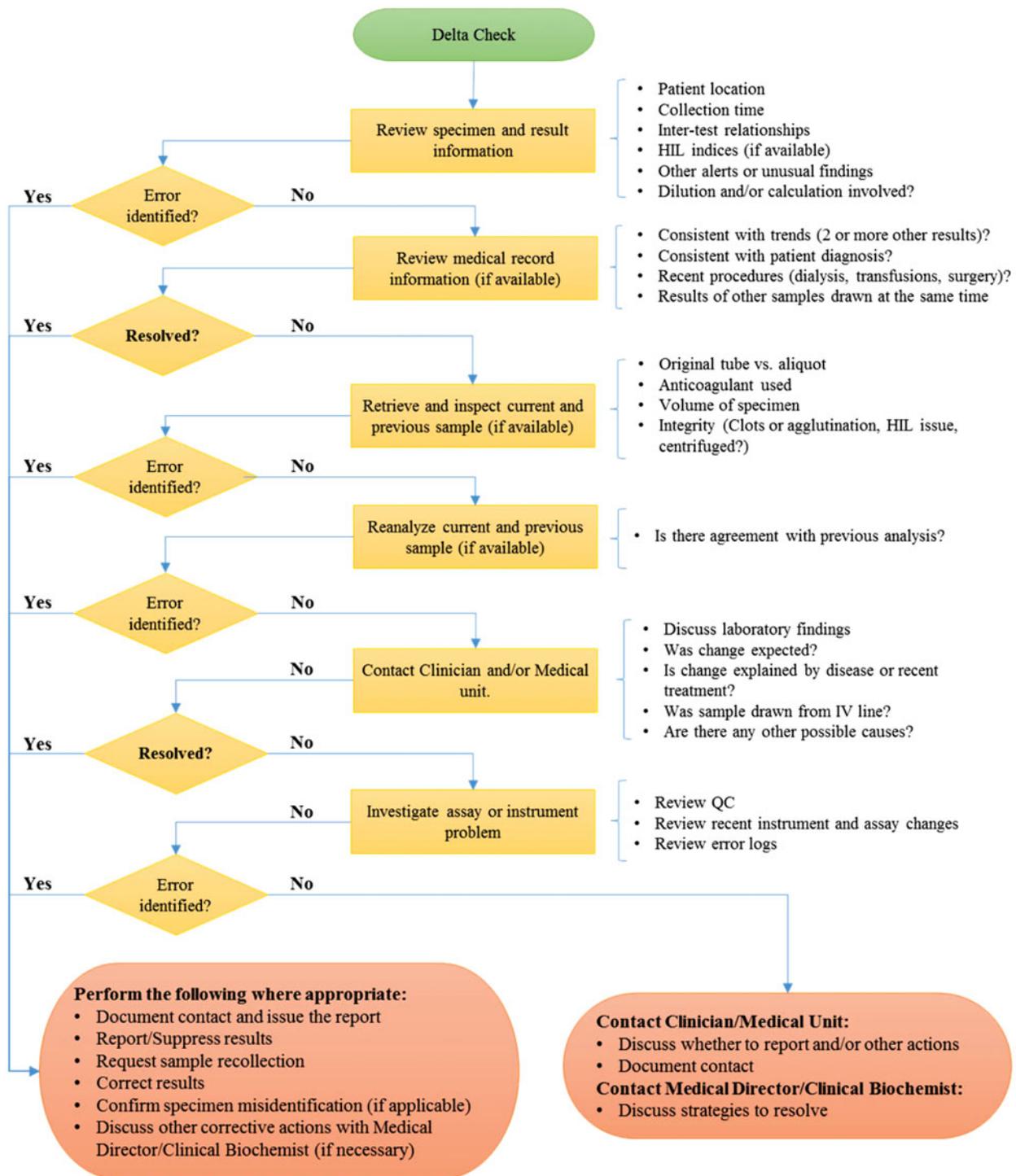
For almost five decades, laboratories have been using delta checks to detect change greater than that expected from normal physiological variation [8]. The following section provides a brief historical perspective on the use of delta checks. Recent work has challenged the usefulness of its most common application in early detection of laboratory error, particularly specimen misidentification errors, and, in spite of its common use, investigations into how to select tests for delta checks and to optimize their use are sparse, although a recent guideline has attempted to help fill this gap [9].

To date, a limited number of review papers and opinion pieces have been published that address delta check strategy and/or use of patient data to identify laboratory error [3–5,10]. The objective of this review is to provide the reader with a comprehensive overview on delta checks, especially their use for error detection, and to promote a rational approach to their use.

## Historical perspective

While routine QC processes have been designed to detect analytical errors, they are ineffective in detecting misidentified specimens or errors that may occur between analysis and reporting of test results. The delta check concept arose in an era prior to the widespread use of barcode-labeled patient wristbands, barcode labeling of tubes, and modern, rigorous patient identification procedures. In the absence of these advances, the frequency of mislabeled primary collection tubes and aliquots was significantly higher than it is today [11–13].

Possibly the first reference to the concept of delta checks occurred more than half a century ago, in 1967, as part of a proposed laboratory computer algorithm to detect unusual observations in relation to the goal of recognizing quality issues occurring at and beyond the test assay workbench [8]. Among the unusual observations were test results deemed incompatible with life; changes in test results for an individual that exceeded certain limits (the delta check concept); and recognition of patterns built on known relationships between different tests on the same specimen. In 1974, Nosanchuk and Gottmann [14] described the retrospective



**Figure 1.** Process workflow for investigation of a delta check alert.

examination of serial consecutive laboratory results of complete blood counts (CBCs) and differentials for individual patients undertaken each time a new analysis was performed, with the goal of detecting discrepancies before results left the laboratory. In contrast to the laboratory information system (LIS) or middleware-based approaches commonly used today, Nosanchuk and Gottmann used manual surveillance, and an

unexpected change in test results was followed up by an examination of the entire process, a query to nurses or physicians, or chart review, depending on the situation. Errors identified were mainly clerical and specimen identification errors. The majority of delta check alerts (98%) were explained by a true change in the patient's condition, and the process represented a new burden of costs for the laboratory.

**Table 1.** Investigations and responses to Delta Check alerts. List compiled from reviewed articles, especially Schiffman et al. [59].

Target	Action
Specimen	Repeat analysis of the current specimen Repeat analysis of the previous specimen Inspect the specimen (original tube versus aliquot, integrity, anticoagulant, volume) Inspect specimen labeling and/or requisition
LIS and laboratory	Note location of the patient during collection Track specimen history/processing (collection time, transport and storage conditions, anticoagulant/container, centrifugation, dilution, calculations) Review other test results including HIL indices on the sample Compare with other tests based on inter-test relationships on the specimen Order and examine other tests Discuss with other laboratory areas Check ABO blood type or other RBC antibody/antigen status on the present and previous samples Check results of any drugs being monitored.
Medical record	Document unusual or suspicious findings on reports Review result trends Review information in patient electronic medical record
Laboratory information	Confirm patient diagnosis Review of routine QC Review recent instrument or assay changes including reagent/calibrator lot number changes Review calculations used
Clinicians and medical units	Discuss with medical unit or health care provider (consider the possibility of contamination by anticoagulant/IV fluid; transfusion history; dialysis; recent treatment or procedure) Discuss and/or request and repeat analyses on a fresh sample Document contact and discussions on report

The work done throughout the mid- and late-1970s and early 1980s showed a maturing of the delta check concept similar to the way it is used today. Of these advances, Ladenson [15] made use of a computer and select clinical chemistry and immunoassay tests, and considered, but did not adopt, thresholds based on biological variability that had been described by Young et al. [16]. Contemporary with this report, Whitehurst et al. [17] described a system involving routine clinical chemistry tests, but they also examined change in the calculated anion gap, the first multivariate approach to delta checking. Up to this point, all delta check thresholds had been set empirically. In a move towards a more systematic approach for identifying delta check thresholds, Wheeler and Sheiner [18] used archived laboratory data to determine probabilities of change affecting six commonly measured clinical chemistry tests and two calculated parameters, the anion gap, and the urea (or blood urea nitrogen, BUN) to creatinine ratio. This approach was more complex than previous ones as it applied different delta check thresholds based on test result categories for the current result, used two different time intervals between specimens, and used seven different probability-based thresholds for which different actions were ascribed. Sher [6] examined a variety of clinical chemistry analytes, including anion gap, using a computerized delta check strategy that extended the delta check interval to a 30-d period; this represented a significant change from shorter intervals of <4 d evaluated by predecessors. The strategy resulted in a 1.6% positivity rate but only

16% of these delta check alerts represented an error. While specimen misidentification errors (22%) were the most common, other errors included specimen mishandling and instrument failures. Later, work by Sheiner et al. [19] and Wheeler and Sheiner [20], which evaluated delta check strategies published over the previous decade, concluded that all approaches showed similar performance in terms of the frequency of alerts and errors detected. Recognized as a challenge by these early studies was balancing error detection with the work required to evaluate and rule-out errors, as changes in the majority of results were explained by pathophysiology or clinical intervention.

### Delta check calculation modes

The simplest application of the delta check concept involves its use for qualitative data, for example, as applied to serum protein electrophoresis gel [21] or for blood grouping results [22]. In these cases, significant qualitative changes could indicate sample misidentification. Different calculations of delta check values are possible when quantitative data is available, including absolute, percentage, the rate of change, and various multivariate approaches [23–26]. Figure 2 shows the common calculation modes used for delta checks. Many calculations are possible including variations of the common percentage difference delta (PΔ) check equation in which the difference between the highest and lowest is expressed as a percentage of the lowest [27] rather than the previous result, or simplified

approaches that apply the ratio of the current value to the preceding one [28] or use the natural log of this ratio [29]. Based on differences in the frequency distribution of results for tests, some equations work better than others when predicting positivity rates [30], but it is not clear how this translates into improved delta check efficiency in error detection.

### Absolute versus percentage change

The most commonly-used calculations for delta checks make use of either an absolute difference (AD $\Delta$ ) or PDA over a predefined time interval. Ladenson's [15] strategy used all PDAs with specified thresholds ranging from 5% for sodium to 99% for creatine kinase (CK), but thresholds around 20% were used for most tests. In contrast, Wheeler and Sheiner [18] used AD $\Delta$ s exclusively. While the use of AD $\Delta$  seems applicable to analytes like chloride or sodium where concentrations are held within narrow homeostatic limits, both within and between individuals, the greater magnitude of change possible for such tests as enzymes, urea and creatinine make AD $\Delta$  less preferred, especially for higher values of these analytes. All other reports of delta check use both AD $\Delta$  and PDA, with some analytes having either or both calculation modes applied. For example, Whitehurst et al. [17] used an AD $\Delta$  check threshold of 44  $\mu\text{mol/L}$  when the current creatinine result was less than 177  $\mu\text{mol/L}$ , but a PDA threshold of 25% when the result was above 177  $\mu\text{mol/L}$ . Sher [6] likewise used two calculation modes. These studies demonstrate the greater utility of AD $\Delta$  for results in the lower part of the measurement range. However, PDA checks allow greater consistency in delta checking for analytes demonstrating a broad dynamic range.

### Multivariate delta checks

Use of multivariate approaches to delta checking compared to univariate approaches can improve error detection [5]. While test pipetting-, reagent- or calibrator-related problems are more likely to result in a single test delta check alert, specimen misidentification and collection related errors are more likely to be the cause of multiple delta check alerts [31]. Multivariate delta checks require measurement of change in two or more tests with univariate delta checks alerts; complex calculations of change involving more than one analyte over consecutive samples; or changes in calculated parameters such as anion gap or the urea to creatinine ratio. Some multivariate delta check approaches take advantage of internal relationships between concentrations

### RCV calculation [40]

$$\%RCV = \sqrt[3]{Z} \times z \times \sqrt{CV_A^2 + CV_I^2}$$

Z = For 2 tailed analyses:

- 1.96 at 95% probability ("significant") for change in either direction;
- 2.58 at 99% probability ("highly significant") for change in either direction;

CV<sub>A</sub> = analytical variation (from QC) = (standard deviation)/Mean \* 100%

CV<sub>I</sub> = intra-individual variation

### Absolute delta (ADA) check

$$ADA = \text{Current test value} - \text{Previous test value}$$

### Percentage delta (PDA) check

$$PDA = \frac{\text{Highest value of Previous and Current} - \text{Lowest value of Previous and Current}}{\text{Lowest of Previous and Current}} \times 100\%$$

or

$$PDA = \frac{\text{Current test value} - \text{Previous test value}}{\text{Previous test value}} \times 100\%$$

### Absolute rate-difference delta (ARDA) check [37]

$$ARDA = \frac{\text{Current test value} - \text{Previous test value}}{\text{time interval between sample}}$$

### Percentage rate-difference delta (PRDA) check [37]

$$PRDA = \left[ \frac{\text{Current test value} - \text{Previous test value}}{\text{Previous test value}} / \text{Time} \right] \times 100\%$$

**Figure 2.** Calculations for RCV and common delta check calculations.

of analytes. The results of many commonly measured analytes correlate with others in a sample [32] and these relationships are held during changes in the levels of correlated tests [33]. Some tests show highly positive correlation (e.g. alanine aminotransferase [ALT] and aspartate aminotransferase [AST], direct and total bilirubin, and total protein and albumin in liver dysfunction; hemoglobin and hematocrit in blood volume and bone marrow dysfunction; or creatinine, urea, and uric acid in renal dysfunction). Other tests, including bicarbonate and chloride, show a negative correlation. High correlation during change for different analytes indicates maintenance of underlying physiological relationships during illness or response to therapeutic intervention. While the presence of these associations may be helpful in confirming real change, discordances in the degree or direction between correlated analytes can indicate an error [33].

Delta check alert criteria based on multiple univariate delta check parameters is also called the "multitem univariate delta check" [34], and can improve identification of errors [18]. In early studies using this approach, two or more delta check alerts among eight tests results yielded true positive rates of 23 to 84% [19, 20], depending on whether the goal was to detect all errors or only misidentified specimens. Improvement in the specificity for error detection using the multiple univariate delta check approach was supported by Dufour et al. [35] in a study of 498 cases in which two or more tests exceeding delta check thresholds identified laboratory errors, misidentified specimens, and specimen mishandling. However, as with purely univariate

alerts, the most common cause was a change in the patient's condition. Using simulation study design, Rheem and Lee [34] concluded that inclusion of at least four items and especially including total cholesterol, albumin, and total protein, reduced false-positive delta check alerts to near zero. This increase in specificity came with a significant loss in sensitivity to <50% for error. Of course, the sensitivity and specificity of this strategy will also vary depending on the specific delta check thresholds used for each test.

Calculated parameters involving more than one analyte leverages physiological relationships between analytes to improve the error yield. Delta checks on calculated parameters can be used in two ways. Calculated parameters can be more sensitive to subtle changes in more than one test compared to univariate delta checks and in situations in which changes in individual test results are insufficient to surpass delta check thresholds. Alternatively, change in the calculated parameter result can be used to determine whether there is a need to address an individual test result which also has a delta check alert. The main clinical use of the anion gap is in the differential diagnosis of acid-base disorders, while the urea to creatinine ratio is helpful in evaluating renal impairment. A significant change in these parameters, especially over short time periods and in the absence of obvious clinical explanation, highlights the potential for error. Cause for skepticism concerning this approach is substantiated by studies which reveal that the performance of anion gap based delta checks is inferior to its univariate components (sodium, potassium, chloride, and bicarbonate) [25]. No detailed examination of delta checks of the urea to creatinine ratio compared to its univariate components has been published.

Several studies have examined more complex calculations involving several analytes in an effort to improve the sensitivity and specificity of analyte changes for specimen misidentification errors. Early work using complex functions derived from analysis of laboratory test panels and changes in sodium, potassium, chloride, bicarbonate, urea, and creatinine to produce an approach superior to univariate delta checks came up empty-handed [19]. Multivariate delta check calculation modes can vary significantly in complexity. Iizuka et al. [26] provided a complex multivariate approach based on the "Mahalanobis distance", or the difference between correlating test result combinations for two samples. By this method, the Mahalanobis difference ( $D^2$ ) was computed and compared against a Chi-squared statistic for probability and designated probability for false rejection. It was demonstrated that up

to 84% of specimen misidentification errors could be identified but with a false-positive rate of 6%. Because of the excessive computer processing times required, three tests (zinc sulfate turbidity, cholinesterase, and cholesterol) were selected for application of this strategy, which revealed an error detection rate of 65% but with a 2.2% false-positive rate. This meant that only 3.3% of delta check alerts represented true positives in an environment where the true specimen misidentification error rate was about 0.1%. This approach, which was later evaluated by Furutani et al. [36], was shown to be effective in error detection (almost 6% of anomalies found representing true error), but the strategy was not shown to be superior to the univariate threshold-based delta checks. Using a different approach, Yamashita et al. [24] corrected for distribution differences and non-Gaussian distribution for common clinical chemistry, hematology, and coagulation tests. This approach involved using a calculated index for specimen misidentification error that weighted summed differences between multiple analytes that had been transformed and converted to z values and then adjusted by weighting to minimize the impact of analytes with larger adjusted standard deviations. Maximum sensitivity was achieved when about 10 different test results were used in the calculated index; sensitivities ranged from about 90–95% for misidentified specimens, but false-positive rates ranged from about 5–10%. Complex multivariate delta check strategies remain a challenge for implementation in contemporary LIS and middleware software, and are without clear evidence for superiority over univariate approaches.

More recent work gives room for optimism concerning the use of multivariate approaches. Miller [23] attempted to improve the sensitivity of delta checks applied to CBC results by using weighed changes in multiple parameters in a calculated parameter called the composite CBC delta (CCD). Selection of CBC components for the CCD was first determined by examining the discriminating power of individual CBC components for inter-individual change over intra-individual differences. The final equation, which was the square root of the sum of squares for changes in hemoglobin, mean corpuscular hemoglobin (MCH), red blood cell distribution width, and platelets, was compared to an assigned threshold and was applied over a 14-d period. A similar hemoglobin, MCH, red blood cell distribution width (HMR) calculation, which dropped the platelet term, was also assigned a threshold and evaluated. A logical delta check (LDC), which was defined, generated an alert when both CCD and HMR were above thresholds when applied over a 35-d period. The discriminatory

power of the CCD and LDC outperformed the univariate mean corpuscular volume (MCV) delta check (increase/decrease by  $>3$  fL over 3 d). The LCD generated alerts for 2% of samples, a third of the number generated using a univariate MCV delta check. The MCV delta check identified only half of presumed or confirmed misidentified specimens detected by the CCD and LCD. The superior performance of this approach has not yet been evaluated in a subsequent published work.

### Delta check thresholds

Delta check thresholds are determined using three strategies. Empiric or “best guess”, based on the experience of the laboratory director, or in consultation with local clinicians, or selected from the literature [6,8,14,15,17] was the earliest approach. Another strategy involved determining percentile-based thresholds from the frequency distribution of differences in selected populations [18,33,37–39]. This rested on the assumption that changes that occurred as a result of error, pathophysiology or clinical intervention represented unusual or statistically significant change. The third approach relied on calculated reference change values (RCV), also called the critical difference [40].

The common calculation for RCV uses the previous test result as a baseline and determines change over sequential samples; then RCV is compared to established statistical bounds. There are at least five different approaches to calculating the RCV, with each performing differently depending on whether the change indicates an increase or decrease; whether the homeostatic set-point (the denominator) is the previous result in a series, the mean of two results, or other estimate; and whether the distribution of results for the analyte is normal or log-normal on transformation [30]. Calculation of the RCV considers analytical imprecision for the analyte of interest and the intra-individual variability component of biological variability (Figure 2).

Biological variability has achieved prominence for development of quality specifications for test assays following the Milan consensus conference held in 2014 [41], and its application to calculation of RCV is supported in the recent Clinical Laboratory Standards Institute (CLSI) guideline, CLSI EP33, on delta checks [9]. Use of intra-individual variation to calculate RCV is based on two assumptions. The first is that analyte concentrations for an individual shows a normal distribution over time, or at least that it can be transformed mathematically to approximate a normal distribution for a patient at steady state. The second assumption, that intra-individual variation is similar across

individuals, allows the use of published databases of reference estimates for calculations. The common RCV calculation (Figure 2) includes terms for only analytical variation and intra-individual variation, as pre- and post-analytical variation add little to total variation relative to these two. The validity of these assumptions, and the reliability of biological variability databases [42,43] widely used to obtain estimates of intra-individual variation for analytes, have been challenged, especially concerning the usefulness of RCVs when applied to acutely and critically ill-patient populations [44,45]. These patient populations are also more likely to have delta criteria applied because of frequent monitoring and when short time intervals of a few days are used for delta checks. Because of medical intervention and rapid change in the clinical course of ill patients, calculated  $P\Delta$  values tend to be significantly larger. In defense of applying biological variability database information determined in healthy individuals to ill-patient populations, Ricos et al. [46] argued that for most analytes the intra-individual variability of patients with the disease is comparable to that of healthy individuals. While true in some patient subpopulations, relatively common treatment regimens like dialysis, fluid and electrolyte replacement, or blood transfusions remain as major causes of false-positive delta check results when thresholds are developed on the basis of RCVs. Furthermore, estimates of biological variability show significant differences across studies [47] due to study limitations. Acknowledging this, the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation and Task and Finish Group for the Biological Variation Database published a checklist for selecting robust estimates of biological variability [48].

The simplest approach to calculating the RCV limits involves using in-laboratory analytical imprecision ( $CV_A$ , analytical variability) for a test analyte and published estimates of intra-individual variability ( $CV_I$ ) obtained from available sources [40], where  $CV_I$  denotes intra-individual variability, the variability resulting from normal changes about the homeostatic set-point. These include changes occurring over periods of minutes to hours, to days, to even longer depending on the analyte.  $CV_A$  describes the analytical variability measured using internal QC procedures. Determination of RCV also depends on *a priori* selection of probability, typically limited to 95% or 99% of results. The inter-individual variability ( $CV_G$ ) defines the variability of homeostatic set points of the group of individuals in a population. The index of individuality, as defined by  $(CV_I^2 + CV_A^2)^{1/2}/CV_G$  [49], can be used as a tool for

selecting analytes for delta checks [25]. Analytes with low indices of individuality (typically  $<0.6$ ) perform better at detecting differences between individuals and allow more effective identification of misidentified specimens [9,25]. Hence, analytes like ALP, MCV, MCH, prothrombin time, and creatinine, all with indices of individuality below 0.4, perform better than most other analytes that have indices of individuality above 0.6. However, differences in assay imprecision in different laboratories coupled with differences in estimated biological variability across studies can lead to significantly different estimates of the index of individuality [47]. Thus, the use of this approach should be done with due diligence [48] and with consideration of the analytical performance differences that may exist in specific laboratories before generalizing delta check rules and criteria across different laboratories [25].

The distribution of changes in analyte results differ across different analytes and is often skewed in one direction. Lee et al. [38] determined delta check limits as 0.5th, 2.5th, 97.5th, and 99.5th percentile estimates for archived patient records in the LIS over a one-year interval, and compared them to delta check thresholds determined by the RCV approach. They noted that population-based percentile limits exceeded limits established by the RCV approach, and showed asymmetry in result change distribution for some tests. Asymmetric distributions of change occurred when intra-individual variability was large and the magnitude of the change for decreasing concentration was significantly different from the increasing concentration based on the analogous percentiles (2.5<sup>th</sup> versus 97.5<sup>th</sup>, for example). It was concluded that single  $PD\Delta$ s worked better for  $CV_I < 5\%$  (e.g. protein, albumin, sodium, potassium, and chloride) compared to tests where the  $CV_I$  was larger (e.g. ALT, AST, creatinine, and glucose). Because  $CV_I$  generally correlated with percentile-based limits based on population distribution, it was also concluded that  $CV_I$  was an appropriate parameter for determining delta check limits, especially when  $CV_I$  was  $<5\%$ . Other approaches to addressing asymmetry, taking into account differences in standard deviations when results were trending upward versus downward for calculating thresholds, were described by Jones [50]. Careful selection of analytes for delta checks as well as thresholds for upward versus downward change are required to optimize error detection and to minimize false alerts.

Delta check thresholds differ in different hospital settings when they are based on local hospital patient populations and when percentiles of change [25] are used. This strategy requires examination of paired patient data for

specific analytes, calculation of differences between sequential pairs among the paired data, typically within specified time limits between the serial values, and analysis of the frequency distribution of the differences. Using this strategy, delta check thresholds can be determined using parametric statistics or non-parametric estimates of percentiles, depending on whether there is a significant skewing of differences favoring the latter. As Sanchez-Navarro et al. [45] suggested, this approach may offer a simpler and more versatile alternative that addresses pathological and iatrogenic variability in the local population. Furthermore, change due to increasing results versus decreasing results could be more easily assessed. Although this pragmatic approach achieved in practice greater correlation with predicted positivity rates, it showed a loss of sensitivity for error detection when more extreme percentile estimates were used and when the frequently monitored very ill patients made up a greater proportion of data used to estimate delta check thresholds.

In a more sophisticated approach to determining delta check thresholds, Houwen and Durrin [27] used a version of the  $PD\Delta$  calculation and selected thresholds based on combinations of the linear discriminant and variable rate non-linear mathematical functions for various hematology parameters. The variable rate non-linear delta check thresholds were determined to address either the bell-shaped or the hyperbolic result distribution observed when percentage change results were plotted against the initial concentration of analyte. Special software was required to recalculate appropriate delta check thresholds for each new initial analyte concentration. The reported incidence of specimen misidentification errors was 0.03% by this strategy. There has been no subsequent follow-up work.

### *Partitioning of delta checks thresholds*

The effectiveness of delta checks for error detection can be improved through control of other factors that contribute to variability in a particular analyte, such as age, sex, outpatient versus inpatient, medical care unit or treatment modality, testing interval between consecutive results [51] or disease state [39]. To better leverage the significance of the time between serial values on the magnitude of change, Wheeler and Sheiner [18] adopted a categorical approach, applying different delta check thresholds over two different time intervals, 0.9–1.5 d and 1.5–2.5 d. No subsequent work using this approach has been reported in the literature. Lezotte and Grams [52] described different sex-specific “clinical delta ranges” for within-day and between-day. The calculated thresholds were set at 2 standard deviations from each side of the mean value for change from

healthy individuals and were partitioned according to sex. However, the numbers of subjects were too few to make conclusions about any differences between the two sexes. Based on the observation of an asymmetric increase in the degree of change in sequential values of hemoglobin A<sub>1C</sub> over time, Tran et al. [51] suggested different delta check thresholds over shorter terms of up to 60 d versus longer time periods of 60–120 d. As acknowledged in the report, the monitoring of hemoglobin A<sub>1C</sub> within a 90-d interval is rarely indicated.

Delta check criteria optimized for in-patients are broader and cover a shorter time interval than out-patients [39,53]. No study to date, however, has evaluated the practical value of applying different delta check criteria to in-patients versus out-patients. While delta checks are commonly used in autoverification routines, they reduce autoverification pass rates by generating alerts in frequently-tested patients, who are also generally those with wide fluctuations in test results that are likely due to treatment (e.g. dialysis, intravenous fluids, organ transplant, or surgery) or disease course (e.g. acute injury or recovery) [54]. In these situations, alerts are not helpful if the intent of the delta check is error detection [54]. No doubt, excluding such populations or using alternative criteria would reduce the incidence of false-positive alerts [55], but to what degree this would translate to a gain in error detection remains to be determined. Furthermore, programming software to exclude certain medical units or disease groups from delta check algorithms remains a challenge for many laboratories.

### **Rate-based delta checks**

Some analytes are more tightly held within narrow homeostatic limits than others. For other analytes, changes can be large over a short period of time, especially if associated with a medical intervention (e.g. changes in sodium during saline infusion), or during an acute pathological process (e.g. changes in cardiac troponins during acute coronary injury). In these cases, the rate of change in analyte concentrations can vary over the course of the intervention or process and can show a time dependency. Ignoring the time interval effect on an applied delta check threshold can lead to loss of information related to the nature of biological change for the analyte. Therefore, measuring change within constrained time limits or expressing change as a rate has the potential to provide more sensitive detection of clinically significant change or unusual change that may indicate an error [56]. However, in practice, the interval between samples is more often related to the acuity of

the patient condition, the test, and the location of the patient than with the previous test result [37,57].

In an attempt to improve error detection, yet reduce the workload associated with frequent delta checks, Lacher and Connolly [37] explored expressing ADA and PDA as rates. The purpose of the rate delta checks was to control for the time interval between analyses for serially monitored analytes that showed significant time dependency. This approach gave rise to absolute (ARDA) and percentage rate difference (PRDA) methods (Figure 2). The introduction of new delta check calculations for analytes such as sodium, MCV, or calcium, which are tightly regulated in time, created the challenge of determining which of the four calculation modes was best for which analytes. To address this issue, Kim et al. [58] recommended that the time dependence of the test item and the coefficient of variation (CV) of the ADA frequency distribution be considered. Tests with a large CV for the ADA, like liver enzymes, glucose, and urea, were expressed as PDA, while tests with a small CV for the ADA remained as ADA. The corresponding rate change delta checks were applied only to analytes showing significant time dependence for change. ARDA checks worked better for analytes where variation was small, but the goal was to determine significant change over a short time period, such as changes in cardiac troponin during the early stages of an acute coronary syndrome. In contrast, methods based on PRDA were better suited for analytes that showed large but slower changes, like uric acid, which could change by over 100 μmol/L but over the course of weeks to months [58]. Delta check thresholds were determined at 2.5% and 97.5% percentiles for changes within the population. Use of this strategy to select delta check calculations resulted in an increase in the overall positive predictive value (PPV) from 0.41% to 1.8% with a 74.3% decrease in workload for review. Park et al. [53] attempted to improve calculation mode selection criteria by also taking into account clinical characteristics of the patients, such as in-patients versus out-patients and those undergoing dialysis or receiving a transcatheter arterial chemo-embolization procedure. They proposed the use of the ratio of the “delta difference” relative to the difference of reference range of the test as a selection tool. These ratios were determined for increasing and decreasing changes separately but they used an approach similar to that of Kim et al. [58] and selected percentile-based limits. Although yielding similar preferred delta calculation and estimated thresholds for most analytes, there were differences in delta check thresholds and preferred calculations for in-patients versus out-patients. The selective use of delta check strategies based on time dependence has a modest impact on the effectiveness in detecting errors and reducing the number of samples that must be inspected.

The application of rate-based delta checks relies more heavily on accurate recording of specimen collection times.

### **Time interval for delta checks**

The time interval over which change is measured has a bearing on the likelihood for error detection. For time-dependent analytes, the longer the time intervals between serial tests, the greater the opportunity for the divergence of test results within the limits of biological variability or because of disease processes. In recognition of this, using the approach described by Park et al. [53], the median of the optimum time intervals for most routine chemistry tests ranged from 1–2 d. Sampson et al. [31] described a method for optimizing the selection of the time interval between tests for delta checks based on setting the specificity of error detection at 99% and using a time adjusted sensitivity score. High test scores were seen for various enzymes, but also creatinine, bilirubin, and urea, while glucose and electrolytes yielded low scores. Using data pairs of clinical chemistry analytes, optimum time intervals of between 2–5 d, depending on the analyte, were determined. This seems to reflect intervals used in contemporary practice among laboratories that responded to a recent College of American Pathologists (CAP) Q-probe study. In this study, Schiffman et al. [59] reported that the time interval for delta check comparisons was analyte dependent but differed across laboratories. The time intervals for delta checks were short, typically 3–7 d for chemistry and hematology tests. Interestingly none of the 46 participant laboratories indicated the use of rate of change calculations. Nevertheless, intervals of more than a week were not uncommon. In contrast, a recent study focusing on improving turnaround time (TAT) to the emergency department made use of extended windows of several years for application of delta checks based on RCVs [60]. Use of extended intervals beyond a few days may also be of value in the application of delta checks to case finding [61]. However, overall these represent exceptions to what is generally supported by the available literature and what is consistent with general practice. No recommendation on specific time interval is made in CLSI EP33 [9].

### **Standards and recommendations**

There is wide acceptance of delta checking to augment internal QC practices but minimal published guidance for evaluating the effectiveness of delta checking strategies. Although not specifically identifying delta checks by name,

the International Organization of Standardization [62], CAP [63, 64], and the Joint Commission International [65] all have issued standards or accreditation requirements that include delta checks, at least indirectly (Table 2). These standards assume a value in applying delta checks and promote their use as a part of good laboratory practice.

The 2005 International Consensus Group for Hematology Review [66] recommended the specific use of delta checks for MCV as a means of verifying sample integrity and misidentified specimens. This guideline provided criteria for the use of delta checking for white blood cells (WBCs) and platelets, alone or in combination with other findings, as criteria for follow-up blood smear review. The guideline fell short of indicating thresholds or a means of evaluating the effectiveness for any of the delta-checked analytes. There is no similar recommendation for any specific clinical chemistry, coagulation, or immunoassay test. In 2016, CLSI published the first guideline, CLSI EP33, on delta checks [9]. This guideline provided the framework for selection, implementation and evaluation of the performance of delta checks. These guidelines also recognized limitations in the literature surrounding the use of delta checks and acknowledged the subjective nature of the set-up of delta checks to a laboratory's overall quality assurance system. It recommended customizing delta checks based on the primary purpose, the prevalence of issues related to the purpose, and the patient population to which the delta checks were to be applied. It offered flexibility in selection of calculation modes, thresholds, time intervals, and patient populations.

CLSI EP33 recommends that laboratories assess the validity of current delta check parameters by reviewing retrospective data on flagged results to determine if errors are being identified efficiently [9] (Figure 3). The definition of efficiency remains open and leaves questions as to what positivity rates, and hence false-positive rates, should be tolerated when weighed against the costs of investigating delta check alerts, and the impact on patient care when the investigation leads to interruption of clinical staff to assist with resolving issues identified through the strategy. While the guideline provides sound recommendations to address the quality of delta check applications, application of the guideline will likely lead to diminished use of delta checks for error detection. Most published reports, especially of the commonly-used univariate delta check approaches, demonstrate low effectiveness for error detection. This lack of effectiveness is also likely to be observed by individual laboratories carrying out studies to evaluate delta check performance.

**Table 2.** Medical laboratory accreditation standards and clauses related to the Delta check.

Organization	Category/clause	Quoted description
College of American Pathologists, 2015 <sup>a</sup> . [64]	Quality management	<i>Automatic "traps" for improbable results. The system for detecting clerical errors, significant analytical errors, and unusual laboratory results should provide for the timely correction of errors, i.e. before results become available for clinical decision making.</i>
College of American Pathologists, 2012 <sup>b</sup> . [63]	GEN.43890 Autoverification Delta Checks	<i>The autoverification process includes all delta checks that the laboratory performs prior to the manual release of test results. NOTE: This requirement does not require delta checking for all autoverified results, but the laboratory's delta-checking procedures should be the same for manually released and autoverified test results.</i>
International Organization of Standards: ISO 15189:2012 <sup>c</sup> . [62]	5.9.2 Automated selection and reporting of results	<i>If the laboratory implements a system for automated selection and reporting of results, it shall establish a documented procedure to ensure that: a) the criteria for automated selection and reporting are defined, approved, readily available and understood by the staff. Items for consideration when implementing automated selection and reporting include changes from previous patient values that require review and values that require intervention by laboratory personnel, such as absurd, unlikely or critical values.</i>
Joint Commission International, 2017 <sup>d</sup> . [65]	Quality control processes: Standard 1.6	<i>The quality control processes of the laboratory include a process for a coordinated review of patient results, quality control results, and instrument function checks.</i>

<sup>a</sup>College of American Pathologists. Master All Common Checklist. 2015. p 17.

<sup>b</sup>College of American Pathologists. Laboratory General Checklist 2012. p 49.

<sup>c</sup>International Organization for Standardization. Medical laboratories: Requirements for quality and competence. ISO 15189:2012. p 35.

<sup>d</sup>Joint Commission International. Accreditation Standards for Laboratories, 3rd Edition, 2017. p 28.

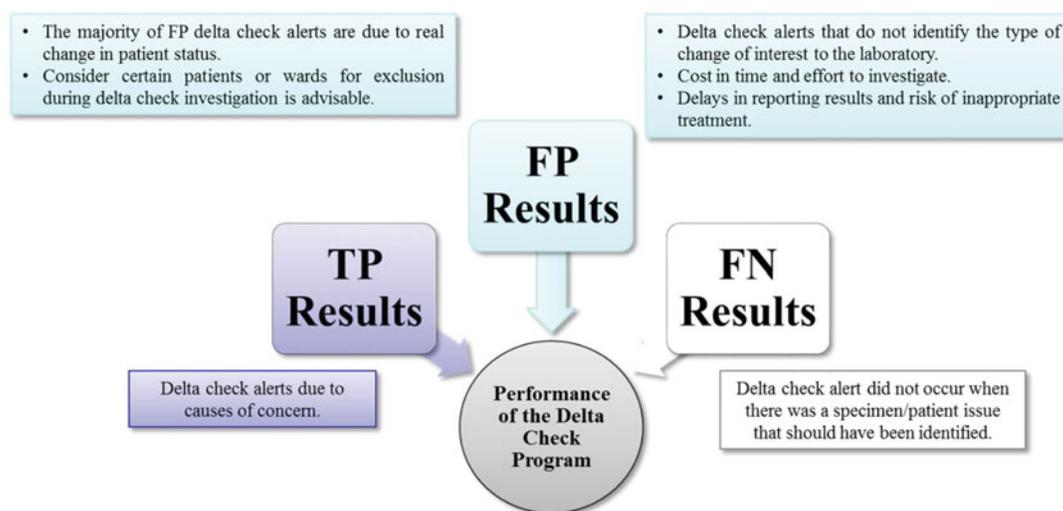
## Applications of delta checks

Delta checks are performed for two main purposes: 1) to identify specimen and collection issues affecting pre-analytical, analytical, or post-analytical quality that are not identified by QC methods; and 2) to assist clinicians with identifying disease and significant change in a patient's status. Apart from identifying true biological change in a patient, the delta check has been used for detecting clerical errors, specimen misidentification errors, contamination (e.g. intravenous fluids, anticoagulant), and mishandling, use of the wrong anticoagulant, presence of hemolysis and other interferences, and a host of problems with sample preparation, instruments, and reagents. In practice, true biological change in a

patient is the main explanation for delta check values exceeding thresholds, especially when thresholds are based on RCV (see section above). As there are many sources of biological variability, it is important that these be considered when developing alerts for clinicians about the need for medical intervention.

## Detection of error

From their origins, delta checks have been used to identify incorrect laboratory test results, and especially to identify mislabeled specimens, as part of the test result validation process. The effectiveness of the delta check concept when applied to mislabeling situations is



**Figure 3.** Monitored parameters recommended for evaluation of Delta Check schemes.

TP, true positives; FP, false positives; FN, false negatives.

based on two prerequisites: (1) test results on consecutive specimens from a given patient are likely to show less difference than results on two different patients and (2) large short-term change is unlikely in an individual. Both prerequisites are challenged in practice. Abrupt short-term changes that are sufficient to exceed typical delta check thresholds are relatively common in the acutely ill and/or patients undergoing certain treatments. Specimens that are misidentified in the pre-analytical phase are likely to go unnoticed if results between patient samples are not significantly different. Yet clinically significant change causing a delta check alert provides an opportunity for investigation and detection of misidentified samples, which, if left unchecked, could lead to inappropriate management. The ideal strategy detects all incorrect test results before release. Practically, the strategies adopted are based on the needs of specific clinical laboratories, their overall error rates, and availability of laboratory staff to review delta check results [45].

#### *Detection of specimen misidentification errors*

Most studies examining the effectiveness of delta checks use simulated data. This approach allows control over the “true error rate” as an experimental variable. Without this, the true error rate remains unknown because the overall effectiveness of all other existing follow-up strategies is unknown. In one of the earliest evaluations of delta checks for identifying mislabeled specimens, Sheiner et al. [19] compared three previously reported approaches [15,17,18]. Assuming a high mislabeling rate of about 1% and accepting a 5% false-positive rate, they concluded that about 5.5% of specimens would have a delta check alert, and about half of the mislabeled specimens would still go undetected. In a follow-up work, Wheeler and Sheiner [20] further evaluated the three delta check techniques by a rigorous multi-step algorithm to detect errors that involved immediate recollection and re-run, and by chart review to better determine performance in real-world situations. Significant false-positive rates in excess of 70% were observed when error affected 1.2–4.0% of samples. Even with high error rates, the time spent evaluating false positives represents major activity.

Advances made through the implementation of reliable, highly automated analyzer systems, standards of practice that require the use of two unique identifiers for patient identification and modern LIS systems have significantly reduced the opportunity for specimen misidentification errors and analytical errors. Nevertheless, specimen mislabeling remains one of the most common pre-analytical errors detected by delta checks, and

patient misidentification is the major cause, affecting almost 72% of adverse outcomes in-laboratory medicine [67]. Specimen misidentification occurs at a rate of 0.04% to as high as 1%, with a significant number of incidents leading to adverse patient outcomes [1,67–70]. In 1969, McSwiney and Woodrow [71] estimated error rates at 2.2%, but over half were clerical and less than 0.3% were due to mislabeling. A similar investigation by Grannis et al. [72] determined the incidence of error due to specimen misidentification at just under 1%, with the incidence of all other errors approximating 3%. In 1985, Tuckerman and Henderson [73] reporting on a computer-assisted approach to error detection that included delta checks and estimated total error rates of 3–5%. Later studies placed the incidence of specimen misidentification errors at <0.1% of samples [74,75], which suggested a low frequency of the type of error most commonly evaluated and most suited to identification by delta checks.

A large proportion of delta check alerts, especially those based on the use of conservative delta check criteria, detect a large number of samples not associated with any error. Addressing these alerts leads to significant extra work for laboratory staff, and also delays in reporting. In evaluating the current state of delta checking practices across laboratories, it has been shown that, among about 4.8% of delta check alerts associated with testing problems, the majority were caused by interferences and sample contamination, and only 0.3% of all delta check alerts were related to specimen misidentification [59]. In almost 95% of testing episodes in which a delta check alert occurred, no change in test results was made and no problem was identified. In cases where delta checks resulted in recollection of a sample (i.e. about 1.6% of all delta check alerts), the change persisted in over two-thirds of instances, indicating that, although some errors were detected, it was not without cost to patients and the system through unnecessary and difficult recollection, retesting, loss of significant blood volume in neonates, and potential delays in treatment or discharge.

A major issue concerning the use of delta checks is whether the number of errors detected is worth the investment of labor and other costs to investigating delta check alerts. While the costs of performing delta check calculations are negligible, the costs associated with investigation are significant. At one extreme, a delta check alert could lead to redrawing and rerunning a fresh sample. This comes at significant costs in time, labor, and quality of patient care [25]. Costs associated with evaluating false positives alone, however, does not negate the value of a prompt review of other

information and ruling out a problem, highlighted by a delta check alert, with specimen integrity or labeling. Discussion of unresolved delta check alerts with clinicians is also important so as to avoid unnecessary repeat sampling and testing. Repeat collection and testing may be warranted if a result cannot be explained on the basis of the clinical state.

Selection of analytes for delta checks and criteria have several notable limitations [25,59]. Typically the selection of analytes with low intra-individual variability is preferred [9]. While this reduces the opportunity for false alerts, it also reduces the potential power for error detection, especially if no test with an associated delta check alert is ordered. Secondly, the interval over which the delta check is applied is generally short, so only tests that are frequently repeated over short time intervals are useful. This limits its use mainly to hospital inpatients and other clinical settings in which frequent testing occurs [9].

An optimized system is one in which the number of false delta check alerts are minimized with acceptable rates of error detection. Based on results from the CAP Q-probe study [59], the median number of analytes with delta checks was 15, but it ranged from 6 to 32 analytes in different laboratories. The most frequently used analytes in delta checking schemes were MCV, hemoglobin, platelet count, sodium, calcium, potassium, creatinine, urea, albumin, and protein. Delta check alerts for sodium, potassium, calcium, magnesium, MCV, MCH, MCH concentration, red blood cell distribution width, hemoglobin or hematocrit, and platelet counts were associated with higher rates of error detection [59,76]. The least effective tests were total protein, bilirubin (total and direct), uric acid, aspartate aminotransferase, alkaline phosphatase, glucose, lactate dehydrogenase, gamma-glutamyl transferase, and cholesterol. Strathmann et al. [25] demonstrated that MCV had the highest PPV with the fewest false-positive results when they evaluated simulated delta check performance for parameters of the CBC and routine chemistries using  $\Delta$  within a 72-h window. A threshold of 3 fL for MCV yielded a sensitivity of 80% but about 2.3 positive results per 1000 samples. In contrast, the best-performing chemistry tests yielded a PPV of 0.44% (urea and creatinine) with 16 positive samples per thousand when the specimen misidentification rate (or true positive rate) was set to 1 in 500, the same conditions that yielded a PPV of 4.1% for MCV. Using a simulation modeling design, Ovens and Naugler [77] examined several different delta check strategies for specimen misidentification errors using sodium, potassium, chloride, total carbon dioxide and creatinine results. The

most commonly reported delta check strategies generally yielded sensitivities less than 30%, except for an  $\Delta$  strategy for creatinine that yielded a sensitivity of about 83% but a low specificity of around 50%. While at an error rate of 1%, the PPV was high at 58.2%, at a more realistic error rate of 0.1%, the PPV was 12.1%. This study concluded that delta check strategies were ineffective overall, and that no univariate strategy yielded high error detection without a significant false-positive rate.

The argument for delta checks rests in the potential to prevent serious harm from a misidentified specimen. From this perspective, the delta check alert represents an actionable finding through the investigation, cancellation of incorrect result, and specimen recollection that follow. However, little guidance is given in the literature on how to address delta check alerts. Some actions are listed in Table 1, and a possible workflow is provided in Figure 1. This list is not exhaustive but illustrates the potential complexity of the investigative process. Schifman et al. [59], and CLSI EP33 [9] also provide lists of possible actions in investigating delta check alerts, and, depending on the purpose of the delta check, the CLSI guideline offers an algorithm for follow-up of delta check alerts [9]. Depending on the frequency of test results held, the investigation process can be time-consuming and challenging, especially in the absence of clear protocols indicating what actions should be taken. This leads to variability in practice among laboratory staff involved in addressing delta check alerts. The practical response to a delta check alert has evolved. For example, Wheeler and Sheiner [20] suggested a preliminary check for transcription errors; if no errors were found, the alert was referred on to a pathologist for review. The outcome actions could variably involve repeating the test or calling the physician to determine if a change in the patient's condition or therapy could explain the change in results. The recent CAP Delta Check Q-Probe study [59], which examined actions taken in response to 6541 delta check alerts from 4505 testing episodes in 49 laboratories, also compared the frequency of various actions to a delta check alert. Based on medians for all laboratory actions, this study showed that 38% of actions involved clinical review, 25% involved retesting of the current sample, 20% involved rechecking of the current sample, 15% involved no action, 5% involved a check of the analytical system, and 2% indicated other actions. Retesting or rechecking the previous specimen represented a small proportion. Schifman et al. [59] recognized this as an area for improvement and recommended greater attention to investigating the previous sample because

almost one-quarter of testing problems originated from it. Except for the rare analytical error, rerunning the sample offers little to the investigation but adds time and costs to reporting critical results [78], and to reporting other unusual results where timely delivery may be important.

Although the majority of sample errors are identified before the sample result leaves the laboratory, the use of delta checks are not associated with significantly lower post-verification error rates nor with a lower number of errors detected pre-verification [69]. Misidentification errors occur as labeling errors on the primary specimen; registration and order entry errors; aliquot-related label errors, and result entry or other clerical errors [69]. These errors can result from failure to follow protocols to identify patients at the time of collection, whether through distraction, being hurried, inadequate staffing, inadequate training, fatigue, or movement of specimens away from the bedside for labeling, especially when multiple specimens are involved [22]. Most misidentification errors are identified by comparing patient identifiers and other information on blood tubes and requisitions. Other strategies for early detection of specimen misidentification errors include flagging new patients for further review, matching results against test requisitions, and matching computerized test orders against requisitions, and also through proactive approaches that require multiple identifiers on samples [69,70]. Errors are also prevented through the use of standard procedures and taking the "one patient at a time" approach to collection, labeling and submission to the laboratory; and by the use of barcode-based patient identification systems during specimen collection [70]. In terms of patient safety, modern protocols typically require the use of at least two patient identifiers, such as name, unique identification numbers and/or date of birth, to correctly identify patients for phlebotomy [79]. Finally, misidentified specimens can be identified by the caregiver through a report that contains an unordered test, or if results are substantively different from expected, or different from a recent previous result.

There are many potential applications of delta checks to qualitative information. Qualitative delta checks can be used for potential error detection by a change in class, type or location of monoclonal protein peaks on sequential serum protein electrophoresis gels [21]; by a change in blood group antigen or antibody status [22]; by the presence of any of a number of disease-specific and unmodifiable increases in biomarkers of biochemical genetic disease; by a change in viral antigen or antibody or vaccination status; by changes

in autoimmune-related antibody testing results; by the absence of a therapeutic drug being monitored; or by the sudden absence or presence of pathologically significant cells on automated hematology systems relative to recent samples [80]. Of course, applications of qualitative delta checks are most useful for investigations likely to be repeated, but over a significantly longer time interval than that used for most quantitative delta check applications. Qualitative delta checks, especially those involving the use of blood group data, can add value for investigating suspected delta check alerts from quantitative tests, and as a means of confirming internal consistency in a suspect sample. However, it should be borne in mind that error may not be the only reason for discordances between samples tested by a qualitative approach, although other reasons are less likely for samples collected over a short time interval of just a few days.

### *Detection of other specimen integrity issues*

Some of the frequently monitored analytes for which delta checks are used also generate alerts because of the presence of hemolysis icterus lipemia (HIL) interferences, contamination of the specimen by intravenous fluids [59], or the use of inappropriate anticoagulant during collection. Delta checks for commonly measured analytes like potassium are more likely to detect sample integrity issues, including hemolyzed or contaminated samples, than misidentified samples. Considering this, analytical systems that perform serum HIL indices checks largely make delta checks for hemolyzed samples redundant. This may not be true for all testing methods as sudden changes in potassium may be useful in identifying hemolysis in whole blood analyses performed by point of care testing devices. Contamination of samples during collection is a relatively common occurrence and a number of strategies have been proposed to detect such samples [81–83]. Preliminary work by Leen et al. [5], on an integrated discriminant system using multivariate statistical techniques, demonstrated greater sensitivity for overall error detection than other approaches; it detected more than twice as many errors as univariate delta checks [5]. Demirci et al. [82] later used an artificial neural network technique involving logistic regression methodology to develop decision algorithms that could identify contamination caused by taking blood from the same arm through which intravenous saline was delivered. These algorithms involved criteria where sodium, potassium and chloride exceeding certain limits ( $>160$  mmol/L,  $<3.5$  mmol/L,  $>110$  mmol/L, respectively) and/or that were based on increases in sodium and chloride coupled with decreases in glucose and

potassium over previous samples, and used delta check thresholds developed from RCVs for the corresponding tests. A similar strategy that was developed to identify contamination with dextrose intravenous fluid involved sodium, chloride, potassium, and glucose exceeding certain limits ( $<136$  mmol/L,  $<98$  mmol/L,  $>5.5$  mmol/L, and  $>6.1$  mmol/L, respectively) and/or negative changes in sodium and chloride with positive changes in potassium and glucose exceeding delta check thresholds. Similarly, by combining individual result limit criteria and delta check thresholds based on RCV, these authors also reported criteria for detecting fibrin and EDTA contamination. It is not known how well these multivariate approaches compare to univariate delta checks alone, but they have the potential to improve identification of specific contamination-related and other specimen integrity errors. The inclusion of delta checks as a tool and a part of the laboratory's quality management plan may be the only means of identifying some types of error (e.g. to identify contamination by intravenous fluids) occurring at earlier stages of the total testing process.

### ***Detection of analytical issues***

In addition to the use of delta checking to identify pre-analytical errors, other applications have been proposed for augmenting analytical quality assurance in the clinical laboratory. Recently, Cervinski and Cembrowski [84] and Jones [85] separately reported on the "average of deltas" concept for detection of a systematic analytical error. The average of deltas concept is similar to "average of normals" but instead involves computing averages of calculated change of sequential patient results for a certain test. The number of deltas in each point calculation is predetermined for the test and depends on the degree of shift requiring detection. Compared with the average of normals concept, the average of deltas tends to require fewer data points [85] and works better for frequently monitored analytes with a low index of individuality. Because of reliance on change over a short fixed interval, the strategy is a better fit for laboratories testing large inpatient populations.

### ***Detection of clinically significant change***

Determining clinically significant change based on empiric judgment can be unreliable [86]. Statistically, significant change may not necessarily be clinically significant, especially if it is expected from the patient status or treatment, or if it is not associated with a worsening clinical course. The selection of delta check thresholds is more relevant if, in addition to biological

and analytical variability, clinical significance is taken into account. The latter is less important if the primary purpose is error detection. Time-series analysis represents a robust statistical approach for evaluating and predicting further change but is difficult to implement because of the large number of serial values required. A minimum time series that is limited to two serial values, as is used for RCV and delta checks, is a practical approach that makes use of estimates of intra-individual variation and takes into account analytical variability. Yet in spite of its theoretical and practical value, few laboratories report RCVs and acceptable change limits, and few physicians use them. A number of studies have suggested values of RCVs applied to laboratory tests for detecting renal transplant rejection [87,88]; for detecting and monitoring heart failure [29,89,90]; in the diagnosis and monitoring of irritable bowel disease [91]; to assist with interpretation of tumor marker results [92–95]; for monitoring monoclonal gammopathies [96]; and to complement reference ranges as a tool for clinical decision making [97]. In spite of the importance of change in particular analytes to diagnosis and monitoring disease, there is a relative dearth of clear direction or published guidelines for the use of delta checks in clinical practices. There are a few exceptions, and in these cases the calculation methods used and the thresholds selected hold merit for improving clinical outcomes. One of the oldest applications of this concept involved the use of "PSA velocity", which is the  $AR\Delta$  ( $\geq 2$   $\mu\text{g/L/year}$ ) applied to serial prostate-specific antigen (PSA) results; it has been used as an alternative to threshold-based reference limits (typically  $>4$   $\mu\text{g/L}$ ). This approach is controversial and may provide little advantage over the threshold-based cutoffs [98].

Recent literature supports the use of delta check based electronic alerts to clinicians to detect pathological changes in patient status. In particular, correction of hyponatremia, identification of acute kidney injury, and the early rule in or rule out of acute myocardial events appear to hold clinical merit. When applied for these purposes, delta check information is reported with a result, is attached as a report comment or is addressed as a critical result that prompts an urgent call to a clinician with notification of the change.

The clinical need for this application of delta checks is illustrated by a recent case report by Chakraborty et al. [99]: too rapid correction of sodium in a hyponatremic patient led to osmotic demyelination syndrome (ODS). The brain normally adjusts to gains in water and hypotonicity by pushing out intracellular solutes, including sodium, potassium, and organic osmolytes, in order to establish a new equilibrium. Adaptation to the

new situation occurs in just a couple of days [100]. However, the osmotic effect of too rapid replacement of sodium pulls intracellular water out of the brain, resulting in the brain shrinking, and in dysarthria, ataxia, convulsions and eventually death. The risk of ODS increases in situations of severe hyponatremia (serum sodium  $<105$  mmol/L) or when sodium  $<120$  mmol/L is coupled with severe liver disease in patients receiving desmopressin therapy. As most cases of ODS are associated with rapid correction (exceeding 10–12 mmol/L/d or more than 18 mmol/L/48 h period [101]), it is recommended that sodium correction should not exceed 8 mmol/L/d in patients at high risk, or 12 mmol/L/d otherwise. These changes exceed ADA thresholds typically used for sodium and create the potential for confusion and miscommunication when a change exceeds both laboratory-based thresholds for error detection and that for clinical action. Chakraborty et al. [99] recommended the use of rate-based delta checks to detect changes in sodium at rates equal to or greater than 1 mmol/L/h for samples within a 6-h period, 0.7 mmol/L/h for a 6–12 h interval, and 0.5 mmol/L/h for a 12–24 h interval between serial values, and to alert clinicians of too rapid correction. Caution is noted in the use of rate-based approaches as the effects of biological variability and analytical imprecision may lead to large numbers of unnecessary alerts if the sampling interval is short. In connection with this, Tormey et al. [102] recommended due consideration of RCV thresholds for monitoring changes in sodium, and the use of direct ISE methodology or the same analyzer for serial measurements to improve the significance of the change in sodium levels over that due to biological and analytical variability.

Acute kidney injury (AKI) is a frequently-unrecognized condition that can lead to a rapid decline in renal function and death [103]. The main strategy for diagnosing AKI relies on the serial measurement of serum creatinine. Serum creatinine has a low index of individuality, and clinically significant changes can occur in an individual while results remain within the reference range. Rapid changes in serum creatinine, e.g.  $\geq 26.5$   $\mu\text{mol/L}$  within 48 h, or by  $\geq 150\%$  within 7 d, signifies AKI [104]. In a study by Garner et al. [105], 11.3% of admitted hospital in-patients were defined as having AKI using a combination of strategies. Using an ADA of  $>26$   $\mu\text{mol/L}$  between successive values and applied over a 30-d interval, the delta check detected 132 (98%) of the cases detected by the other strategies. The authors acknowledged that in situations of high baseline creatinine ( $>153$   $\mu\text{mol/L}$ ), a change in serum creatinine around the 26  $\mu\text{mol/L}$  threshold may represent change that is still

within the limits of the RCV for creatinine; and speculated that adoption of the RCV as a threshold for creatinine may provide a more effective means for identifying AKI. Flynn and Dawney [61], who used similar but less stringent criteria, showed that a 50% increase in creatinine to more than 50  $\mu\text{mol/L}$  within a 90-d period flagged only 0.75% of samples, but of these, 70% were cases of AKI. The authors alerted clinicians to cases by a comment on reports. Thomas et al. [106] also examined an AKI delta check alert system involving an increase in serum creatinine  $\geq 75\%$  over the previous result but with an unlimited time interval. The alert system included an outreach team of nephrologists and nurses calling patients when delta check thresholds were exceeded. Following implementation of the alert system, the study showed a small but statistically insignificant improvement in survival, which converged with the cumulative survival before implementation, after about 4 years. To provide a sensitive early alert system for AKI, Baron et al. [107] developed an algorithm based on the use of a 72-h tracked minimum result (a delta check technique where the minimum test result over a limited time interval, rather than the previous result, is used in calculating the ADA). For tracked minimum results  $\leq 176$   $\mu\text{mol/L}$ , an increase of  $>26$   $\mu\text{mol/L}$  was used as the alert criteria, and for results above 176  $\mu\text{mol/L}$ , an increase  $>44$   $\mu\text{mol/L}$  was used. This strategy demonstrated 94% sensitivity and 95% specificity for AKI and proved superior to simple delta checks in sensitivity.

Chest pain patients represent a significant proportion of patients presenting to emergency departments. The implementation of high sensitivity cardiac troponin assays provides an opportunity for effective use of delta checks to aid in the early rule in and rule out of acute myocardial syndromes and acute myocardial infarction (AMI) and for more timely triage. Jaffe et al. [108] described the potential value of delta checks, especially when addressing situations of chronic elevations in cardiac troponins, existing gaps, and a way forward for study and establishing a firmer place for delta checks in the diagnosis of acute coronary syndromes. The gains in analytical sensitivity of cardiac troponins allow quantification above the limit of detection in a majority of healthy individuals, but the improved precision for low values allows leveraging of changes in results to become part of diagnostic decision making when applied to very short time intervals. Many laboratories using high sensitivity cardiac troponin assays also use ADA or PDA criteria to support the clinical interpretation of short-term cardiac troponin changes. Early rule-out criteria are based on finding undetectable levels of troponin at presentation or lack of significant change in

cardiac troponin values based on serial sampling at time intervals as short as 1 h after the initial sample [109–113]. There are no universal thresholds for delta changes that can be used across all high sensitivity cardiac troponin I or cardiac troponin T assays, but both  $P\Delta$  and  $AD\Delta$  calculations are of value depending on the interval between collection of serial samples and whether the interval is 1, 2, or 3 h relative to the first sample. For example, for the high sensitivity cardiac troponin T assay, it has been recommended that a 50% change over a 3-h interval can be applied for rule-in when the initial value is near the upper limit of normal (99<sup>th</sup> percentile), but a smaller proportionate change of 20% would be significant if the initial value is above this limit [114]. As part of its rule-out algorithm, the European Society of Cardiology suggests an  $AD\Delta$  for high sensitivity cardiac troponin T of <3 ng/L on Roche Diagnostic platforms, or for high sensitivity cardiac troponin I of <2 ng/L on Abbott Architect analyzers, after allowing at least 1 h between serial measurements. This strategy has a high negative predictive value for AMI and can be used for early rule out of >50% of early presenters [109,111,112,115].

Significant heterogeneity in diagnostic thresholds and time intervals for clinical application of delta checks based on different assays, such as in the case of cardiac troponins, increases opportunity for error, or the blurring of diagnostic criteria, when different assays are used. These situations highlight the need for harmonization to ensure that not only are data from different laboratories comparable and consistent, but also that interpretative information like reference limits, critical values, and other alerts provide clinicians with the same information regardless of the laboratory doing the analysis. Recent calls for harmonization, not only of aspects related to analysis but also of the total testing process, and of all subdisciplines of laboratory medicine [116, 117] are also applicable to delta checks when applied to clinical purposes.

### Delta checks and autoverification

Rapid TAT for test results is required by clinical services like intensive care units and emergency departments. Overcrowding of emergency departments is a problem that spans international borders and affects both the quality of and access to healthcare. Increases in patients presenting to emergency departments negatively impact wait times and the quality of care, which compounds problems related to bed and staff shortages and inpatient boarding [118]. Excessive laboratory TATs have the potential to delay treatment and increase the length of stay [119]. Considerable effort and time are

required for result validation, yet the proportion of suspicious samples is small. To address these burdens, many laboratories have adopted autoverification protocols that use extreme or unusual result criteria, instrument flags, HIL indices, critical values, consistency checks including the use of bivariate ratios [32] and delta checks to filter test results that require closer attention. With well-developed strategies, most results pass through from analyzer to patient electronic record without interruption. This approach has the potential to improve workflow and error detection and reduce variability among laboratory staff involved in test result validation [120]. The delta check has been a common part of autoverification schemes from the earliest reports [28,120]. As a characteristic of design, delta check alerts, especially using conservative thresholds, are a frequent cause for tests being held for manual review [54,121–123]. Using a delta check threshold based on 99% RCV, Fraser et al. [124] achieved an autoverification rate of approximately 60% of sample reports. Delta check alerts are usually not the major factor impacting autoverification rates. The Krasowski et al. [125] strategy to high-performance autoverification of >99% of tests used few delta checks; those used were mainly conditional and were applied to specific result ranges (e.g. potassium <2.9 mmol/L or >6.2 mmol/L). Randell et al. [121,122] achieved 91–95% sample autoverification using less stringent delta check threshold limits that allowed 97.5–99.5% of test result changes over a 72-h period to pass. In this investigation, over 99% of delta check alerts showed concordance on repeat and <0.5% of samples held for inspection were found to have associated error. This result is qualitatively like that of Gruenberg et al. [126] who showed a low error yield after applying a 60% delta check threshold to 23,410 creatinine results. Out of 254 results held for review, only 1.2% of delta check alerts represented laboratory error. Two instances were identified as suspicious and were followed with repeat testing. The remaining cases represented true pathological changes, changes due to dialysis, or non-pathological changes. Despite relatively low rates of errors detected by repeat, some laboratories still recommend the use of delta checks to identify discordances due to any cause and repeat testing on the same sample, at least for hematology profiles [127]. A recent study by Fernandez-Grande et al. [60] described an autoverification process based largely on the use of delta check thresholds determined by 95% RCVs but applied over a 3-year period. While this study, with about 56% of tests autoverified, demonstrated improved verification time, TAT, and physician satisfaction following implementation, the effectiveness of this

strategy in identifying error was not indicated. When applied to error detection in autoverification schemes, delta check criteria increase TAT and consume significant amounts of laboratory staff time investigating specimens with alerts but yield few defects.

### Future directions

Despite the wide application of delta checks in laboratory medicine, there is still little data supporting its effectiveness. Most studies to date examining effectiveness make use of data modeling with an artificial introduction of error, typically specimen misidentification error, but few have examined performance in real life applications, which would require rigorous follow-up of outcomes for reported results. One recent attempt using a relatively high delta check threshold for serum creatinine showed a low error output [126]. More studies that examine other analytes, thresholds and calculation modes with rigorous follow-up are required. Some multivariate delta check approaches, such as the use of CCD or LDC by Miller [23], or various mixes of delta checks with other indices of specimen compromise, such as those described by Demirci et al., [82] hold promise of greater effectiveness in identifying specimen collection and processing error or contamination error. The effectiveness of these approaches requires further evaluation and validation under real-world conditions. Clearer guidance is required for laboratories to determine an acceptable alert rate to true positive error rate ratio for delta checks. The present literature shows controversy as to where this balance should lie.

This review describes three clinical situations, acute coronary syndromes, osmotic neurologic injury, and AKI, where delta checks could be used for identifying patients at high risk for disease. While much work is being done in examining strategies for use of cardiac troponins in acute coronary syndromes, studies are generally lacking that would allow determination and selection of the most effective approach to screen for AKI, or for risk for osmotic neurologic injury and complications by rapid correction of electrolyte imbalance. As the use of delta checks gains ground in application to clinical diagnostic problems and across different assay systems, harmonization of delta check criteria also becomes necessary [116,117].

### Conclusions

Contemporary use of delta checks shows differences in practice with little clarity concerning strategies that support the best outcomes. Several calculation types for

delta checks have been described, and some guidance on matching calculation modes with tests has been provided; however,  $AD\Delta$  and  $PD\Delta$  check calculations dominate [59]. Many analytes can be delta checked, but few are of significant value for identifying error. For identification of specimen misidentification errors, analytes that have a low index of individuality ( $<0.6$ ), and especially analytes that are frequently monitored, have the potential to perform better. However, there are few commonly measured analytes that meet these criteria, and most perform poorly when examined individually, collectively, and as part of automated autoverification routines. Essentially all univariate delta check strategies carry a low PPV. The larger the number of consecutive univariate delta check alerts on a sample, the higher the probability of a specimen misidentification error. The rate of change delta check calculations are preferred for tests showing significant time dependency, but they are used infrequently. Furthermore, the relatively short time interval over which delta checks are typically applied virtually excludes the vast majority of out-patients, leaving this patient group with little to no opportunity for detecting specimen misidentification errors following collection.

Delta checks using multiple univariate alert criteria increases specificity for error but at a cost to sensitivity. Multivariate approaches that consider the complex interrelationships between different analytes in health and disease show promise for improved delta check efficiency, but none of these have been independently evaluated nor adopted into practice.

Compliance with published guidelines requires setting up delta checks only when there is a clear purpose in mind. Several points are of note when considering if or how to apply the delta check concept within a laboratory. First, delta checking identifies errors that would otherwise go undetected by other means. The most significant and studied of these are specimen misidentification errors. Secondly, commonly-used delta check strategies have high positivity rates that on review are mainly false positives. Addressing delta check alerts is a challenge, because intensive review of results is time-consuming, and scrutiny of the previous sample is also required. Among the factors to be considered when resolving delta check alerts (Table 1 and Figure 1), the experience of the reviewer can be significant in the final determination. Use of standard procedures, checklists, or workflow algorithms to investigate with clear directions and customized actions by analyte are approaches that may reduce inter-observer variability. Moreover, a systematic approach that is incorporated into the laboratory's quality management plan and that aligns the application of delta checks to

identified pre-analytical problems, such as contamination by intravenous fluids, also contributes to establishing credibility for delta checks as a useful quality assurance tool. Most delta check alerts are generated by pathophysiological and iatrogenic causes, so attempts to resolve alerts must consider these by review of the sample and medical record information, and when necessary through discussion with clinicians. Other major causes arise from the presence of interference or contamination, or other specimen integrity issues. These can be evaluated through inspection of the sample tube and its labeling, or through a call to the medical unit. Finally, confirmation of specimen misidentification errors can be aided by investigating the sample for unique patient-specific characteristics such as blood type.

Several delta check strategies hold value for more efficient disease detection. Noted examples reviewed here include strategies for early identification of AKI, acute myocardial injury, or warning of neurologic injury due to too rapid correction of electrolyte abnormalities. Such strategies can be implemented with minimal expense to the laboratory and in the form of result comments linked to a delta check alert. Such approaches must be formalized after consultation with local clinicians because all such strategies denote clear expectation for clinical action once delta check thresholds are exceeded.

Both undetected errors and response to false-positive delta check alerts come at significant cost. Efforts to detect misidentified specimens in the laboratory cannot replace vigilance concerning appropriate patient and specimen identification at the time of blood collection. In the modern hospital laboratory setting, proactive approaches to prevent patient misidentification and specimen misidentification may provide a greater yield in error reduction than time-consuming error identification systems, including delta checks, in the laboratory. However, even the most well-intentioned efforts to prevent misidentification errors do not eliminate the need for tools to catch errors escaping these efforts. Hence, the key question remains as to how to best identify misidentification errors without an overwhelming number of false-positive delta check alerts. The common approach of using delta check thresholds taken from published sources or peer-reviewed literature does not address the significance of the location, disease state, or diagnostic interventions applicable to a given patient. For example, patients receiving dialysis are especially prone to drastic changes in urea, creatinine, and potassium levels. Yet these analytes are commonly selected for delta check processes.

Furthermore, patient populations used in studies may not resemble the patient population of all laboratories or even the same laboratory over time. These challenges support the importance of regularly verifying delta check parameters using data from the patient population served by the laboratory, and customizing thresholds based on the local population and local laboratory needs.

Despite the requirements and expectation of accreditation bodies, and the guidance provided by CLSI EP33, there remains significant controversy concerning the use of delta checks for error detection. Sound practice in the face of controversy requires a careful and balanced approach through selection, implementation, monitoring, review and update, based on local learning, on a pathway of continual improvement in error or disease detection but in a manner that minimizes the negative impacts of false-positive alerts. The optimum delta check approach is therefore a customized one for individual laboratories, where causes and outcomes of delta check alerts are frequently monitored and are a part of a process improvement program.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### References

- [1] Bonini P, Plebani M, Ceriotti F, et al. Errors in laboratory medicine. *Clin Chem*. 2002;48:691–698.
- [2] Snyderman LK, Harubin B, Kumar S, et al. Voluntary electronic reporting of laboratory errors: an analysis of 37,532 laboratory event reports from 30 health care organizations. *Am J Med Qual*. 2012;27:147–153.
- [3] Kazmierczak SC. Laboratory quality control: using patient data to assess analytical performance. *Clin Chem Lab Med*. 2003;41:617–627.
- [4] Straseski JA, Strathmann FG. Patient data algorithms. *Clin Lab Med*. 2013;33:147–160.
- [5] Leen TK, Erdogmus D, Kazmierczak S. Statistical error detection for clinical laboratory tests. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2012; 2720–2723.
- [6] Sher PP. An evaluation of the detection capacity of a computer-assisted real-time delta check system. *Clin Chem*. 1979;25:870–872.
- [7] Soloway HB, Kaplan HS. Mathematical implications of inadvertent sample swapping and the perception of apparent rates of gross error. *Clin Chem*. 1990;36: 1525–1526.
- [8] Lindberg DA. Collection, evaluation, and transmission of hospital laboratory data. *Methods Inf Med*. 1967;6: 97–107.

- [9] CLSI. Use of delta checks in the medical laboratory; Approved Guideline – First Edition. CLSI Document EP33. Wayne, PA: Clinical and Laboratory Standards Institute; 2016.
- [10] Karger AB. To Delta Check or Not to Delta Check? That is the question. *Jrnl App Lab Med.* 2017;1: 457–459.
- [11] Bologna LJ, Lind C, Riggs RC. Reducing major identification errors within a deployed phlebotomy process. *Clin Leadersh Manag Rev.* 2002;16:22–26.
- [12] Howanitz PJ, Renner SW, Walsh MK. Continuous wristband monitoring over 2 years decreases identification errors: a College of American Pathologists Q-Tracks Study. *Arch Pathol Lab Med.* 2002;126: 809–815.
- [13] Snyder SR, Favoretto AM, Derzon JH, et al. Effectiveness of barcoding for reducing patient specimen and laboratory testing identification errors: a laboratory medicine best practices systematic review and meta-analysis. *Clin Biochem.* 2012;45:988–998.
- [14] Nosanchuk JS, Gottmann AW. CUMS and delta checks: A systematic approach to quality control. *Am J Clin Pathol.* 1974;62:707–712.
- [15] Ladenson JH. Patients as their own controls: use of the computer to identify "laboratory error". *Clin Chem.* 1975;21:1648–1653.
- [16] Young DS, Harris EK, Cotlove E. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects: IV. Results of a study designed to eliminate long-term analytic deviations. *Clin Chem.* 1971;17:403–410.
- [17] Whitehurst P, Di Silvio TV, Boyadjian G. Evaluation of discrepancies in patients' results—an aspect of computer-assisted quality control. *Clin Chem.* 1975;21: 87–92.
- [18] Wheeler LA, Sheiner LB. Delta check tables for the Technicon SMA 6 continuous-flow analyzer. *Clin Chem.* 1977;23:216–219.
- [19] Sheiner LB, Wheeler LA, Moore JK. The performance of delta check methods. *Clin Chem.* 1979;25: 2034–2037.
- [20] Wheeler LA, Sheiner LB. A clinical evaluation of various delta check methods. *Clin Chem.* 1981;27:5–9.
- [21] Vyas SG, Singh G. Location of monoclonal peak as a tool in checking specimen integrity. *Pract Lab Med.* 2017;8:49–51.
- [22] Makroo RN, Bhatia A. Delta check for blood groups: a step ahead in blood safety. *Asian J Transfus Sci.* 2017;11:18–21.
- [23] Miller I. Development and evaluation of a logical delta check for identifying erroneous blood count results in a tertiary care hospital. *Arch Pathol Lab Med.* 2015;139:1042–1047.
- [24] Yamashita T, Ichihara K, Miyamoto A. A novel weighted cumulative delta-check method for highly sensitive detection of specimen mix-up in the clinical laboratory. *Clin Chem Lab Med.* 2013;51:781–789.
- [25] Strathmann FG, Baird GS, Hoffman NG. Simulations of delta check rule performance to detect specimen mislabeling using historical laboratory data. *Clin Chim Acta.* 2011;412:1973–1977.
- [26] Iizuka Y, Kume H, Kitamura M. Multivariate delta check method for detecting specimen mix-up. *Clin Chem.* 1982;28:2244–2248.
- [27] Houwen B, Duffin D. Delta checks for random error detection in hematology tests. *Lab Med.* 1989;20: 410–417.
- [28] Valdiguie PM, Rogari E, Philippe H. VALAB: expert system for validation of biochemical data. *Clin Chem.* 1992;38:83–87.
- [29] Fokkema MR, Herrmann Z, Muskiet FA, et al. Reference change values for brain natriuretic peptides revisited. *Clin Chem.* 2006;52:1602–1603.
- [30] Lund F, Petersen PH, Fraser CG, et al. Different percentages of false-positive results obtained using five methods for the calculation of reference change values based on simulated normal and ln-normal distributions of data. *Ann Clin Biochem.* 2016;53:692–698.
- [31] Sampson ML, Rehak NN, Sokoll LJ, et al. Time adjusted sensitivity analysis: a new statistical test for the optimization of delta check rules. *J Clin Ligand Assay.* 2007;30:44–54.
- [32] Kanno T, Takeuchi I, Sudo K. Bivariate ratio monitoring in clinical laboratories. *Am J Clin Pathol.* 1981;76: 782–787.
- [33] Lacher DA. The relationship between delta checks for selected chemistry tests. *Clin Chem.* 1990;36: 2134–2136.
- [34] Rheem I, Lee KN. The multi-item univariate delta check method: a new approach. *Stud Health Technol Inform.* 1998;52:859–863.
- [35] Dufour DR, Cruser DL, Buttolph T, et al. The clinical significance of delta checks. (Abstract 60.). *Am J Clin Path.* 1998;110:531.
- [36] Furutani H, Kitazoe Y, Yamamoto K, et al. Evaluation of the "Mahalanobis' generalized distance" method of quality control: monitoring system of multivariate data. *Am J Clin Pathol.* 1984;81:329–338.
- [37] Lacher DA, Connelly DP. Rate and delta checks compared for selected chemistry tests. *Clin Chem.* 1988; 34:1966–1970.
- [38] Lee J, Kim SY, Kwon HJ, et al. Usefulness of biological variation in the establishment of delta check limits. *Clin Chim Acta.* 2016;463:18–21.
- [39] Castro-Castro MJ, Dot-Bach D, Candás-Estébanez B, et al. Estimation of alert and change limits and its application in the plausibility control. *Accred Qual Assur.* 2011;16:643–647.
- [40] Fraser CG. Reference change values. *Clin Chem Lab Med.* 2012;50:807–812.
- [41] Sandberg S, Fraser CG, Horvath AR, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med.* 2015;53: 833–835.
- [42] Carobene A. Reliability of biological variation data available in an online database: need for improvement. *Clin Chem Lab Med.* 2015;53:871–877.
- [43] Braga F, Ferraro S, Ieva F, et al. A new robust statistical model for interpretation of differences in serial test results from an individual. *Clin Chem Lab Med.* 2015;53:815–822.

- [44] Ko DH, Park HI, Hyun J, et al. Utility of reference change values for delta check limits. *Am J Clin Pathol*. 2017;148:323–329.
- [45] Sánchez-Navarro L, Castro-Castro MJ, Dot-Bach D, et al. Estimation of alert and change limits of haematological quantities and its application in the plausibility control. *EJIFCC*. 2014;25:115–127.
- [46] Ricós C, Iglesias N, García-Lario JV, et al. Within-subject biological variation in disease: collated data and clinical consequences. *Ann Clin Biochem*. 2007;44:343–352.
- [47] Zhang P, Tang H, Chen K, et al. Biological variations of hematologic parameters determined by UniCel DxH 800 hematology analyzer. *Arch Pathol Lab Med*. 2013;137:1106–1110.
- [48] Aarsand AK, Fernandez-Calle P, Ricos C, et al. The biological variation data critical appraisal checklist: A standard for evaluating studies on biological variation. *Clin Chem*. 2018;64:501–514.
- [49] Harris EK. Effects of intra-and interindividual variation on the appropriate use of normal ranges. *Clin Chem*. 1974;20:1535–1542.
- [50] Jones GRD. Critical difference calculations revised: inclusion of variation in standard deviation with analyte concentration. *Ann Clin Biochem*. 2009;46:517–519.
- [51] Tran DV, Cembrowski GS, Lee T, et al. Application of 3-D  $\Delta$  check graphs to HbA1c quality control and HbA1c utilization. *Am J Clin Pathol*. 2008;130:292–298.
- [52] Lezotte D, Grams RR. Determining clinical significance in repeated laboratory measurements. *J Med Syst*. 1979;3:175–191.
- [53] Park SH, Kim SY, Lee W, et al. New decision criteria for selecting delta check methods based on the ratio of the delta difference to the width of the reference range can be generally applicable for each clinical chemistry test item. *Ann Lab Med*. 2012;32:345–354.
- [54] Rimac V, Lopic I, Kules K, et al. Implementation of the autovalidation algorithm for clinical chemistry testing in the laboratory information system. *Lab Med*. 2018;49:284–291.
- [55] Kampfrath T. Delta checks checkup: Optimizing cut-offs with lab-specific inputs. *Clinical Laboratory News* 2017, August 1. [cited 2018 Jul 13] Available from: <https://www.aacc.org/publications/cln/articles/2017/august/delta-checks-checkup>
- [56] Min WK, Lee JO. The development of online delta check system using the workstation. *J Clin Pathol Qual Control*. 1996;18:133–142.
- [57] Valenstein PN, Cadoff E, Burke MD. Computer based audit of potassium monitoring. *Inform Pathol*. 1987;2:102–106.
- [58] Kim JW, Kim JQ, Kim SI. Differential application of rate and delta check on selected clinical chemistry tests. *J Korean Med Sci*. 1990;5:189–195.
- [59] Schiffman RB, Talbert M, Souers RJ. Delta check practices and outcomes: a Q-probes study involving 49 health care facilities and 6541 delta check alerts. *Arch Pathol Lab Med*. 2017;141:813–823.
- [60] Fernández-Grande E, Varela-Rodríguez C, Sáenz-Mateos L, et al. Impact of reference change value (RCV) based autoverification on turnaround time and physician satisfaction. *Biochem Med (Zagreb)*. 2017;27:342–349.
- [61] Flynn N, Dawney A. A simple electronic alert for acute kidney injury. *Ann Clin Biochem*. 2015;52:206–212.
- [62] International Organization for Standardization (2012). *Medical laboratories: requirements for quality and competence*. ISO Standard No 15189.
- [63] College of American Pathologists (CAP). *Laboratory General Checklist*. CAP Accreditation Program. Northfield (IL). CAP; 2012.
- [64] College of American Pathologists (CAP). *Master All Common Checklist*. CAP Accreditation Program. College of American Pathologists, Northfield (IL). CAP; 2015.
- [65] Joint Commission International. *Accreditation Standards for Laboratories*. Oak Brook (IL). Joint Commission International, 2017. [cited 2018 Jul 12] Available from: [https://www.jointcommissioninternational.org/assets/3/7/JCI\\_Standards\\_for\\_Laboratories\\_STANDARDS-ONLY.pdf](https://www.jointcommissioninternational.org/assets/3/7/JCI_Standards_for_Laboratories_STANDARDS-ONLY.pdf)
- [66] Barnes PW, McFadden SL, Machin SJ, et al. The International Consensus Group for Hematology Review: suggested criteria for action following automated CBC and WBC differential analysis. *Lab Hematol*. 2005;11:83–90.
- [67] Dunn EJ, Moga PJ. Patient misidentification in laboratory medicine: a qualitative analysis of 227 root cause analysis reports in the Veterans Health Administration. *Arch Pathol Lab Med*. 2010;134:244–255.
- [68] Carraro P, Zago T, Plebani M. Exploring the initial steps of the testing process: frequency and nature of pre-preanalytic errors. *Clin Chem*. 2012;58:638–642.
- [69] Valenstein PN, Raab SS, Walsh MK. Identification errors involving clinical laboratories: a College of American Pathologists Q-Probes study of patient and specimen identification errors at 120 institutions. *Arch Pathol Lab Med*. 2006;130:1106–1113.
- [70] Lippi G, Blanckaert N, Bonini P, et al. Causes, consequences, detection, and prevention of identification errors in laboratory diagnostics. *Clin Chem Lab Med*. 2009;47:143–153.
- [71] McSwiney RR, Woodrow DA. Types of errors within a clinical laboratory. *J Med Lab Technol*. 1969;26:340–346.
- [72] Grannis GF, Grümer HD, Lott JA, et al. Proficiency evaluation of clinical chemistry laboratories. *Clin Chem*. 1972;18:222–236.
- [73] Tuckerman JF, Henderson AR. The clinical biochemistry laboratory computer system and result entry: validation of analytical results. *Comput Meth Prog Bio*. 1985;20:103–116.
- [74] Wagar EA, Tamashiro L, Yasin B, et al. Patient safety in the clinical laboratory: a longitudinal analysis of specimen identification errors. *Arch Pathol Lab Med*. 2006;130:1662–1668.
- [75] Murphy MF, Stearn BE, Dzik WH. Current performance of patient sample collection in the UK. *Transfus Med*. 2004;14:113–121.
- [76] Ford A. Delta checks as a safety net: how used, how useful. *CAP Today*. 2015;29:22–24.

- [77] Ovens K, Naugler C. How useful are delta checks in the 21st century? A stochastic-dynamic model of specimen mix-up and detection. *J Pathol Inform.* 2012;3:5. doi: [10.4103/2153-3539.93402](https://doi.org/10.4103/2153-3539.93402).
- [78] Deetz CO, Nolan DK, Scott MG. An examination of the usefulness of repeat testing practices in a large hospital clinical chemistry laboratory. *Am J Clin Pathol.* 2012;137:20–25.
- [79] Joint Commission. National Patient Safety Goals. Oak Brook (IL). The Joint Commission. [cited 2018 Jul 12]. Available from: [https://www.jointcommission.org/standards\\_information/npsgs.aspx](https://www.jointcommission.org/standards_information/npsgs.aspx).
- [80] Lantis KL, Harris RJ, Davis G, et al. Elimination of instrument-driven reflex manual differential leukocyte counts: optimization of manual blood smear review criteria in a high-volume automated hematology laboratory. *Am J Clin Pathol.* 2003;119:656–662.
- [81] Mancuso JV, Short P, Manilich EA. Learning to classify contaminated blood tests: a Neural Approach. Workshop track - ICLR 2018. [cited 2018 Jul 12] Available from: <https://openreview.net/pdf?id=S1vuwKJPz>.
- [82] Demirci F, Akan P, Kume T, et al. Artificial neural network approach in laboratory test reporting: learning algorithms. *Am J Clin Pathol.* 2016;146:227–237.
- [83] Baron JM, Mermel C, Lewandowski KB, et al. Detection of preanalytic laboratory testing errors using a statistically guided protocol. *Am J Clin Pathol.* 2012;138:406–413.
- [84] Cervinski M, Cembrowski G. Detection of systematic error using the average of Deltas. *Am J Clin Pathol.* 2017;147:S162–S163.
- [85] Jones GR. Average of delta: a new quality control tool for clinical laboratories. *Ann Clin Biochem.* 2016; 53:133–140.
- [86] Queraltó JM. Intraindividual reference values. *Clin Chem Lab Med.* 2004;42:765–777.
- [87] Biosca C, Ricós C, Lauzurica R, et al. Reference change value concept combining two delta values to predict crises in renal posttransplantation. *Clin Chem.* 2001;47:2146–2148.
- [88] Ozturk OG, Paydas S, Balal M, et al. Biological variations of some analytes in renal posttransplant patients: A different way to assess routine parameters. *J Clin Lab Anal.* 2013;27:438–443.
- [89] Bruins S, Fokkema MR, Römer JW, et al. High intraindividual variation of B-type natriuretic peptide (BNP) and amino-terminal proBNP in patients with stable chronic heart failure. *Clin Chem.* 2004;50:2052–2058.
- [90] Meijers WC, van der Velde AR, Muller Kobold AC, et al. Variability of biomarkers in patients with chronic heart failure and healthy controls. *Eur J Heart Fail.* 2017;19:357–336.
- [91] Padoan A, D’Incà R, Scapellato ML, et al. Improving IBD diagnosis and monitoring by understanding pre-analytical, analytical and biological fecal calprotectin variability. *Clin Chem Lab Med.* 2018;56:1926–1935.
- [92] Sölétormos G, Schiøler V, Nielsen D, et al. Interpretation of results for tumor markers on the basis of analytical imprecision and biological variation. *Clin Chem.* 1993;39:2077–2083.
- [93] Erden G, Barazi AO, Tezcan G, et al. Biological variation and reference change values of CA 19-9, CEA, AFP in serum of healthy individuals. *Scand J Clin Lab Invest.* 2008;68:212–218.
- [94] Sölétormos G, Duffy MJ, Hayes DF, et al. Design of tumor biomarker-monitoring trials: a proposal by the European group on tumor markers. *Clin Chem.* 2013;59:52–59.
- [95] Qi Z, Zhang L, Chen Y, et al. Biological variations of seven tumor markers. *Clin Chim Acta.* 2015;450:233–236.
- [96] Salamatmanesh M, McCudden CR, McCurdy A, et al. Monoclonal protein reference change value as determined by gel-based serum protein electrophoresis. *Clin Biochem.* 2018;51:61–65.
- [97] Bugdayci G, Oguzman H, Arattan HY, et al. The use of reference change values in clinical laboratories. *Clin Lab.* 2015;61:251–257.
- [98] Vickers AJ, Savage C, O’Brien MF, et al. Systematic review of pretreatment prostate-specific antigen velocity and doubling time as predictors for prostate cancer. *Jco.* 2009;27:398–403.
- [99] Chakraborty S, Bhattacharya C, Das S, et al. Delta check: a must in the management of hyponatremia. *Clin Chem Lab Med.* 2013;51:e161–e162.
- [100] Verbalis JG, Gullans SR. Hyponatremia causes large sustained reductions in brain content of multiple organic osmolytes in rats. *Brain Res.* 1991;567: 274–282.
- [101] Sterns RH. Overview of the treatment of hyponatremia in adults. UpToDate; 2018. [cited 2018 Oct 5]. Available from: <https://www.uptodate.com/contents/overview-of-the-treatment-of-hyponatremia-in-adults>.
- [102] Tormey WP, Carney M, Cuesta M, et al. Reference change values for sodium are ignored by the American and European treatment guidelines for hyponatremia. *Clin Chem.* 2015;61:1430–1432.
- [103] Stewart J, Findlay G, Smith N, et al. Adding insult to injury: a review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury (acute renal failure). A report by the National Confidential Enquiry into Patient Outcome and Death. London (UK). NCEPOD;2009. [cited 2018 Oct 5]. Available from <https://www.ncepod.org.uk/2009aki.html>
- [104] Kidney Disease: Improving Global Outcomes (KDIGO). Clinical practice guideline update for the diagnosis, evaluation, prevention, and treatment of chronic kidney disease – mineral and bone disorder (CKD-MBD). *Kidney Int Suppl.* 2017;7:1–59.
- [105] Garner AE, Lewington AJ, Barth JH. Detection of patients with acute kidney injury by the clinical laboratory using rises in serum creatinine: comparison of proposed definitions and a laboratory delta check. *Ann Clin Biochem.* 2012;49:59–62.
- [106] Thomas ME, Sitch A, Baharani J, et al. Earlier intervention for acute kidney injury: evaluation of an outreach service and a long-term follow-up. *Nephrol Dial Transpl.* 2015;30:239–244.
- [107] Baron JM, Cheng XS, Bazari H, et al. Enhanced creatinine and estimated glomerular filtration rate reporting to facilitate detection of acute kidney injury. *Am J Clin Pathol.* 2015;143:42–49.

- [108] Jaffe AS, Moeckel M, Giannitsis E, et al. In search for the Holy Grail: suggestions for studies to define delta changes to diagnose or exclude acute myocardial infarction: a position paper from the study group on biomarkers of the Acute Cardiovascular Care Association. *Eur Heart J Acute Cardiovas Care*. 2014; 3:313–316.
- [109] Pickering JW, Greenslade JH, Cullen L, et al. Assessment of the European Society of Cardiology 0-Hour/1-hour algorithm to rule-out and rule-in acute myocardial infarction clinical perspective. *Circulation*. 2016;134:1532–1541.
- [110] Morrow DA. Clinician's guide to early rule-out strategies with high-sensitivity cardiac troponin. *Circulation*. 2017;135:1612–1616.
- [111] Jaeger C, Wildi K, Twerenbold R, et al. One-hour rule-in and rule-out of acute myocardial infarction using high-sensitivity cardiac troponin I. *Am Heart J*. 2016; 171:92–102.
- [112] Twerenbold R, Boeddinghaus J, Nestelberger T, et al. Clinical use of high-sensitivity cardiac troponin in patients with suspected myocardial infarction. *J Am Coll Cardiol*. 2017;70:996–1012.
- [113] Apple FS, Sandoval Y, Jaffe AS, et al. Cardiac troponin assays: guide to understanding analytical characteristics and their impact on clinical care. *Clin Chem*. 2017;63:73–81.
- [114] Thygesen K, Mair J, Giannitsis E, et al. How to use high-sensitivity cardiac troponins in acute cardiac care. *Eur Heart J*. 2012;33:2252–2257.
- [115] Roffi M, Patrono C, Collet JP, et al. 2015 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: task force for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J*. 2016;37:267–315.
- [116] Plebani M. Harmonization in laboratory medicine: the complete picture. *Clin Chem Lab Med*. 2013;51: 741–751.
- [117] Plebani M. Harmonization in laboratory medicine: more than clinical chemistry? *Clin Chem Lab Med*. 2018;56:1579–1586.
- [118] Hoot NR, Aronsky D. Systematic review of emergency department crowding: causes, effects, and solutions. *Ann Emerg Med*. 2008;52:126–136.
- [119] Steindel SJ, Howanitz PJ. Physician satisfaction and emergency department laboratory test turnaround time: observations based on College of American Pathologists Q-Probes studies. *Arch Pathol Lab Med*. 2001;125:863–871.
- [120] Oosterhuis WP, Ulenkate HJ, Goldschmidt HM. Evaluation of LabRespond, a new automated validation system for clinical laboratory test results. *Clin Chem*. 2000;46:1811–1817.
- [121] Randell EW, Short G, Lee N, et al. Autoverification process improvement by Six Sigma approach: clinical chemistry & immunoassay. *Clin Biochem*. 2018;55: 42–48.
- [122] Randell EW, Short G, Lee N, et al. Strategy for 90% autoverification of clinical chemistry and immunoassay test results using six sigma process improvement. *Data Brief*. 2018;18:1740–1749.
- [123] Feitosa MS, Bucker DH, Santos SME, et al. Implementation of criteria for automatic release of clinical chemistry test results in a laboratory at an academic public hospital. *J Bras Patol Med Lab*. 2016;52:149–156.
- [124] Fraser CG, Stevenson HP, Kennedy IM. Biological variation data are necessary prerequisites for objective autoverification of clinical laboratory data. *Accred Qual Assur*. 2002;7:455–460.
- [125] Krasowski MD, Davis SR, Drees D, et al. Autoverification in a core clinical chemistry laboratory at an academic medical center. *J Pathol Inform*. 2014;5:13. Available from: <http://www.jpathinformatics.org/text.asp?2014/5/1/13/129450>
- [126] Gruenberg JM, Stein TA, Karger AB. Determining the utility of creatinine delta checks: a large retrospective analysis. *Clin Biochem*. 2018;53:139–142.
- [127] Lyashchenko AK, Wontakal SN, Reynafarje GM, et al. Trust but verify: repeat CBC sample measurements help identify automated hematology analyzer errors. *Am J Clin Pathol*. 2018;149:S172–S173.