



## REVIEW ARTICLE

# Diagnostic methods for neonatal hyperbilirubinemia: benefits, limitations, requirements, and novel developments

Christian V. Hulzebos<sup>1</sup>, Libor Vitek<sup>2</sup>, Carlos D. Coda Zabetta<sup>3</sup>, Aleš Dvořák<sup>2</sup>, Paul Schenk<sup>4</sup>, Eline A. E. van der Hagen<sup>5,6</sup>, Christa Cobbaert<sup>4</sup> and Claudio Tiribelli<sup>7</sup>

Invasive bilirubin measurements remain the gold standard for the diagnosis and treatment of infants with severe neonatal hyperbilirubinemia. The present paper describes different methods currently available to assess hyperbilirubinemia in newborn infants. Novel point-of-care bilirubin measurement methods, such as the BiliSpec and the Bilistick, would benefit many newborn infants, especially in low-income and middle-income countries where the access to costly multi-analyzer in vitro diagnostic instruments is limited. Total serum bilirubin test results should be accurate within permissible limits of measurement uncertainty to be fit for clinical purposes. This implies correct implementation of internationally endorsed reference measurement systems as well as participation in external quality assessment programs. Novel analytic methods may, apart from bilirubin, include the determination of bilirubin photoisomers and bilirubin oxidation products in blood and even in other biological matrices.

*Pediatric Research* (2021) 90:277–283; <https://doi.org/10.1038/s41390-021-01546-y>

**IMPACT:**

- Key message: Bilirubin measurements in blood remain the gold standard for diagnosis and treatment of severe neonatal hyperbilirubinemia (SNH). External quality assessment (EQA) plays an important role in revealing inaccuracies in diagnostic bilirubin measurements.
- What does this article add to the existing literature? We provide analytic performance data on total serum bilirubin (TSB) as measured during recent EQA surveys. We review novel diagnostic point-of-care (POC) bilirubin measurement methods and analytic methods for determining bilirubin levels in biological matrices other than blood.
- Impact: Manufacturers should make TSB test results traceable to the internationally endorsed total bilirubin reference measurement system and should ensure permissible limits of measurement uncertainty.

**INTRODUCTION**

International guidelines for the management of neonates with unconjugated hyperbilirubinemia include treatment thresholds that are based on total serum bilirubin (TSB) concentrations. Bilirubin measurements are key to the management of neonatal jaundice. An essential prerequisite for bilirubin measurements is that they are accurate and precise to manage jaundiced newborn infants appropriately. Over the past decades, bilirubin measurement for severe neonatal hyperbilirubinemia (SNH) identification has constituted a major challenge. A vast number of methods to determine bilirubin in human serum have been developed since it was first reported in 1858 by Frerichs (Gmelin reaction).<sup>1,2</sup> Van den Bergh and Snapper described their important diazo reaction in 1913<sup>3</sup> and the colorimetric determination by Malloy and Evelyn was published in 1937.<sup>4</sup> This was followed by Jendrassik and Grof who refined and modified the diazo reactions in 1938.<sup>5</sup> For a thorough understanding in the subsequent paragraphs of the so-called science of measurement and of terms such as TSB accuracy, trueness, metrological traceability, and measurement uncertainty, we

refer to the definitions and terminology provided by the International Vocabulary of Metrology.<sup>6</sup>

**DIAGNOSTIC METHODS—INVASIVE TOTAL SERUM BILIRUBIN MEASUREMENTS****Multiparameter instruments**

Routine laboratory bilirubin measurement is commonly performed with multichannel instruments that provide values for conjugated (direct), unconjugated (indirect), and TSB concentrations. They usually base the direct determination of bilirubin on diazo (Jendrassik Grof) and vanadate oxidase chemical reactions, or on variants.<sup>7</sup> On account of the high costs involved and the requirement of specialized personnel, these instruments are used mainly in laboratories of large hospitals. Even though the concentrations supplied by these instruments were considered as the clinical “reference” for TSB, inconsistencies amongst the different laboratory methods have been observed for decades.<sup>8,9</sup> In 2010, in the Netherlands, Cobbaert and colleagues analyzed the accuracy of TSB levels nationwide as measured by the most

<sup>1</sup>Beatrix Children's Hospital, University Medical Center Groningen, Groningen, The Netherlands; <sup>2</sup>Institute of Medical Biochemistry and Laboratory Diagnostics, and 4th Department of Internal Medicine, 1st Faculty of Medicine, Charles University, Prague, Czech Republic; <sup>3</sup>Bilimetrix s.r.l., Trieste, Italy; <sup>4</sup>Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands; <sup>5</sup>Dutch Foundation for Quality Assessment in Medical Laboratories (SKML), Nijmegen, The Netherlands; <sup>6</sup>MCA Laboratory, Queen Beatrix Hospital, Winterswijk, The Netherlands and <sup>7</sup>Italian Liver Foundation, Trieste, Italy  
Correspondence: Christian V. Hulzebos (c.v.hulzebos@umcg.nl)

Received: 6 August 2020 Revised: 20 February 2021 Accepted: 10 April 2021  
Published online: 4 May 2021

commonly used multiparameter instruments and in vitro diagnostic devices (IVDs).<sup>10</sup> In this study, a pooled human serum was supplemented with unconjugated bilirubin to obtain target values of 26.7 µmol/L (95% confidence interval (CI) range of 26.1–27.3 µmol/L), and 68.7 µmol/L (95% CI range of 67.2–70.2 µmol/L), as assigned with the Doumas reference measurement procedure (RMP) of the Joint Committee for Traceability in Laboratory Medicine (JCTLM) listed reference laboratory in Hannover, Germany. The Doumas RMP can be considered as the gold standard. The two value-assigned specimens were measured in 183 medical laboratories and in-house by IVD manufacturers using their respective routine methods. This procedure allows the accuracy of results produced by Dutch medical laboratories and IVD manufacturers to be assessed and to be compared to the target values assigned with an internationally recognized reference method. The interlaboratory variability and inaccuracy of TSB levels observed among manufacturers and individual laboratories were substantial. This indicates that the concept of metrological traceability, which leads to exchangeable TSB results, was not uniformly adopted in the commercially available IVDs. Also, in-house results produced by individual IVD manufacturers demonstrated significant differences. Similar discrepancies were observed by Greene and colleagues when comparing the performance of a Beckman AU680 instrument versus Ortho-Clinical-Diagnostics' VITROS device.<sup>11</sup>

It is common in neonatal intensive care units (NICUs) to find blood gas analyzer instruments like Radiometer ABL models or the RAPIDPoint models produced by Siemens HealthCare Diagnostics Inc.<sup>12</sup> Blood gas analyzer instruments estimate TSB indirectly by using whole blood co-oximetry, whereby whole blood is hemolyzed and serum equivalent bilirubin concentration calculated. Previously, the comparability of blood gas analysis-derived TSB levels and TSB levels measured on routine laboratory instruments was assessed. In 2018, Lano and colleagues reported a comparative analysis of neonatal TSB levels by whole blood co-oximetry (Radiometer® ABL90) against plasma bilirubin methods (Roche Diagnostics Cobas® C-601 and Ortho Clinical Diagnostics VITROS® 350).<sup>13</sup> Results showed good correlation in comparison to the Roche plasma diazo method, with a mean bias of -1.0 µmol/L across the bilirubin range examined and a 95% confidence interval range of -20.00 to 19.00 µmol/L. However, a statistically significant underestimation was found against the VITROS® 350 machine with a mean bias of -4.4 µmol/L over the bilirubin range examined and a 95% confidence interval of -29.90 to 21.10 µmol/L. Performance of the GEM Premier 4000® blood gas analyzer (Instrumentation Laboratory, Bedford, MA) was also assessed showing a wide range of differences compared to VITROS, with a negative bias at low concentrations of bilirubin and a positive bias at higher concentrations. Moreover, hemoglobin concentration and hemolysis affected the recovery of the GEM blood gas analyzer results.<sup>14</sup>

#### Bench-top bilirubinometers

Other commonly used instruments for TSB measurement are bench-top bilirubinometers. Based on direct spectrophotometry, these are simple and rapid alternatives for assessing TSB that

require a minimal sample for analysis. In practice, an undiluted serum is used to measure the bilirubin absorbance (at 454 nm) and hemoglobin (at 454 and 528 nm). Hence, subtracting the absorbance at 528 nm from that at 454 nm yields a value that can be attributed largely to bilirubin. The prevalence of the other forms of bilirubin and chromophores in older children and adults limits the use of this technique to neonates younger than 2–3 weeks of age.<sup>15</sup> Two types of direct spectrometry instruments are available commercially: those using sample cuvettes such as UNISTAT (Reichert Technologies, USA), and those using hematocrit capillary tubes like One Beam (Ginevri, Italy). Although there are many options available commercially of instruments based on direct spectrophotometry, validation studies of this method are limited. The advantage of bilirubinometers is the short turnaround times for results as shown in Table 1. The requirement of sample processing, however, and the need for additional instrumentation such as centrifuge and trained laboratory personnel limits the use of this method for TSB determination substantially.<sup>16,17</sup>

#### Hand-held point-of-care (POC) bilirubin instruments

Neonatal jaundice identification has always posed a challenge, mainly in LMICs.<sup>18,19</sup> Over the last years, with the advancement of technology, different solutions have emerged for SNH screening. In 2017, Keahey and colleagues reported validation data of a new screening device under development, the BiliSpec, in 94 blood samples of 67 newborn infants between the day of birth and 24 days.<sup>20</sup> This screening method is based on a battery-powered low-cost reader designed to quantify serum bilirubin levels from whole blood applied to a lateral flow card.<sup>20</sup> From a maintenance point of view, a drawback is that this device requires daily calibration for routine use. The study showed a high correlation ( $r = 0.97$ ) of BiliSpec against a bench-top bilirubinometer (UNISTAT; Reichert Technologies) within a TSB range from 19 to 393 µmol/L with a mean ( $\pm$  SD) value of  $181 \pm 68$  µmol/L and differences up to 51 µmol/L (68% of total samples deviated  $\leq 17$  µmol/L). The mean bias between bench-top TSB and BiliSpec bilirubin readings was 5 µmol/L, with 95% limits of agreement of -29 to 38 µmol/L. In practice, lateral flow cards are designed to accept drops of whole blood obtained directly from a heel or finger prick. The separation from corpuscular components of the blood allows the flow of plasma into the nitrocellulose (NC) membrane by capillarity. Once the operator has visually interpreted that the NC membrane is saturated, the card is inserted into the reader for bilirubin measurement. The authors report that the design of the card allows controlling the volume of blood applied. The variability, however, in bilirubin test results could be the effect of an undersaturation or oversaturation of the NC membrane. In recent years, TSB measurement by another POC diagnostic method on capillary or venous blood samples became available with the Bilistick® System 1.0 (BM-BS 1.0 - Bilimetrix, Italy).<sup>21</sup> This direct method consists of a hand-held, rechargeable battery reflectance reader and test strips composed of a blood plasma separator coupled with an NC membrane, both encased in a plastic cassette. After loading the whole blood sample on the strip (35 µL -

**Table 1.** Characteristics of commonly used instruments to measure bilirubin concentrations in neonates.

	Multiparameter instruments and in vitro diagnostic devices (IVDs)	Bench-top bilirubinometers	Hand-held POC instruments
Costs	High	Moderate-low	Low
Time to test result	Hour	Minutes	Minutes
Blood volume	Large	Moderate-small	Small
Accuracy and precision	Acceptable	Variable	Variable
Assessment in external quality assessment (EQA) programs	Yes	No	No

hematocrit up to 70%), it requires less than 100 s for serum separation and NC membrane saturation, depending on the hematocrit of the sample (identified automatically by the reader using light reflectance measurements to detect serum flow stabilization). The reader measures reflected light from the plasma-saturated NC membrane, using a light-emitting diode (LED) with an emission peak at 465 nm for quantifying bilirubin. A second LED of 570 nm detects whether hemoglobin contamination is present. The instrument is internally calibrated to optimize sensitivity and provides TSB measuring in a range of 17 to 684  $\mu\text{mol/L}$ . The accuracy of the Bilistick® System 1.0 device was documented by comparing results with TSB measurements from routine laboratories. In 2018, Greco and colleagues reported the performance of the Bilistick® System 1.0 for identifying SNH in a multi-country approach in 1911 newborns. They showed that the TSB level measured by Bilistick® System 1.0 was not significantly different from laboratory TSB values in all four countries.<sup>22</sup> The Bilistick® System had a positive predictive value (PPV) of 92.5% and a negative predictive value (NPV) of 92.8%. When Greco and colleagues compared the Bilistick® System 1.0 with both transcutaneous bilirubinometry (JM-103) and laboratory TSB results (Synchron CX PRU 16360, Beckman-Coulter, USA), they found the Bilistick® System 1.0 to be a good alternative to transcutaneous bilirubin determination for early diagnosis and proper management of neonatal jaundice.<sup>23</sup> In 2018, Thielemans and colleagues reported a rather high failure rate for the Bilistick® System 1.0, especially in highly humid climatic conditions and at high hematocrit values.<sup>24</sup> In 2020, Kamineni and colleagues called to further improve the accuracy of the Bilistick.<sup>25</sup> Despite of these observations, the reliability and clinical use of the Bilistick® System 1.0 for measuring TSB was considered to be appropriate in other studies performed under similar weather conditions.<sup>22,26,27</sup> POC diagnosis of hyperbilirubinemia has also been claimed by Tabatabaee and colleagues who reported fast and reproducible TSB measurements in whole blood with a recently developed smartphone-based bilirubin assay kit using photoluminescent bacterial cellulose nanopaper.<sup>28</sup> One of the main advantages of the portable POC bilirubin instruments is the much shorter turnaround time, that is the interval between collecting the specimen and reporting the TSB result, compared to commonly used multi-analyzer instruments for TSB tests (Table 1).<sup>27</sup> Low-cost POC instruments appear to be an effective alternative for the measurement of TSB in newborns, particularly when conventional laboratory methods are not available or inaccessible.

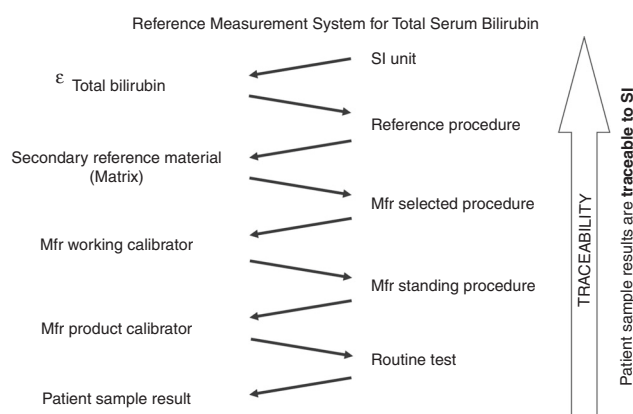
As shown by comparative analysis, the unacceptable high variability in TSB measurement among methods continues to pose a real challenge to result harmonization of clinical routine methods. When Lo and colleagues evaluated the trueness of neonatal TSB using value-assigned, commutable specimens in four major instrument groups (Dimension, Olympus, Synchron, and VITROS), they found a systematic error in TSB measurement associated primarily with the failure of instrument manufacturers to produce reliable bilirubin calibrators.<sup>17</sup> High variabilities were also observed when comparing TSB levels in patient specimens on multiparameter analyzers, transcutaneous bilirubin, and direct spectrophotometry instruments.<sup>29</sup> Other potential sources of inaccuracy include sample integrity and sample handling. These were, however, typically identified as random errors. An in-depth analysis of results reported by the College of American Pathologists Neonatal Bilirubin PT Survey from 2011 to 2015, showed how changes in TSB test results—when trying to recalibrate instruments—can lead to completely opposite clinical interpretations.<sup>30</sup> Standardization of TSB measurement remains a formidable challenge for laboratory medicine.<sup>31,32</sup>

The relevance of correctly implementing the internationally endorsed reference measurement systems  
Inaccuracy and non-equivalence of TSB results among IVD manufacturers are well known.<sup>17,30</sup> But why are TSB tests not

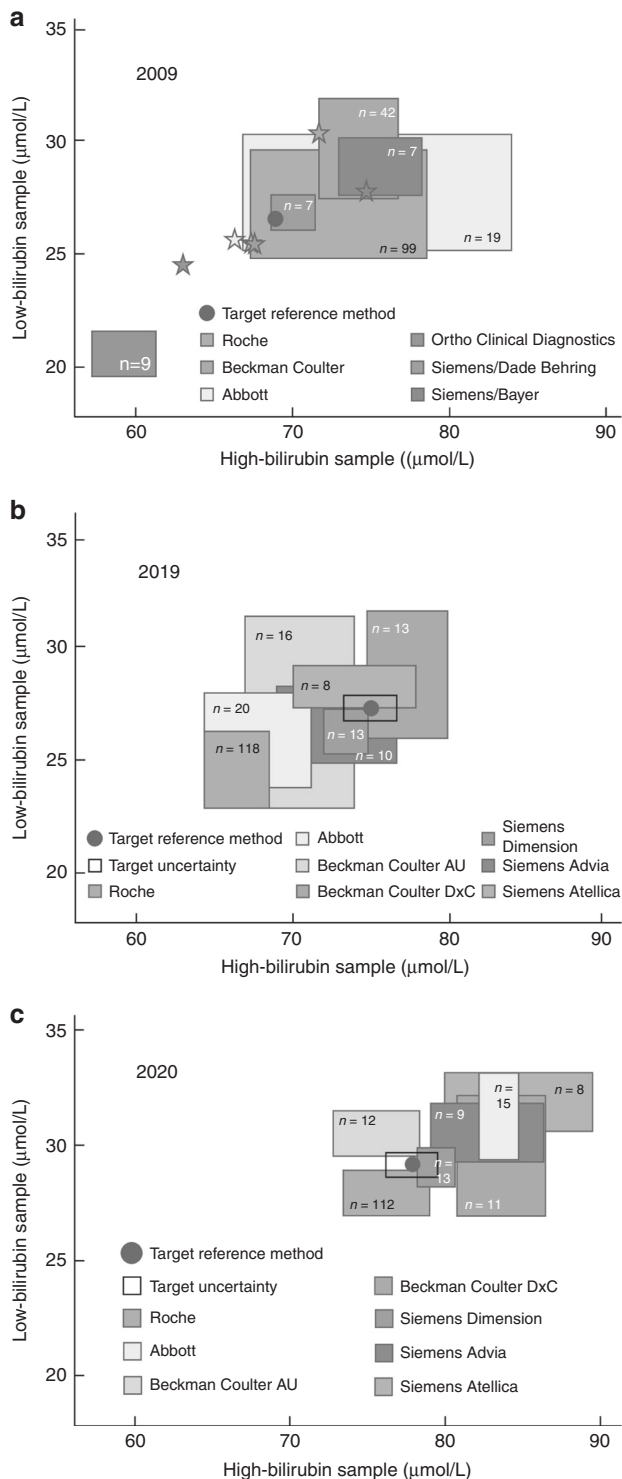
standardized adequately? It appears that there is insufficient awareness of the metrological traceability concept and its essential implementation through an unbroken chain of calibration hierarchies. Standardization is key to guaranteeing that TSB test results correspond properly to internationally agreed standards of a higher order (Fig. 1).<sup>33,34</sup> To achieve global standardization of measurement results in medical laboratories, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and the JCTLM promote the concept of metrological traceability of test results to internationally accepted standards. For proper commercial test calibration, medical test results must be anchored to higher-order reference materials and higher-order reference procedures. In the case of total bilirubin, tests were standardized with a Doumas reference method for bilirubin (using diazo-based spectrophotometry) and National Institute of Standards and Technology Standard Reference Material (NIST SRM) 916a reference material according to a strict calibration hierarchy. The former NIST SRM 916a reference material went missing years ago (it ran out of stock) and consisted of three isomers, two of which were not present in native human sera. It was decided therefore to use the specific molar extinction coefficient ( $\epsilon$ ) of 7649  $\text{m}^2/\text{mol}$  for bilirubin quantitation as it is superior to control the accuracy of a standard solution. Consequently, an updated and extended reference measurement system (RMS) was established and has been in place since 2018. To that end, Klauke and colleagues re-evaluated the Doumas candidate Reference Method and established a next-generation RMS without any need for a calibrator or SRM.<sup>35</sup> Currently, in this RMS total bilirubin is described as a so-called operative measurand, defined by a set of measurement parameters.

The German Society for Clinical Chemistry and Laboratory Medicine runs the Reference Laboratories International Federation of Clinical Chemistry (RELA IFCC) external quality assessment scheme. It periodically checks how to reference laboratories perform using their RMP, i.e., the gold standard for total bilirubin (<http://www.dgkl-rfb.de:81>).

Notwithstanding the existence of approved reference measurement systems (the former Doumas and AACC RMS (1985) versus the new RMS (2018)),<sup>35</sup> the levels of standardization and trueness



**Fig. 1 Reference measurement system and metrological traceability chain for TSB.** This measurement system was adapted from ISO 17511:2020 and based on the re-evaluated reference measurement procedure from Klauke and colleagues.<sup>35</sup> Total serum bilirubin is described as an operative measurand, as defined by the reference measurement procedure. Through this reference measurement system, with its unbroken chain of calibrator materials and methods, bilirubin test results are anchored and made traceable to the International System of Units. Measurement uncertainty should remain within allowable limits to make the test fit for the purpose. SI International System of Units;  $\epsilon$  the molar absorption coefficient of bilirubin (conventional quantity value of 7649  $\text{m}^2/\text{mol}$ ); Mfr, IVD manufacturer who supplies CE-IVD kits for routine analyses to the European market.



**Fig. 2 Total serum bilirubin measured in EQA samples by ~185 medical laboratories, common IVD manufacturers, and a JCTLM-endorsed reference laboratory.** Pooled human serum was supplemented with bilirubin. Serum samples were dispensed, frozen below  $-70^{\circ}\text{C}$ , and shipped on dry ice to the laboratories participating in the regular general clinical chemistry external quality assessment program of the Dutch Foundation for Quality Assessment in Medical Laboratories (SKML). In 2009 (a) and 2020 (c), samples were supplemented with unconjugated bilirubin alone (>98%, mixed isomers, Sigma-Aldrich). In 2019 (b), both unconjugated bilirubin and conjugated bilirubin were added. Each participating laboratory-measured total bilirubin. Results are plotted with the high-concentration sample on the x axis and the low-concentration sample on the y axis ( $17.1\ \mu\text{mol/L} = 1\ \text{mg/dL}$ ). Colored squares represent mean bilirubin concentrations  $\pm 1$  standard deviation in both samples, as measured by the routine medical laboratories in the quality assessment program and using the methods by the respective manufacturers as indicated. The numbers inside the colored squares indicate the numbers of participants for each category. In panel a, stars represent bilirubin concentrations measured in-house by the corresponding manufacturers. The red dots are the targets set in the JCTLM-listed reference laboratory of Prof. Gerhard Schumann and Dr. Denis Grote-Koska in Hannover, Germany (Institut für Klinische Chemie - Zentrallabor, Medizinische Hochschule Hannover). Panel a was adapted from Cobbaert and colleagues.<sup>10</sup> Data for panels b and c were supplied by the Dutch Foundation for Quality Assessment in Medical Laboratories and used with their permission.

error of 31.1% (<https://biologicalvariation.eu/>).<sup>36</sup> For neonatal bilirubin no biological variation data are available. Consequently, pediatricians can determine the analytic performance criteria needed to make the neonatal bilirubin test fit for purpose themselves.

Relevance of accuracy-based external quality assessment for trueness verification  
External quality assessment (EQA) plays an essential role in helping to assure the quality of laboratory medicine on a daily basis. EQA schemes may reveal significant and systematic between-method variability for measurements of the same analyte in the same specimen. Detection of between-method variability through EQA is also a major driver for further improvement of test standardization.

Medical laboratories are obliged to perform an EQA and when used effectively it can provide many opportunities for improving test accuracy. In the Netherlands, an accuracy-based EQA was developed by the Dutch Foundation for Quality Assessment in Medical Laboratories (Stichting Kwaliteitsbewaking Medische Laboratoria, SKML) for general chemistry analytes. Thus far, however, these did not include TSB. To develop an accuracy-based EQA, native, commutable, value-assigned EQA samples are essential to give medical laboratories insight into trueness and imprecision of their bilirubin tests as well as other chemistry tests. In this way, inaccuracy, and absolute bias as a result of, for example, lot-to-lot variation or method changes, can be monitored in a longitudinal and sustainable way by every participating laboratory. To date, no proven commutable EQA samples have been developed for TSB. This has not been accomplished yet because the EQA samples used until now are spiked with either unconjugated and/or conjugated bilirubin. It was demonstrated that the TSB recoveries are affected by the spiking material, hampering trueness verification. Nevertheless, interlaboratory and intermethod variations are monitored bi-weekly for TSB in approximately 185 Dutch laboratories. Figure 2 demonstrates the interlaboratory and intermethod variations in Dutch medical laboratories using common reagents. In the recent EQA surveys, SKML 2019.4 and SKML 2020.1, overall interlaboratory CVs ranged from 3% to 6% for TSB in the concentration range of 13 to 110

of routine total bilirubin test results consistently differ amongst manufacturers—as demonstrated in external quality assessment (EQA) programs that use commutable EQA samples—and has often been questioned. Yet, TSB tests can only be declared fit for clinical purposes if results are accurate within permissible limits of measurement uncertainty. The latter is often translated into desirable analytic performance specifications derived from biological variation data. In the case of TSB, rational and desirable analytic performances in adults are a  $\text{CV}_a$  (analytical variation) of <12.8%, an allowable bias of less than 10%, and a total allowable

µmol/L. The interlaboratory spreading has improved compared to the situation in 2009 (data not shown).

Other methods to study bilirubin metabolism—determination of bilirubin in tissues and cells

It is well known that the concentration of unconjugated free bilirubin correlates better with bilirubin neurotoxicity in comparison to total serum bilirubin concentration. Because bilirubin behaves like a real signaling molecule,<sup>37</sup> its blood concentrations can serve only as a rough surrogate marker of bilirubin metabolism within the cells. Although serum or plasma bilirubin concentrations are in dynamic equilibrium with other biological compartments, its tissue and cell concentrations differ substantially even within one organ as proved, for example, in the brain.<sup>38</sup> Knowledge of bilirubin kinetics and dynamics within individual human body compartments is essential to understand its role in the pathophysiology of various clinical conditions. Besides, bilirubin undergoes extensive metabolism, in particular, due to oxidation and photooxidation processes, forming tetra-, tri-, di- as well as monopyrrolic oxidation derivatives, which are likely to exert biological activities and may serve as important biomarkers of pathological conditions. Because cellular bilirubin concentrations are within the submicromolar range, the standard analytic, mostly diazo reaction-based methods, which are used in routine clinical chemistry, have insufficient sensitivity, and do not enable quantitation of bilirubin in the cells, tissues, and organs. This drawback of standard clinical chemistry methods is eliminated by using high-performance liquid chromatography (HPLC) techniques. These techniques enable accurate separation and quantitation of individual bilirubin fractions, such as delta bilirubin, unconjugated bilirubin, bilirubin monoglucuronosides, and diglucuronosides, and can be used under specific circumstances even in clinical settings.<sup>39</sup> Simultaneously, HPLC methods overcome the overestimation of bilirubin concentrations caused by the presence of unidentified diazo-positive compounds distinct from bilirubin.<sup>40</sup> High-performance liquid chromatography methods are also capable of differentiating various bilirubin isomers present under certain conditions in the biological systems.<sup>41</sup> Importantly, determination of the bilirubin subfractional changes, including detection of delta bilirubin, may help in the prediction of the risk of human diseases, such as cholestasis or gallstone disease, or in the differential diagnosis of such diseases.<sup>42,43</sup> Interestingly, the first method for separation and quantitative estimation of serum and biliary bilirubin fractions from serum and of three bilirubin fractions from bile was published as early as 1966.<sup>44</sup> Since that time, various chromatographic approaches were explored and a number of methods developed, including separation of native as well as derivatized bilirubins, such as ethyl anthranilate azo pigments or bilirubin methyl esters (for a review see ref. <sup>44</sup>). These methods were gradually improved. The separation was enhanced by stepping from isocratic, normal-phase HPLC, to various gradient, reverse-phase systems.<sup>45</sup> The use of internal standards, such as xanthobilirubic acid<sup>40</sup> or mesobilirubin,<sup>46</sup> led to improvements in accuracy. Further enhancement was reached with the employment of mass spectrometric detection.<sup>47</sup> With these analytic advances, methods for the detection of bilirubin and its metabolites in tissues and cells were established. In contrast to the early insufficient attempts to determine bilirubin in brain tissue, which were based on spectrophotometry,<sup>48</sup> HPLC-based methods demonstrated much higher sensitivity and accuracy. Using a newly developed HPLC method based on C8-column separation with the implementation of the methanol/water/tetrabutyl ammonium hydroxide mobile phase and equipped with the diode array detector, it was possible to detect as little as 10 pmol of bilirubin per gram of tissue.<sup>46</sup> This method was used in experimental studies quantifying bilirubin in numerous organs, including the heart,<sup>49</sup> and especially brain tissues,<sup>38,50–56</sup> which

are essential to understand the pathophysiology of bilirubin neurotoxicity.

Determination of bilirubin photoisomers and oxidation products With increasing knowledge on the biological importance of bilirubin derivatives formed during oxidation processes,<sup>57</sup> sensitive and accurate analytic methods are becoming essential. These derivatives include tetra-, tri-, di-, and monopyrrolic bilirubin oxidation derivatives. Probably the most clinically important are bilirubin photoisomers formed during PT of severe unconjugated hyperbilirubinemia. However, the determination of these bilirubin derivatives in biological material is not trivial, because of the lack of commercial standards as well as instability of the pigments. In terms of the determination of bilirubin photoisomers, several methods were published.<sup>58</sup> Previously, an improved analytic HPLC method for the simultaneous determination of major bilirubin photoisomers, lumirubin, Z,E- and E,Z-bilirubins, and bilirubin was described using lumirubin as well as internal standards with tandem mass detection.<sup>57</sup> The method was validated on serum samples of jaundiced neonates treated with PT. It has the potential of facilitating our understanding of the kinetics and biology of bilirubin photoproducts, which to date are practically unknown. Research into other bilirubin oxidation products is also progressing. Tripyrrolic biopyrrins, which are clinically relevant markers of increased oxidative stress, can be analyzed by immunochemical methods using specific anti-bilirubin monoclonal antibodies.<sup>59</sup> Reliable analytic methods for dipyrrolic propendyopents and monopyrrolic bilirubin oxidation products, Z-BOX A and B, with potential clinical implications were also published recently.<sup>60,61</sup> Finally, tetrapyrrolic compounds and their oxidation products were also studied in plants recently using these novel analytic methods and it will certainly improve our understanding of the biological relevance of these pigments.<sup>62,63</sup>

#### Summary and conclusion

Invasive TSB measurements remain the gold standard on which the definitive diagnosis of SNH is based. According to international guidelines for neonatal jaundice management, the clinical decision for treatment of neonatal hyperbilirubinemia should be based on bilirubin levels measured in blood by diagnostic instrumentation. Any non-invasive bilirubin estimation must be confirmed by an invasive diagnostic method. Bench-top bilirubinometers and hand-held POC instruments have a few advantages over multiparameter instruments of being cheaper and faster. Test results are available immediately compared to results of TSB measurements using multiparameter instruments in a central laboratory. In addition, less blood is needed as the bench-top bilirubinometers and hand-held POC instruments require minimal sample volumes. Novel POC bilirubin measurement methods, such as the BiliSpec and the Bilistick® System are of interest for many newborn infants, especially in LMICs, where the access to costly multiparameter instruments is limited. The main disadvantages of hand-held POC instruments and bench-top bilirubinometers are that agreement with routine laboratory TSB varies and that they are still not included in EQA programs. TSB test results on these instruments should be accurate within permissible limits of measurement uncertainty to be fit for clinical purposes. The key to accomplish this is anchoring TSB test results to the latest internationally endorsed RMS for bilirubin. In addition, participation in EQA programs for TSB in the neonatal range, and close interdisciplinary cooperation between physicians and clinical chemists are needed to assure the desired analytic and clinical performance of TSB testing. It is surprising that after bilirubin was first measured two centuries ago, uncertainty still exists on how to correctly assess the concentration of this yellow pigment. Universal implementation of endorsed calibration hierarchies for test standardization remains a daunting task. Recently, analytic methods for bilirubin measurement in biological matrices, such as HPLC thermal lens

spectrometry, spectrophotometric, molecular imprinting, and piezoelectric techniques were developed. These methods employ novel techniques that could further accelerate bilirubin research to improve the management of newborn infants with SNH.

## ACKNOWLEDGEMENTS

We greatly appreciate the help of T. van Wulfften Palthe in correcting the English grammar and language. The study was supported by grants NV18-07-00342 and RVO-VFN64165/2020 from the Czech Ministry of Health. The support of an intramural grant of Fondazione Italiana Fegato to Claudio Tiribelli is appreciated.

## AUTHOR CONTRIBUTIONS

Study concept and design, and retrieving and analyzing the literature: C.V.H., L.V., C.D. C.Z., C.C., and C.T. Drafting and critical revision of the manuscript for important intellectual content: C.V.H., L.V., C.D.C.Z., A.D., P.S., E.A.E.v.d.H., C.C., and C.T. All authors approved the final manuscript as submitted.

## ADDITIONAL INFORMATION

**Competing interests:** Claudio Tiribelli is the President and Carlos D. Coda Zabetta is the CTO of Bilimetrix, the company responsible for the development of the Bilistick® System. The remaining authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Tiedemann, F. & Gmelin, L. In *Die Verdauung nach Versuchen* Ch. 10 (ed. Groos, K.) 79 (1826).
2. Frerichs, F. T. *Klinik der Leberkrankheiten* 1st edn, Vol. I of 2 volumes (1858).
3. Van den Bergh, A. A. H. & Snapper, J. Die Farbstoffe des Blutserums. *Deut. Arch. Klin. Med.* **110**, 540–561 (1913).
4. Malloy, H. T. & Evelyn, K. A. The determination of bilirubin with the photoelectric colorimeter. *J. Biol. Chem.* **119**, 481–490 (1937).
5. Jendrassik, L. & Grof, P. Vereinfachte photometrische Methoden zur Bestimmung des Blutbilirubins. *Biochem. Z.* **297**, 81–89 (1938).
6. BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP, ISO, OIML. *The International Vocabulary of Metrology—Basic And General Concepts and Associated Terms (VIM)* 3rd edn. [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_200\\_2012.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf) (2012).
7. Ngashangva, L., Bachu, V. & Goswami, P. Development of new methods for determination of bilirubin. *J. Pharm. Biomed. Anal.* **162**, 272–285 (2019).
8. Vreman, H. J. et al. Interlaboratory variability of bilirubin measurements. *Clin. Chem.* **42**, 869–873 (1996).
9. Dumas, B. T. & Eckfeldt, J. H. Errors in measurement of total bilirubin: a perennial problem. *Clin. Chem.* **42**, 845–848 (1996).
10. Cobbaert, C., Weykamp, C. & Hulzebos, C. V. Bilirubin standardization in the Netherlands: alignment within and between manufacturers. *Clin. Chem.* **56**, 872–873 (2010).
11. Greene, D. N., Liang, J., Holmes, D. T., Resch, A. & Lorey, T. S. Neonatal total bilirubin measurements: Still room for harmonization. *Clin. Biochem.* **47**, 1112–1115 (2014).
12. Kirk, J. M. Neonatal jaundice: a critical review of the role and practice of bilirubin analysis. *Ann. Clin. Biochem.* **45**, 452–462 (2008).
13. Lano, I. M., Lyon, A. W., Wang, L., Ruskin, R. & Lyon, M. E. Comparative evaluation of neonatal bilirubin using radiometer whole blood co-oximetry and plasma bilirubin methods from Roche diagnostics and ortho clinical diagnostics. *Clin. Biochem.* **53**, 88–92 (2018).
14. Wang, L. et al. Limitations and opportunities of whole blood bilirubin measurements by GEM premier 4000®. *BMC Pediatr.* **17**, 92 (2017).
15. Kazmierczak, S. C. et al. Direct spectrophotometric method for measurement of bilirubin in newborns: comparison with HPLC and an automated diazo method. *Clin. Chem.* **48**, 1096–1097 (2002).
16. Barko, H. A., Jackson, G. L. & Engle, W. D. Evaluation of a point-of-care direct spectrophotometric method for measurement of total serum bilirubin in term and near-term neonates. *J. Perinatol.* **26**, 100–105 (2006).
17. Lo, S. F., Jendrzyszczak, B. & Dumas, B. T. Laboratory performance in neonatal bilirubin testing using commutable specimens: a progress report on a college of american pathologists study. *Arch. Pathol. Lab Med.* **132**, 1781–1785 (2008).

18. Olusanya, B. O., Ogunlesi, T. A. & Slusher, T. M. Why is kernicterus still a major cause of death and disability in low-income and middle-income countries? *Arch. Dis. Child* **99**, 1117–1121 (2014).
19. Slusher, T. M. et al. Burden of severe neonatal jaundice: a systematic review and meta-analysis. *BMJ Paediatr. Open* **1**, e000105 (2017).
20. Keahey, P. A. et al. Point-of-care device to diagnose and monitor neonatal jaundice in low-resource settings. *PNAS* **114**, E10965–E10971 (2017).
21. Coda Zabetta, C. D. et al. Bilistick: a low-cost point-of-care system to measure total plasma bilirubin. *Neonatology* **103**, 177–181 (2013).
22. Greco, C. et al. Diagnostic performance analysis of the point-of-care Bilistick system in identifying SNH by a multi-country approach. *EClinicalMedicine* **1**, 14–20 (2018).
23. Greco, C. et al. Comparison between Bilistick System and transcutaneous bilirubin in assessing total bilirubin serum concentration in jaundiced newborns. *J. Perinatol.* **37**, 1028–1031 (2017).
24. Thielemans, L. et al. Laboratory validation and field usability assessment of a point-of-care test for serum bilirubin levels in neonates in a tropical setting. Version 2. *Wellcome Open Res.* **3**, 110 (2018).
25. Kamineni, B. et al. Accuracy of Bilistick (a point-of-care device) to detect neonatal hyperbilirubinemia. *J. Trop. Pediatr.* **66**, 630–636 (2020).
26. Rohsiswatmo, R. et al. Agreement test of transcutaneous bilirubin and bilistick with serum bilirubin in preterm infants receiving phototherapy. *BMC Pediatr.* **18**, 315 (2018).
27. Boo, N.-Y. et al. The point-of-care Bilistick method has very short turn-around-time and high accuracy at lower cutoff levels to predict laboratory-measured TSB. *Pediatr. Res.* **86**, 216–220 (2019).
28. Tabatabaee, R. S., Golmohammadi, H. & Ahmadi, S. H. Easy diagnosis of jaundice: a smartphone-based nanosensor bioplatfrom using photoluminescent bacterial nanopaper for point-of-care diagnosis of hyperbilirubinemia. *ACS Sens.* **4**, 1063–1071 (2019).
29. Grohmann, K. et al. Bilirubin measurement for neonates: comparison of 9 frequently used methods. *Pediatrics* **117**, 1174–1183 (2006).
30. Lo, S. F. Laboratory accuracy in neonatal bilirubin: the search for truth in laboratory medicine. *JAMA Pediatr.* **170**, 529–530 (2016).
31. Stepman, H. C. M. et al. Measurements for 8 common analytes in native sera identify inadequate standardization among 6 routine laboratory assays. *Clin. Chem.* **60**, 855–863 (2014).
32. Vitek, L. Bilirubin as a predictor of diseases of civilization. Is it time to establish decision limits for serum bilirubin concentrations? *Arch. Biochem. Biophys.* **672**, 108062 (2019).
33. Jones, G. R. & Jackson, C. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) - its history and operation. *Clin. Chim. Acta* **453**, 86–94 (2016).
34. Cobbaert, C., Smit, N. & Gillery, P. Metrological traceability and harmonization of medical tests: a quantum leap forward is needed to keep pace with globalization and stringent IVD-regulations in the 21st century! *Clin. Chem. Lab Med.* **56**, 1598–1602 (2018).
35. Klauke, R. et al. Reference measurement procedure for total bilirubin in serum re-evaluated and measurement uncertainty determined. *Clin. Chim. Acta* **481**, 115–120 (2018).
36. Fraser, C. G. & Peake, M. J. Problems associated with clinical chemistry quality control materials. *CRC Crit. Rev. Clin. Lab Sci.* **12**, 59–86 (1980).
37. Vitek, L. Bilirubin as a signaling molecule. *Med. Res. Rev.* **40**, 1335–1351 (2020).
38. Gazzin, S. et al. Bilirubin accumulation and Cyp mRNA expression in selected brain regions of jaundiced Gunn rat pups. *Pediatr. Res.* **71**, 653–660 (2012).
39. Adachi, Y. et al. Clinical application of serum bilirubin fractionation by simplified liquid chromatography. *Clin. Chem.* **34**, 385–388 (1988).
40. Muraca, M. & Blanckaert, N. Liquid-chromatographic assay and identification of mono- and diester conjugates of bilirubin in normal serum. *Clin. Chem.* **29**, 1767–1771 (1983).
41. Aziz, S. et al. Bilirubin-IX alpha and -IX beta pigments, coproporphyrins and bile acids in meconium and stools from full-term and preterm neonates during the first month of life. *Acta Paediatr.* **90**, 81–87 (2001).
42. Spivak, W. & Yuey, W. Application of a rapid and efficient h.p.l.c. method to measure bilirubin and its conjugates from native bile and in model bile systems. Potential use as a tool for kinetic reactions and as an aid in diagnosis of hepatobiliary disease. *Biochem. J.* **234**, 101–109 (1986).
43. Ostrea, E. M. Jr, Ongtengco, E. A., Tolia, V. A. & Apostol, E. The occurrence and significance of the bilirubin species, including delta bilirubin, in jaundiced infants. *J. Pediatr. Gastroenterol. Nutr.* **7**, 511–516 (1988).
44. Kuenzle, C. C., Sommerhalder, M., Ruttner, J. R. & Maier, C. Separation and quantitative estimation of four bilirubin fractions from serum and of three bilirubin fractions from bile. *J. Lab. Clin. Med.* **67**, 282–293 (1966).
45. Heirwegh, K. P., Fevery, J. & Blanckaert, N. Chromatographic analysis and structure determination of biliverdins and bilirubins. *J. Chromatogr.* **496**, 1–26 (1989).

46. Zelenka, J. et al. Highly sensitive method for quantitative determination of bilirubin in biological fluids and tissues. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **867**, 37–42 (2008).
47. Jašprova, J. et al. A novel accurate LC-MS/MS method for quantitative determination of Z-lumirubin. *Sci. Rep.* **10**, 4411 (2020).
48. Sawasaki, Y., Yamada, N. & Nakajima, H. Developmental features of cerebellar hypoplasia and brain bilirubin levels in a mutant (Gunn) rat with hereditary hyperbilirubinaemia. *J. Neurochem.* **27**, 577–583 (1976).
49. Muchova, L. et al. Statin treatment increases formation of carbon monoxide and bilirubin in mice: a novel mechanism of in vivo antioxidant protection. *Can. J. Physiol. Pharm.* **85**, 800–810 (2007).
50. Cuperus, F. J. et al. Beyond plasma bilirubin: the effects of phototherapy and albumin on brain bilirubin levels in Gunn rats. *J. Hepatol.* **58**, 134–140 (2013).
51. Schreuder, A. B. et al. Albumin administration protects against bilirubin-induced auditory brainstem dysfunction in Gunn rat pups. *Liver Int.* **33**, 1557–1565 (2013).
52. Schreuder, A. B. et al. Optimizing exchange transfusion for severe unconjugated hyperbilirubinemia: studies in the Gunn rat. *PLoS ONE* **8**, e77179 (2013).
53. Bortolussi, G. et al. Life-long correction of hyperbilirubinemia with a neonatal liver-specific AAV-mediated gene transfer in a lethal mouse model of Crigler-Najjar syndrome. *Hum. Gene Ther.* **25**, 844–855 (2014).
54. Vodret, S. et al. Albumin administration prevents neurological damage and death in a mouse model of SNH. *Sci. Rep.* **5**, 16203 (2015).
55. Vodret, S., Bortolussi, G., Jasprova, J., Vitek, L. & Muro, A. F. Inflammatory signature of cerebellar neurodegeneration during neonatal hyperbilirubinemia in Ugt1<sup>-/-</sup> mouse model. *J. Neuroinflamm.* **14**, 64 (2017).
56. Bockor, L. et al. Modulation of bilirubin neurotoxicity by the Abcb1 transporter in the Ugt1<sup>-/-</sup> lethal mouse model of neonatal hyperbilirubinemia. *Hum. Mol. Genet.* **26**, 145–157 (2017).
57. Jasprova, J. et al. The biological effects of bilirubin photoisomers. *PLoS ONE* **11**, e0148126 (2016).
58. Itoh, S., Isobe, K. & Onishi, S. Accurate and sensitive high-performance liquid chromatographic method for geometrical and structural photoisomers of bilirubin IX alpha using the relative molar absorptivity values. *J. Chromatogr. A* **848**, 169–177 (1999).
59. Vitek, L., Kraslova, I., Muchova, L., Novotny, L. & Yamaguchi, T. Urinary excretion of oxidative metabolites of bilirubin in subjects with Gilbert syndrome. *J. Gastroenterol. Hepatol.* **22**, 841–845 (2007).
60. Joerk, A. et al. Propentdyopents as heme degradation intermediates constrict mouse cerebral arterioles and are present in the cerebrospinal fluid of patients with subarachnoid hemorrhage. *Circ. Res.* **124**, e101–e114 (2019).
61. Clark, J. F., Loftspring, M., Wurster, W. L. & Pyne-Geithman, G. J. Chemical and biochemical oxidations in spinal fluid after subarachnoid hemorrhage. *Front. Biosci.* **13**, 1806–1812 (2008).
62. Konickova, R. Anti-cancer effects of blue-green alga *Spirulina platensis*, a natural source of bilirubin-like tetrapyrrolic compounds. *Ann. Hepatol.* **13**, 273–283 (2014).
63. Ritter, M. et al. Pyrrolic and dipyrrolic chlorophyll degradation products in plants and herbivores. *Chemistry* **26**, 6205–6213 (2020).