IFCC Paper

Christian V. Hulzebos, Johanna E. Camara, Miranda van Berkel, Vincent Delatour, Stanley F. Lo, Agnès Mailloux, Marcel C. Schmidt, Mercy Thomas, Lindsey G. Mackay* and Ronda F. Greaves*, on behalf of the IFCC Working Group Neonatal Bilirubin

Bilirubin measurements in neonates: uniform neonatal treatment can only be achieved by improved standardization

https://doi.org/10.1515/cclm-2024-0620
Received May 18, 2024; accepted June 28, 2024;
published online July 26, 2024

Abstract: Measurement of total bilirubin (TBil) concentration in serum is the gold standard approach for diagnosing neonatal unconjugated hyperbilirubinemia. It is of utmost importance that the measured TBil concentration is sufficiently accurate to prevent under treatment, unnecessary escalation of care, or overtreatment. However, it is widely recognized that TBil measurements urgently require improvement in neonatal clinical chemistry. External quality assessment (EQA) programs for TBil assess for differences between laboratories and provide supporting evidence of significant differences between various methods, manufacturers and measurement platforms. At the same time, many countries have adopted or only slightly adapted the neonatal hyperbilirubinemia management guidelines from the USA or UK, often without addressing differences in the methodology of TBil measurements. In this report, we provide an overview of the components of bilirubin that are measured by laboratory platforms, the availability of current reference measurement procedures and reference materials, and the role of EQA surveys in this context. Furthermore, the current status of agreement in neonatal bilirubin against clinical decision thresholds is reviewed. We advocate for enhancements in accuracy and comparability of neonatal TBil measurements, propose a path forward to accomplish this, and reflect on the position of the International Federation for Clinical Chemistry and Laboratory Medicine (IFCC) Working Group Neonatal Bilirubin (WG-NB) in this matter.

Keywords: neonatal bilirubin; treatment; accuracy

Introduction

Measurement of total bilirubin (TBil) concentration in serum is considered the gold standard for diagnosing neonatal unconjugated hyperbilirubinemia [1]. Current international clinical guidelines utilize a TBil-based, hour-specific nomogram, and treatment thresholds depend on the presence of specific hyperbilirubinemia and/or neurotoxicity risk factors [2, 3]. As such, accurate TBil measurements are essential to diagnose the severity of hyperbilirubinemia, and to monitor the response to any installed treatment [4]. However, it is

***Corresponding authors: Ronda F. Greaves**, Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Parkville, VIC 3052, Australia; and Department of Paediatrics, University of Melbourne, Parkville, VIC 3052, Australia, E-mail: ronda.greaves@mcri.edu.au. https://orcid.org/0000-0001-7823-8797; and **Lindsey G. Mackay**, National Measurement Institute, Sydney, NSW 2113, Australia, E-mail: lindseygmackay@gmail.com. https://orcid.org/0000-0002-7652-4682

Christian V. Hulzebos, Department of Paediatrics, Division of Neonatology, Beatrix Children's Hospital, University Medical Center Groningen, Groningen, The Netherlands. https://orcid.org/0000-0002-8159-7501

Johanna E. Camara, Chemical Sciences Division, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA. https://orcid.org/0000-0002-9415-8452

Miranda van Berkel, Department of Laboratory Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands. https://orcid.org/0000-0002-6500-3419

Vincent Delatour, Laboratoire National de Métrologie et d'Essais (LNE), Paris, France. https://orcid.org/0000-0003-2447-0341

Stanley F. Lo, Department of Pathology and Laboratory Medicine, Children's Wisconsin and Medical College of Wisconsin, Milwaukee, WI, USA

Agnès Mailloux, Centre National de Référence en Hémobiologie Périnatale, Unit of Biologie, Pole Biology, Hopital Saint Antoine (Public Assistance Hospitals of Paris (AP-HP)), Paris, France

Marcel C. Schmidt, Roche Diagnostics GmbH, Penzberg, Germany

Mercy Thomas, The Royal Children's Hospital, Parkville, VIC, Australia; Murdoch Children's Research Institute, Parkville, VIC, Australia; and School of Health Sciences, Swinburne University of Technology, Melbourne, VIC, Australia. https://orcid.org/0000-0001-5324-3119

widely recognized that TBil measurement urgently requires improvement in terms of reducing bias and thereby propagating comparability between different methods currently in use in clinical chemistry.

Over 60 years ago, it was stated that "bilirubin determinations are perhaps the most notoriously unreliable of any in clinical chemistry" [4]. Though there have been significant improvements in TBil measurement, relatively high variability of measurement between manufacturers and measurement platforms continues to be repeatedly shown in neonatal patient comparisons and external quality assessment (EQA) programs [1, 4–7]. As such, the IFCC working group neonatal bilirubin (WG-NB) was formed in 2022 to improve neonatal bilirubin measurements [8].

In this paper, we review the availability and appropriateness of reference materials, reference measurement procedures (RMPs) and services in order to define needs and opportunities to improve comparability and traceability of neonatal bilirubin testing to meet clinical needs. We also aim to understand the current status of agreement in neonatal bilirubin (i.e. usually TBil) measurement at clinical decision concentrations, and EQA performance.

To address the aims of this paper, we first highlight the treatment cascade and potential consequences of neonatal jaundice management (Figure 1). Then we pose six key questions to explore the current status of TBil measurement clinical need and its traceability:

(1) Do all central laboratory platforms measure the same components of bilirubin?
(2) What is the status of reference measurement systems?
(3) What is the reported accuracy of bilirubin measurement procedures in national proficiency testing surveys?
(4) What about other technologies for assessment of neonatal jaundice?
(5) How do clinical guidelines interpret TBil results?
(6) What is the impact of bias in neonatal bilirubin measurement on clinical decisions?

Treatment cascade

Neonatal jaundice management (Figure 1) is a matter of debate in virtually all newborns because the majority develop clinically visible jaundice in the first postnatal week. Most often, jaundice will resolve spontaneously in a well-baby with sufficient feeding volume. Unfortunately, and usually not foreseen, jaundice may progress to severe

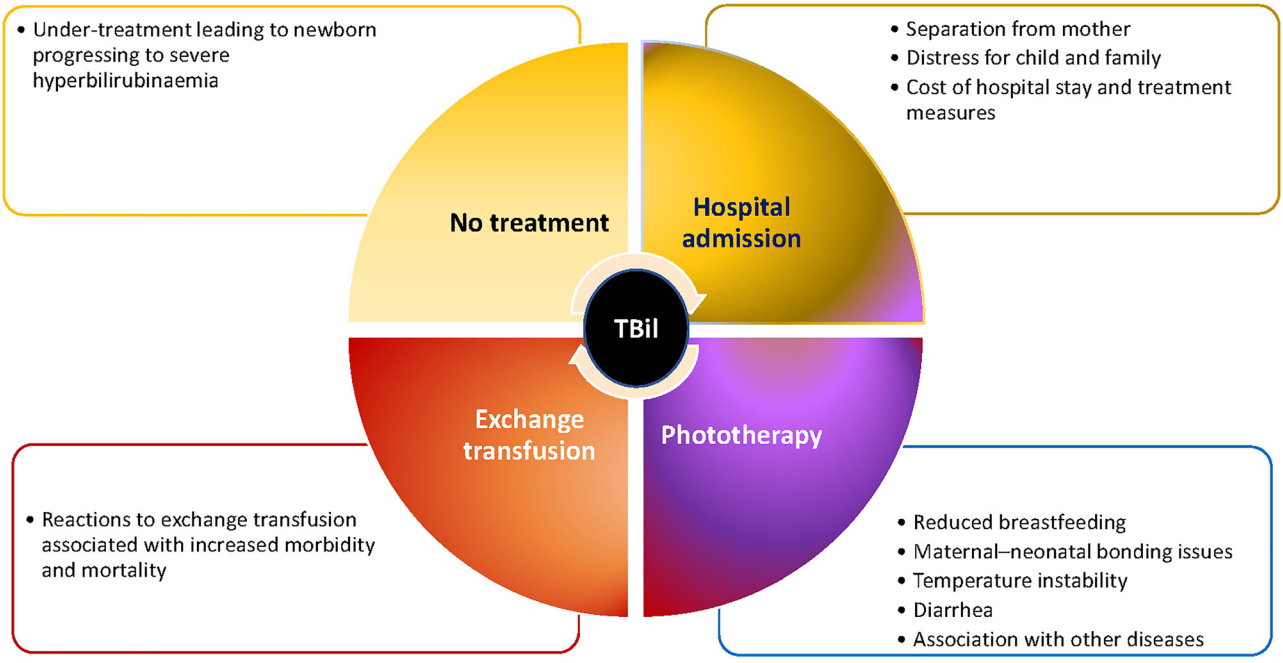

Figure 1: Treatment cascade and potential consequences of neonatal jaundice management. TBil, total bilirubin.

neonatal hyperbilirubinemia. When left untreated, acute bilirubin encephalopathy and potentially devastating long-term neurological sequelae, i.e., kernicterus-spectrum disorders, may occur [9]. Neonatal jaundice is an important reason for hospital admission of newborns. Hospital admission may, next to in-hospital phototherapy (PT), lead to parental emotional stress, disturbances in parent-infant bonding, difficulties in breastfeeding, and contribute to escalating hospital costs and resource utilization. Although generally considered safe, PT is known to have a few side effects. Apart from the pharmacological effect of the delivered light photons on bilirubin molecules, PT can lead to temperature instability or diarrhea. Furthermore, a tendency towards increased mortality in mechanically ventilated infants ≤750 g receiving PT has been reported and – more recently an association between PT and epilepsy in boys [10–12]. Thus, unnecessary PT exposure should be prevented and when indicated, the intensity, frequency and duration of PT should be carefully and individually tailored [13]. The most radical decision in neonatal jaundice management is an exchange transfusion (ET), that is associated with morbidity, including necrotizing enterocolitis, infection, thrombocytopenia, and even mortality [14].

Key questions

Q1. Do all central laboratory platforms measure the same components of bilirubin?

In general, for both adult and pediatric samples, the TBil in native samples is a variable mixture of different bilirubin species and isomers. It comprises non-covalently albumin-bound unconjugated bilirubin, free unconjugated bilirubin, mono- and di-glucuronated (conjugated) bilirubin, the covalently albumin-linked delta bilirubin and, if present, photo-isomers. The bilirubin species are distinguished based on their reaction behavior in the classic detection method using an azo-dye. Bilirubin species that react directly with the reagent, i.e. without support from an accelerator, are conjugated bilirubin and delta bilirubin as well as a tiny amount of free, non-albumin-linked unconjugated bilirubin [15]. Unconjugated bilirubin needs an accelerator to make the reaction site on the bilirubin accessible for rapid reaction with the diazo reagent. This is mediated by caffeine in the classic reaction following Jendrassik-Grof or by detergent in modern automated TBil assays. For note, the various bilirubin measurement principles have been reviewed in the manuscript by Thaler et al. and Ngashangva et al. [16, 17].

Discussion

The nomenclature of indirect and direct bilirubin is based on the chemically reactive behavior in the diazo reaction, even if direct bilirubin is often incorrectly equated only with structurally defined conjugated bilirubin. Assays for Tbil measure fractions that react both directly and indirectly in the diazo reaction. Commercially available bilirubin assays for automated clinical chemistry analyzers are usually available for Tbil and direct bilirubin. The majority of automated clinical chemical analyzers use methods for the determination of bilirubin using diazo reagents. Bilirubin assays based on the azo-coupling principle are unable to distinguish conjugated or delta-bilirubin from other bilirubin species that react directly in the assay. Bilirubin assays using the vanadate oxidation method are also offered as total and direct bilirubin assays. The specificity for direct bilirubin or for TBil depends on the presence or omittance of detergent in the reagent formulation. The vanadate method does also potentially detect “all” bilirubin species [18, 19], however these assays are usually not designed to distinguish the individual bilirubin species.

A third option for routine bilirubin determination is multi-layer film technology where slides for TBil and specific slides for the differentiation of unconjugated and conjugated (not “direct”) bilirubin are available. While the TBil slide follows the diazo technique, the second slide (BuBc) uses spectrophotometry to differentiate unconjugated bilirubin from the sum of mono- and di-conjugated bilirubin by the application of different layers and an algorithm calculating the individual concentrations from the measured signals. Due to the nature of the layers, delta bilirubin cannot be detected with the BuBc slide [16, 20].

In general, the difference between the measurement of mono- and di-glucuronide from the TBil provides an indirect measurement of the delta bilirubin fraction. Additional structural and configurational bilirubin isomers may also be formed quickly under phototherapy but are not routinely measured [21].

Currently no analyzers directly measure the delta bilirubin form or the more polar bilirubin isomers formed during PT, however these have been assessed by HPLC methods [22, 23]. There are different clinical impacts from the differing forms of bilirubin, for example unconjugated hyperbilirubinemia and photo-isomers are relevant to bilirubin-induced neurological damage (kernicterus spectrum disorders) whereas conjugated forms are relevant to specific diseases including biliary atresia [24, 25]. Thus,

different measurands would ideally be considered for different intended uses. Certainly, the specific measurands being determined by different methods may warrant further consideration. In the case of blood gas analyzers, it is likely that the analytes being actually measured may be different compared to traditional clinical analyzers and this will be of even greater complexity for hand-held near patient testing instruments. Thus, the effective linkage of measurements across these different platforms may be challenging.

Q2. What is the status of reference measurement systems?

The reference measurement system (RMS) for bilirubin is supported by reference materials, RMPs, and reference services. Some but not all of the components of this RMS are currently listed on the Joint Committee for Traceability in Laboratory Medicine's (JCTLM) database [26].

Discussion

National Institute of Standards and Technology (NIST) Standard Reference Material® (SRM) 916b bilirubin is a certified reference material consisting of neat unconjugated bilirubin with a certified mass fraction purity of (94.7±0.7) %. The mass purity determination was made by quantitative ^{1}H nuclear magnetic resonance spectroscopy using an internal standard (q^{1}H-NMR$_{IS}$) and is metrologically traceable to the International System of Units (SI). Comparison of qNMR results to diazo spectrophotometric reference measurement procedure (RMP) results demonstrated that the methods are consistent and molar absorptivity values calculated from the spectrophotometric RMP provide evidence for the continuity of the reference system initially established for bilirubin using NIST SRM 916 and NIST SRM 916a. The bilirubin IX-α isomer is predominant in NIST SRM 916b. Bilirubin III-α content is estimated to be 3 % (g/g, mol/kg) greater than that of the XIII-α isomer, which is consistent with the isomer composition observed for SRM 916a. Organic impurities structurally similar to bilirubin, including biliverdin or mesobilirubin, were not identified. Observations from ^{1}H(^{13}C)-NMR spectra suggested that the NIST SRM 916b material contains a greater quantity of paramagnetic impurities vs. NIST SRM 916. The total content of paramagnetic and ferromagnetic elements is approximately 0.1 % (g/g). The total mass fraction of all elements readily attributable to impurity components, determined by semiquantitative ICP-MS and X-ray fluorescence spectroscopy methods, is approximately 0.6 %. Combustion analyses and elemental microanalysis suggest that the unidentified impurities are largely of organic nature and may include macromolecular species or larger organic particles not observable via ^{1}H-NMR [27].

While the previous batch of NIST SRM 916a bilirubin was listed in the JCTLM database, it was delisted in 2014 after it was out of stock. The new batch of NIST SRM 916b bilirubin has been available since 2021 and was submitted for the 2024 JCTLM review cycle. NIST also provides a frozen plasma matrix material with a non-certified reference value for bilirubin.

The JCTLM database lists two entries of RMPs published for bilirubin. However, both methods are based on spectrophotometry and provide the same reference so there is only one method on the JCTLM database [28]. In 2018, the original Doumas method was adapted to be applied without the use of the bilirubin calibrant due to the lack of available NIST SRM 916a bilirubin at that time [29]. This variation of the method relies on the use of a molar absorption coefficient for calibration, rather than a higher order reference material. It is not currently listed in the JCTLM database and thus it still needs to be formally assessed as a candidate RMP.

Historically, precision for the Doumas RMP has been reported to achieve a coefficient of variation of <1 % for values above 100 µmol/L (divide by 17.1 to convert to mg/dL) [28]. Inter-laboratory Doumas RMP coefficients of variation at concentrations >34.2 µmol/L were less than 2.4 % [30]. While many laboratories report use of the Doumas spectrophotometric method, only three of these laboratories are currently listed in the JCTLM database for providing Reference Measurement Services for bilirubin [26]. A list of all of the laboratories that participate in the "RELA External Quality Control for Reference Laboratories" scheme for total bilirubin and their performance is available on the RELA website [31]. Typically, around 23 laboratories participate each year (noting this is not a dedicated neonatal program), three of which are the JCTLM-listed reference service providers. The typical CVs for these results are ≈5–10 %.

Q3. What is the reported accuracy of bilirubin measurement procedures in national proficiency testing surveys?

EQA surveys have been implemented for decades within the clinical laboratory community and serve as a useful tool to assess the status of agreement among results within and between end-user measurement procedures. A number of these programs attempt to include the 'true' value of the samples by targeting the EQA samples with a RMP. In these cases, and provided that EQA materials are commutable, method specific bias can be accurately estimated. Since most

developed countries have an EQA survey, this can provide a valuable tool to assess and compare the status of standardization of the neonatal TBil results across different countries. In the Netherlands, a neonatal EQA program with human serum spiked with unconjugated bilirubin, value assigned using the Doumas RMP is in place. These EQA based results increasingly report differences between methods [5, 7]. Similarly, the College of American Pathologists (CAP) neonatal bilirubin survey utilizes an RMP determined target value that is assigned to human serum samples spiked with unconjugated bilirubin. CAP data reported a 27 % difference between maximum and minimum peer group mean values obtained by individual laboratories [32].

Discussion

Although various EQA programs consistently report differences between methods and the assignment of RMP target values makes it possible to assess bias, there are some major prerequisites to this method of post market surveillance. The commutability of the samples used is often not, or not formally, demonstrated and typically, EQA providers are often limited to the use of adult human serum due to obvious practical restraints in obtaining neonatal material in sufficient amount at relevant concentrations. In EQA programs the serum is often spiked with unconjugated bilirubin, to obtain elevated concentrations encountered in the biological materials from neonates at risk for jaundice. As the conjugation function of neonates is not fully mature [32, 33], it remains uncertain to what extent serum from adult donors spiked with unconjugated bilirubin properly mimics the characteristics of serum from neonates. For TBil, the lack of commutable calibrators and EQA materials likely contributes to discrepancies reported in bias of bilirubin results between native neonatal material and in adult serum-based control samples. For instance, in a French study, it was shown that some EQA materials have insufficient commutability: in this case, results of an individual laboratory can only be compared with its peer laboratories using the same analyzer [34]. In this study, the authors also showed that addition of some cryoprotectants like polyethylene glycol (PEG) could alter commutability of calibration materials specifically for some assays. Additionally, effects of added conjugated bilirubin seems to also affect results provided by some specific platforms, which has not been well investigated [6, 7, 35].

Most EQA programs for bilirubin testing are established for all patient samples and not specific to neonates, which may be inappropriate given that differences may exist between samples from adults and from neonates (e.g., differences in the amount of conjugated vs. unconjugated bilirubin etc.). In addition, some EQA providers lack a target value as established by an RMP, allowing only comparability to other methods, rather than assessment of bias vs. the 'true' value. Ultimately, we advocate that where possible best practice would be the establishment of EQA programs specifically dedicated to neonatal TBil measurement with commutable control materials value assigned with an RMP.

Q4. What about other technologies for assessment of neonatal jaundice?

The technologies discussed in this manuscript focus on central laboratory instruments used for diagnosis and monitoring of neonatal jaundice measured in serum. In the neonatal ward whole blood bilirubin blood gas analyzers, which use spectrophotometric methods, are used extensively for monitoring. Likewise transcutaneous bilirubin (TcB) devices are used extensively for screening in postnatal wards and community settings. Bilirubin measurement on blood gas devices has a different traceability chain compared to mainframe laboratory instruments, the traceability of TcB instruments has not been well investigated.

Discussion

Over the last decade there has been development in near patient testing bilirubin devices, encompassing both blood based and non-invasive approaches, proposed for community screening of neonatal jaundice [36]. Appropriate screening tests will have a decision cascade of e.g., screen negative (i.e., no further action), uncertain (i.e., monitor with a repeat test) or screen positive (i.e., refer for tertiary clinical care). Considerations of traceability and how the decision limits are formed for these devices is beyond the scope of this position paper, but will continue to be an important consideration as more of these devices are adopted into patient care pathways. Furthermore, when wanting to integrate these devices into a patient care pathway, additional challenges on performing the traditional mandatory method comparisons prevail. Importantly, the accuracy of TBil results provided by the central laboratory is likely to be integral to the standardization of these devices.

Q5. How do clinical guidelines interpret TBil results?

Clinical practice guidelines (CPGs) play an important role in assessing and managing newborn jaundice. The most

referenced guidelines are the American Academy of Pediatrics (AAP) Guidelines (2004 and 2022), the National Institute for Health and Care Excellence (NICE) Guidelines, and the Canadian Pediatric Society (CPS) Guidelines [2, 37–40]. These CPGs include recommendations for universal or selective bilirubin screening, treatment thresholds for PT and ET, and the follow-up of infants with jaundice. Many countries have adopted or only slightly adapted the AAP guidelines without taking into account the specific underlying methodology of TBil measurement.

Discussion

In 1999, the 'Bhutani nomogram', a tool for universal TBil screening in term and late preterm infants was published [37]. This nomogram consists of various TBil–based curves and four risk zones. The risk zones predict the likelihood of a subsequent TBil result falling into the high-risk zone of the nomogram, providing a clinical interpretation of the risk of developing severe hyperbilirubinemia. It has been instrumental in guiding the decision-making process on neonatal jaundice follow-up care in many newborns [37]. The Bhutani curve was originally constructed with TBil measurements from a diazo assay on the Hitachi 747 platform [37]. Whilst there were efforts to ensure accuracy of TBil measurement for the Bhutani curve, inter- and intra-laboratory variability in TBil measurements already existed at that time [41, 42]. Consequently, TBil test results obtained on different platforms may lead to different risk zone assignments and corresponding follow-up policy to prevent imminent severe hyperbilirubinemia. Although recently updated with data from 15 years of universal bilirubin screening, the '1999' screening nomogram continues to be in use in countries outside the USA [40]. The newly revised hour-specific nomogram includes 140 times more infants compared to the previous nomogram, slightly increased the TBil-defined risk thresholds, and added the influence of neurotoxic risk factors [43] The update does not take into account the difference in TBil measurement methods; nearly all of the TBil measurements were done on Beckman-Coulter instruments which use a variation of the diazo method.

The hour-specific treatment nomograms of CGPs are not based on a single specific measurement method of bilirubin [39]. In fact, treatment thresholds vary between guidelines, and countries, reflecting the lack of evidence. The threshold TBil concentration at which treatment is initiated is largely based on observational studies that assess TBil concentrations associated with an increased risk of neurodevelopmental impairment. These studies are derived from a variety of populations, and hospitals using different measurement platforms and methods, whereas fixed medical decision limits require harmonized results between methods. With the advent of new data since the 2004 AAP guidelines, the graphs for initiation of PT and ET have been updated and are described in the 2022 AAP guidelines [2]. In all CGPs, PT is the treatment of choice if the TBil exceeds the corresponding PT threshold based on the newborn's gestational age and specific risk factors. Additionally, PT may be considered even when the TBil is below the threshold if the rate of TBil increment is higher than expected. When single-sided PT with normal intensity fails to reduce TBil, escalation of care, such as double-sided PT, an increase in PT intensity and even ET, may be considered. An ET is an invasive procedure initiated when TBil is higher than predefined TBil values, often ≥ 85 μmol/L above the PT threshold [2], or in cases of acute bilirubin encephalopathy regardless of TBil test results. PT is usually discontinued when the TBil is well below the hour-specific treatment threshold, usually ≥35–50 μmol/L lower, to prevent the risk of rebound hyperbilirubinemia [44], again underlining the importance of any measurement of TBil being close to the true value.

It is known that differences in treatment thresholds result in variations in PT rates: up to 11 % of newborns born at ≥35 weeks' gestation and up to 97 % of extremely low birth weight infants (birthweight <1,000 g) require PT for hyperbilirubinemia [11, 12, 38, 41, 45]. But even when similar screening or treatment nomograms are followed, disparity in neonatal jaundice management may occur due to lack of comparability and inaccuracy of bilirubin assays from different manufacturers. Imprecision within instrument peer groups is typically less than 4 %. And while many peer group means are within 10 % of the target value, others exceed this, resulting in variation of care [46].

Ultimately, biased TBil test results will influence the overall burden of TBil determinations, PT utilization, hospital visits and admissions, and healthcare costs, potentially affecting patient outcomes and family well-being. Whilst the International Consortium for Harmonization of Clinical Laboratory Results considers that an adequate reference measurement system is in place for TBil, they do not specifically consider the intended use for neonatal care [47]. Therefore, we consider improvement of TBil standardization an urgent and important matter.

Q6. What is the impact of bias in neonatal bilirubin measurement on clinical decisions?

Biased TBil test results may have profound implications on decisions in neonatal jaundice management. The TBil concentration informs the clinician not only on the need for treatment, but also on follow-up measurements of TBil,

and eventually escalation of care, and additional relevant laboratory parameters to detect any underlying etiology of hemolysis [48]. In analogy to unnecessary treatment, unnecessary blood collection procedures should be avoided, because these are associated with pain, risk of infection and, in preterm infants, with abnormal cognitive development [49].

Discussion

So far, only a few studies have described differences between bilirubin measurement methods leading to substantially different clinical decisions in neonatal care (Table 1). A national study in 36 French laboratories with 11 measurement platforms reported on accuracy in a reference standard, control and patient samples, the maximal TBil concentration being 315 µmol/L [34]. Whereas variation in TBil concentration <150 µmol/L were considered acceptable, variation in TBil values above 150 µmol/L were 10 % or even higher. The authors proposed using a correction factor for TBil test results >150 µmol/L for particular instruments to prevent undertreatment. Australian colleagues studied interlaboratory variability of 11 neonatal samples with TBil concentrations between 75 and 360 µmol/L across seven platforms in four laboratories [6]. Mean differences ranged from −12 and +20 % relative to a particular platform, resulting in altered screening and treatment decisions in three and four of 11 individual neonatal samples, respectively. Canadian researchers studied analytical differences across four different instruments for 40 adult samples as well as the impact of switching TBil assays on neonatal hyperbilirubinemia risk classification and PT rates for >65,000 Canadian neonates [50]. Mean differences of up to 17 % were observed between two platforms. These differences in TBil concentrations resulted in PT rate increases of up to 22 % in the neonate cohort. Subsequent switching of platforms led to reductions in PT rates.

Table 1: Recent published studies demonstrating the clinical impact of neonatal total bilirubin concentration differences.

Authors, country, year	Calibrator or reference material	Comparator assay or RMP	Total bilirubin, µmol/L	Mean difference absolute, µmol/L or %	CV, %	Effect on clinical decision
Mailloux, France, 2020 [34]	Calibrator: BE1449 LNE, France[a]	Proposed RMP using extinction coefficient (Hannover) [29]	Adult: <50 (min 7.5) Neonatal: <150; 150–250; > 250 (max 315) Control: 25 (C1) and 261.9 (C2)	Adult, neonatal (>150): +30 µmol/L to −25 µmol/L Control: +35 µmol/L to −35 µmol/L	Adult, neonatal (>150): <8 %[b] Interlab and inter-technique CV 15 % (C1) and 7 % (C2)	PT according to TBil on one platform, but not according to another (not otherwise specified)
Thomas, Australia, 2022 [6]	EQA value assigned by Reference Standards Laboratory, Children's Wisconsin, WI, USA	Beckman AU5822 bilirubin assay (traceable to NIST SRM 916a)	Neonatal: 75 to 360 EQA[c]: 80–320	Neonatal: −12 to +20 % EQA: −9 to +20 %[d]; >15 % (within-method bias range) in 8 out of 20 instrument – sample combinations	0–2.4 %	Different treatment decisions in 3–4 of 11 samples
Kittanakom, Canada, 2023 [50]	NIST SRM 916a	Beckman AU680	Adult: <60 Neonatal: 61–253.7[e]	Adult (<60 µM): +7.2 % and +17.2 % Neonatal: −25 to +15 %[f]	Not specified	PT rate varied: 8.3–22.4 %
Oostendorp, Netherlands, 2023 [7]	NIST SRM 916	AACC RMP according to Doumas [28]; for EQA samples: proposed RMP using extinction coefficient (Hannover) [29]	Adult: up to 550 Neonatal: up to 520 Control (EQA)[g] 233–435	Adult: 3.1 µmol/L (3.7 %) Neonatal: 34.4 µmol/L (17.4 %) Control (EQA): −3.6 to +20.2 %)	5.2–16 % (mean within method CV for EQA samples)	PT rate increased 3-fold

[a]LNE, Laboratoire National de Métrologie et d'Essais. [b]CV for all tested platforms without Roche-Cobas platform. [c]Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP) samples. [d]Derived from Figure 5, reference [6]. [e]Lowest and highest quartile of TBil. [f]Mean differences in population of neonates at 48 h postnatal age (n=65,719).[g]External quality assessment (EQA) samples provided by the Dutch EQA Organization in Medical Laboratories (Stichting Kwaliteitsbewaking Medische Laboratoria, SKML, Nijmegen, The Netherlands). AACC, American Association for Clinical Chemistry.

Finally, a recent Dutch study revealed a 3-fold increase in PT when switched from a diazo to a vanadate oxidase method within the same laboratory [7]. The mean differences of the vanadate vs. the previously used diazo method was much higher for neonatal samples when compared with adult ones: ≈+17 % vs. ≈+4 %, respectively. External quality assessment (EQA) data analysis of six samples supplemented with unconjugated bilirubin (range 233–435 μmol/L) from 66 laboratories (131 analyzers from 9 manufacturers) showed that the bias of commercial TBil methods varied from ≈−4 to ≈+20 %. In addition to these studies, an Italian case report of a preterm infant with (un)conjugated hyperbilirubinemia highlighted discrepant TBil measurements up to 244 μmol/L between two platforms [51]. Differences in results from different platforms are clearly altering some clinical outcome decisions and authors of these publications have highlighted that working to resolve this needs to be a priority.

Immediate needs

This section presents the clinical needs for neonatal bilirubin measurement and reflects the position of the IFCC WG-NB [8]. Four immediate needs have been identified:

Need 1 – Traceability chain – align methods to give the same result

The different potential measurands and different measurement methods for neonatal bilirubin mean that the metrological traceability of these measurements is complex. In a recent study, Mailloux et al. showed that recalibration of bilirubin assays with commutable calibrators would substantially improve result harmonization and accuracy, advocating for proceeding with a standardization program [34]. Now that the NIST SRM 916b bilirubin pure material is available for calibration, ideally this would be used as the higher order calibrant to ensure traceability to the SI. One question however is whether the JCTLM listed Doumas method is in fact an operationally defined method, with the measurand dependent upon specific procedure conditions. The 2018 version of this method [29] utilizes a molar extinction coefficient in place of a calibrators and would certainly be typically considered as operationally defined and has been described as such [1]. Traceability for both the traditional Doumas method and the 2018 version thus needs to be considered. The use of a conventional quantity value (absorptivity coefficient) in the 2018 method with no uncertainty does not facilitate achieving SI-traceability but may be used, when suitable, in the absence of the SRM. In reality, it may be that ensuring that these methods are fit for purpose rather than SI traceable at this point in time, should be the priority over formal standardization, which may be a longer-term goal.

Importantly, there are some points for further consideration and discussion:

(a) The proposed extinction coefficient reference value (without stated uncertainty) was determined through an interlaboratory study consensus for an operationally defined measurand.
(b) The clinical community may choose to adopt a conventional quantity value (absorptivity coefficient) if deemed fit for purpose; however, lack of measurement system calibration could lead to insufficient accounting for measurement variation (different equipment, minor variations in procedures, laboratories, etc.).
(c) Disregard for uncertainty in the proposed extinction coefficient based on a consensus value negates the possibility to claim traceability to the SI.
(d) Traceability of measurements using this proposed extinction coefficient would be to the consensus-derived absorbance value of NIST SRM 916a bilirubin (not NIST SRM 916b), corrected for impurities, determined via the RMP.

In considering the traceability of measurements, the measurand should also be considered to ensure the most fit for purpose measurand for current clinical needs is focused on and that any differences in measurands across platforms is acknowledged and taken into account.

Need 2 – Commutable EQA materials

As described above, commutability of EQA materials is rarely evaluated and it is difficult to source sufficient number and volume of clinical specimens obtained from neonates with endogenously elevated TBil concentrations. Therefore, accuracy and harmonization of TBil assays cannot be documented according to the best practices, which prevents conducting optimal post-market-surveillance.

Improved availability of commutable EQA materials will make it possible to properly document the state of the art in harmonization and accuracy of results provided by the different commercially available assays for TBil [52]. This will help identify assays in need for improvement and aid in determining whether such improvements should concern specificity and/or calibration. This will guide assay manufacturers to make the appropriate changes. Producers of

CRMs and EQA materials will also gain information regarding the characteristics of materials needed to calibrate TBil assays and monitor their performance. After the necessary improvements, if any, have been made by IVD producers, commutable EQA materials will once again be needed to verify the effectiveness of the changes that were made. This will also help in monitoring the performance and comparability of TBil assays on a regular basis to ensure that they are fit for purpose in the context of neonatal care.

Therefore, accuracy-based neonatal EQA programs relying on materials with RMP target values and properly demonstrated commutability should be propagated. In order to properly assess accuracy and harmonization across end user groups [53–55] efforts are currently being undertaken by one group using neonatal material-based control samples and properly characterize their commutability in a study occurring in 2024 (personal communication MvB).

Need 3 – Review fitness for purpose of current reference measurement procedures

The current JCTLM-listed RMPs for bilirubin are based on spectrophotometry. In addition, some manufacturers are using the 2018 "calibrant-free" variant as an RMP. A review of the robustness of extinction coefficients and also the appropriateness of spectrophotometry based RMPs should be considered, particularly considering advances in mass spectrometry-based techniques for many clinical RMPs.

Many more recently developed RMPs for other target measurands are based on mass spectrometry principles, often in tandem with separation first by gas chromatography (GC) or liquid chromatography (LC). Several mass spectrometry-based methods for research applications of bilirubin have been developed. Putluru et al. developed a novel LC-MS/MS method for estimating glucuronides in *in vitro* assays to study glucuronidation kinetics [56]. This method is a qualitative/quantitative approach that utilizes a UV to MS correction factor to estimate concentrations of mono- and di-glucuronides based on the assumption that the UV response for the conjugated forms are similar to the unconjugated form of bilirubin. Another novel LC-MS/MS method was developed, utilizes mesobilirubin as an internal standard, to determine concentrations of unconjugated bilirubin and the photoisomer, Z-lumirubin, in serum after phototherapy [57].

Looking ahead to the possible development of an MS-based RMP, it seems most current resources for such a method are focused on unconjugated bilirubin. Ideally, a stable isotope labeled internal standard would be used for quantification and some commercial sources currently advertise availability of bilirubin-d_4. However, it would need to be determined whether currently available labeled materials are of sufficient chemical and isotopic purity to serve as internal standards in an MS-based method. NIST SRM 916b is currently available as a high purity calibration material comprised of unconjugated bilirubin. There is a lack of authentic standards for bilirubin glucuronides due to stability issues during synthesis and storage [56]. The previously cited methods support advances in separating various forms of bilirubin by LC with MS/MS detection. However, the reported precision is rather high compared to those typically associated with LC-MS/MS RMPs. Since any MS-based method will target specific masses and/or mass transitions for target analytes, the measurand(s) for bilirubin determination will need to be clearly defined to ensure the proper target analytes are incorporated into candidate RMP development.

Need 4 – Continued cooperation and support from diagnostic manufacturers

Fortunately, through continued cooperation and support with diagnostic manufacturers, improvement of several MPs has been achieved over the years: the interlaboratory spread of results has decreased compared to the situation at the start of this millennium [1, 58]. Also, the production of reliable bilirubin calibrators is demonstrated by the recent availability of NIST SRM 916b into the calibration chain.

IVD manufacturers are primarily interested in providing accurate, safe and traceable test results for their customers, laboratories and patients. They also see the advantage for patients to achieve comparable results worldwide, irrespective of the provider. Harmonization can be achieved by using a reference material and RMP to establish results with metrological traceability. Commutable EQA materials will allow us to better assess the comparability of test results between manufacturers [6]. It is not only the patients' interest, but also in the interest of the manufacturers of *in vitro* diagnostic medical devices. It would be all the more gratifying if this comparability could be established not only in the central laboratory, but also on point of care applications across all platforms for TBil.

The IVD manufacturers are operating in a globalized world and have to meet the requirements of global markets. The availability of certified reference materials and RMPs has implications for IVD manufacturers. Reference materials have

important prerequisites: they must be accessible worldwide, and above all, recognized for use by local regulatory authorities worldwide. The latter is of course also true for reference methods. The material shall be available for standardization activities of the manufacturers in a sufficiently large volume of constant and appropriate quality at an appropriate price and it shall be ensured that, when the material is updated, the quality is maintained at the same concentration.

TBil in blood is a more complex and variable analyte than most of the other well-defined analytes, being it is a mixture of different subspecies. It is a challenge to reflect this with a reference material or RMP. Consequently, it must be proven that the chosen reference system maintains calibration traceability for all sample types as – in this particular case – for newborn and adult bilirubin samples and at all expected analyte concentrations.

When a new reference material or RMP is introduced, they must be assessed for their influence on manufacturer's method results or on the concentration of established decision limits. It would also be necessary to examine the influence of a shift in reference values on the established treatment guidelines. If new reference value studies become necessary, these are complex and expensive. Regarding bilirubin in newborns, reference value studies have an additional more complex dimension in terms of sample availability and the associated ethical issues. Consideration should be given towards applying indirect methods for calculating new reference intervals in children and in adults.

In principle, IVD manufacturers operate within a tight corset of worldwide and locally specific regulatory requirements. Every change to established, approved tests results in an extensive approval process. This can be equal to the new registration of a new test. Thus, changing the traceability may be associated with high costs and a high time expenditure. It is therefore important to understand that even after agreement on a reference system, a lengthy change process for its implementation will follow, the duration of which is measured in years.

From the IVD Manufacturer's point of view, it is highly welcomed that a higher-level organization, i.e., the IFCC, is willing to take on a coordinating role to bring together all stakeholders, such as manufacturers, authorities, standardization institutions and users at an early stage, thereby facilitating communication and coordination among themselves and exploiting synergies.

Room for improvement

The IFCC Neonatal Bilirubin Working Group has been recently launched to inform strategies for improving alignment of neonatal bilirubin for current and emerging tests. Unbiased and precise TBil measurements are essential for uniform clinical decisions to provide PT and to escalate care, to determine the frequency of monitoring TBil concentrations and to inform the clinician when it is safe to lower PT intensity or to discontinue PT without risk of imminent bilirubin neurotoxicity. The need for enhancement in bilirubin measurement is acknowledged by manufacturers, authorities, standardization institutions and users. Important steps for improvement include equivalence of reference materials and RMPs (including consideration of the measurand and appropriate calibration approaches and traceability) and the use of commutable materials in EQA programs to periodically assess the state of standardization of bilirubin measurements in neonatal care. Finally, for the overall benefit of the baby, it is essential as part of this improvement journey for TBil that we establish links and collaborate with the clinical societies who develop the treatment recommendations.

Acknowledgments: Improvements to TSB measurements require the cooperation of our manufacturing partners. We wish to acknowledge the contributions from participating industry representatives for their insight and information. Clearly, small changes to the measurement process can be significant challenges and their information has helped to educate our group and elucidate potential solutions. We hope to continue this working relationship until the clinical need for appropriate TSB measurements has been achieved. Michael A. Nelson (NIST) is acknowledged for his suggestions and expert review in regard to metrological traceability.

Research ethics: Not applicable as this is a position paper and does not contain any original research.

Informed consent: Not applicable as this is a position paper and does not contain any original research.

Author contributions: All authors are full members of the IFCC neonatal bilirubin working group and collectively developed the concepts of the paper and reviewed the entire content. Specific activities included: CH and MT developed the clinical content. RG developed figure 1, drafted the style of the paper and coordinated editions; JC, LM and VD as metrologists lead the NIST/JCTLM discussions; SL provided insight into the current reference method listed on JCTLM; MVB and AM provided information on the current performance between manufacturers; and MS provided the manufacturer's perspective. Together all authors developed the key questions and agreed on the immediate needs, which forms the direction of the WG-NB. The authors

have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: All authors are full members of the IFCC neonatal bilirubin working group.

Research funding: None declared.

Data availability: Not applicable as this is a position paper and does not contain any original data.

Disclaimer: Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

References

1. Hulzebos CV, Vitek L, Coda Zabetta CD, Dvořák A, Schenk P, van der Hagen EAE, et al. Diagnostic methods for neonatal hyperbilirubinemia: benefits, limitations, requirements, and novel developments. Pediatr Res 2021;90:277–83.
2. Kemper AR, Newman TB, Slaughter JL, Maisels MJ, Watchko JF, Downs SM, et al. Clinical practice guideline revision: management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. Pediatrics 2022;150. https://doi.org/10.1542/peds.2022-058859.
3. Maisels MJ, Bhutani VK, Bogen D, Newman TB, Stark AR, Watchko JF. Hyperbilirubinemia in the newborn infant > or =35 weeks' gestation: an update with clarifications. Pediatrics 2009;124:1193–8.
4. Mather A. Reliability of bilirubin determinations in icterus of the newborn infant. Pediatrics 1960;26:350–4.
5. Cobbaert C, Weykamp C, Hulzebos CV. Bilirubin standardization in the Netherlands: alignment within and between manufacturers. Clin Chem 2010;56:872–3.
6. Thomas DH, Warner JV, Jones GRD, Chung JZY, Macey DJ, Screnci A, et al. Total bilirubin assay differences may cause inconsistent treatment decisions in neonatal hyperbilirubinaemia. Clin Chem Lab Med 2022; 60:1736–44.
7. Oostendorp M, Ten Hove CH, van Berkel M, Roovers L. A significant increase in the incidence of neonatal hyperbilirubinemia and phototherapy treatment due to a routine change in laboratory equipment. Arch Pathol Lab Med 2024;148:e40–7.
8. International Federation of Clinical Chemistry and Laboratory Medicine. IFCC Working Group Neonatal Bilirubin; 2023. Available from: https://ifcc.org/ifcc-emerging-technologies-division/etd-working-groups/wg-neonatal-bilirubin-wg-nb-in-cooperation-with-ifcc-scientific-division-sd/.
9. Le Pichon JB, Riordan SM, Watchko J, Shapiro SM. The neurological sequelae of neonatal hyperbilirubinemia: definitions, diagnosis and treatment of the kernicterus spectrum disorders (KSDs). Curr Pediatr Rev 2017;13:199–209.
10. Newman TB, Wu YW, Kuzniewicz MW, Grimes BA, McCulloch CE. Childhood seizures after phototherapy. Pediatrics 2018;142. https://doi.org/10.1542/peds.2018-0648.
11. Brown AK, Kim MH, Wu PY, Bryla DA. Efficacy of phototherapy in prevention and management of neonatal hyperbilirubinemia. Pediatrics 1985;75:393–400.
12. Morris BH, Oh W, Tyson JE, Stevenson DK, Phelps DL, O'Shea TM, et al. Aggressive vs. conservative phototherapy for infants with extremely low birth weight. N Engl J Med 2008;359:1885–96.
13. Lamola AA. A pharmacologic view of phototherapy. Clin Perinatol 2016; 43:259–76.
14. Steiner LA, Bizzarro MJ, Ehrenkranz RA, Gallagher PG. A decline in the frequency of neonatal exchange transfusions and its effect on exchange-related morbidity and mortality. Pediatrics 2007;120: 27–32.
15. Doumas BT, Wu TW. The measurement of bilirubin fractions in serum. Crit Rev Clin Lab Sci 1991;28:415–45.
16. Thaler M, Luppa PB, Schlebusch H. Bilirubin measurement – an updated survey. LaboratoriumsMedizin 2008;32. https://doi.org/10.1515/jlm.2008.005et.
17. Ngashangva L, Bachu V, Goswami P. Development of new methods for determination of bilirubin. J Pharm Biomed Anal 2019;162:272–85.
18. Kiuchi S, Ihara H, Osawa S, Ishibashi M, Kinpara K, Ohtake K, et al. A survey of the reactivity of in vitro diagnostic bilirubin reagents developed in Japan using artificially prepared bilirubin materials: a comparison of synthetic delta, unconjugated, and taurine-conjugated bilirubin. Ann Clin Biochem 2021;58:563–71.
19. Kawamoto S, Koyano K, Ozaki M, Arai T, Iwase T, Okada H, et al. Effects of bilirubin configurational photoisomers on the measurement of direct bilirubin by the vanadate oxidation method. Ann Clin Biochem 2021;58:311–17.
20. Wu TW, Dappen GM, Spayd RW, Sundberg MW, Powers DM. The Ektachem clinical chemistry slide for simultaneous determination of unconjugated and sugar-conjugated bilirubin. Clin Chem 1984;30: 1304–9.
21. Mreihil K, Madsen P, Nakstad B, Benth J, Ebbesen F, Hansen TW. Early formation of bilirubin isomers during phototherapy for neonatal jaundice: effects of single vs. double fluorescent lamps vs. photodiodes. Pediatr Res 2015;78:56–62.
22. Osawa S, Sugo S, Yoshida T, Yamaoka T, Nomura F. An assay for separating and quantifying four bilirubin fractions in untreated human serum using isocratic high-performance liquid chromatography. Clin Chim Acta 2006;366:146–55.
23. Adachi Y, Inufusa H, Yamashita M, Kambe A, Yamazaki K, Sawada Y, et al. Clinical application of serum bilirubin fractionation by simplified liquid chromatography. Clin Chem 1988;34:385–8.
24. Watchko JF, Tiribelli C. Bilirubin-induced neurologic damage – mechanisms and management approaches. N Engl J Med 2013;369: 2021–30.
25. Harpavat S, Finegold MJ, Karpen SJ. Patients with biliary atresia have elevated direct/conjugated bilirubin levels shortly after birth. Pediatrics 2011;128:e1428–33.
26. Joint Committee for Traceability in Laboratory Medicine (JCTLM). Database of higher-order reference materials, measurement methods/procedures and services. Sevres Cedex, France: International Bureau of Weights and Measures (BIPM).
27. Nelson MA, Brown TJ, Lang BE, Mulloor J, Sieber JR, Toman B, et al. Certification of Standard Reference Material 916b Bilirubin. Gaithersburg, MD.: National Institute of Standards and Technology, 2021, NIST Special Publication (SP), 260–212.
28. Doumas BT, Kwok-Cheung PP, Perry BW, Jendrzejczak B, McComb RB, Schaffer R, et al. Candidate reference method for determination of

total bilirubin in serum: development and validation. Clin Chem 1985; 31:1779–89.
29. Klauke R, Kytzia HJ, Weber F, Grote-Koska D, Brand K, Schumann G. Reference measurement procedure for total bilirubin in serum re-evaluated and measurement uncertainty determined. Clin Chim Acta 2018;481:115–20.
30. Lo SF, Kytzia HJ, Schumann G, Swartzentruber M, Vader HL, Weber F, et al. Interlaboratory comparison of the Doumas bilirubin reference method. Clin Biochem 2009;42:1328–30.
31. RELA external quality control for reference laboratories total bilirubin interlaboratory comparison participants (2020–2022). http://www.dgkl-rfb.de:81/index.shtml [Accessed 8 Nov 2023].
32. Coughtrie MW, Burchell B, Leakey JE, Hume R. The inadequacy of perinatal glucuronidation: immunoblot analysis of the developmental expression of individual UDP-glucuronosyltransferase isoenzymes in rat and human liver microsomes. Mol Pharmacol 1988;34:729–35.
33. Kawade N, Onishi S. The prenatal and postnatal development of UDP-glucuronyltransferase activity towards bilirubin and the effect of premature birth on this activity in the human liver. Biochem J 1981;196: 257–60.
34. Mailloux A, Cortey A, Delatour V, Poupon C, Rota M, Schmitt F, et al. Analytical and clinical guidelines on neonatal bilirubinemia. Ann Biol Clin 2020;78:383–97.
35. Lo S, Jendrzejczak B, Doumas BT. Bovine serum-based bilirubin calibrators are inappropriate for some diazo methods. Clin Chem 2010; 56:869–72.
36. Thomas M, Greaves RF, Tingay DG, Loh TP, Ignjatovic V, Newall F, et al. Current and emerging technologies for the timely screening and diagnosis of neonatal jaundice. Crit Rev Clin Lab Sci 2022;59:332–52.
37. Bhutani VK, Johnson L, Sivieri EM. Predictive ability of a predischarge hour-specific serum bilirubin for subsequent significant hyperbilirubinemia in healthy term and near-term newborns. Pediatrics 1999;103:6–14.
38. National Institute for Health and Clinical Excellence. Jaundice in newborn babies under 28 days Clinical Guideline. NICE; 2010 (Updated 2023).
39. Subcommittee on Hyperbilirubinemia. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. Pediatrics 2004;114:297–316.
40. Guidelines for detection, management and prevention of hyperbilirubinemia in term and late preterm newborn infants (35 or more weeks' gestation) – summary. Paediatr Child Health 2007;12: 401–18.
41. Schreiner RL, Glick MR. Interlaboratory bilirubin variability. Pediatrics 1982;69:277–81.
42. Vreman HJ, Verter J, Oh W, Fanaroff AA, Wright LL, Lemons JA, et al. Interlaboratory variability of bilirubin measurements. Clin Chem 1996; 42:869–73.
43. Bahr TM, Henry E, Christensen RD, Minton SD, Bhutani VK. A new hour-specific serum bilirubin nomogram for neonates ≥35 weeks of gestation. J Pediatr 2021;236:28–33.e1.
44. Chang PW, Newman TB. A simpler prediction rule for rebound hyperbilirubinemia. Pediatrics 2019;144. https://doi.org/10.1542/peds.2018-3712.
45. Chang PF, Lin YC, Liu K, Yeh SJ, Ni YH. Identifying term breast-fed infants at risk of significant hyperbilirubinemia. Pediatr Res 2013;74:408–12.
46. Lo S, Hubbard L, Koranda M, Doumas B. Standardization of neonatal bilirubin testing and the lack of a primary reference material. In AACC Annual Scientific Meeting 2019: Anaheim, CA, USA. Poster: (B-176).
47. International Consortium for Harmonization of Clinical Laboratory Results; 2023. https://www.harmonization.net/measurands/ [Accessed 14 Dec 2023].
48. Balasundaram P, Campbell D. Monitoring neonatal bilirubin after discharge. Pediatr Rev 2023;44:110–13.
49. Vinall J, Miller SP, Bjornson BH, Fitzpatrick KP, Poskitt KJ, Brant R, et al. Invasive procedures in preterm children: brain and cognitive development at school age. Pediatrics 2014;133:412–21.
50. Kittanakom S, Shea J, Leung F, Ly V, Panchadcharadevan S, Kathuria S, et al. Impact of switching total bilirubin assays on the classification of neonates at high risk for hyperbilirubinemia. Clin Chem Lab Med 2023; 61:e175–8.
51. Colombo G, Szoke D, Aloisio E, Cavigioli F, Dolci A, Panteghini M. Clinicians at crossroads for a dangerous interference in neonatal bilirubin determination at the point-of-care. Clin Chem 2022;68:887–91.
52. Miller GW, Myers GL, Gantzer LM, Kahn SE, Schönbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. Clin Chem 2011;57:1108–17.
53. Delatour V, Clouet-Foraison N, Jaisson S, Kaiser P, Gillery P. Beware of noncommutability of external quality assessment materials for hemoglobin A1c. Clin Chem 2020;66:390–1.
54. Jones GRD, Delatour V, Badrick T. Metrological traceability and clinical traceability of laboratory results – the role of commutability in external quality assurance. Clin Chem Lab Med 2022;60:669–74.
55. Badrick T, Miller WG, Panteghini M, Delatour V, Berghall H, MacKenzie F, et al. Interpreting EQA-understanding why commutability of materials matters. Clin Chem 2022;68:494–500.
56. Putluru SP, Matta MK, Ahire D, Subramanian M, Sinz M, Mandlekar S. A novel liquid chromatography tandem mass spectrometry method for the estimation of bilirubin glucuronides and its application to in vitro enzyme assays. Drug Metab Lett 2017;10:264–9.
57. Jašprová J, Dvořák A, Vecka M, Leníček M, Lacina O, Valášková P, et al. A novel accurate LC-MS/MS method for quantitative determination of Z-lumirubin. Sci Rep 2020;10:4411.
58. Lo SF, Jendrzejczak B, Doumas BT. Laboratory performance in neonatal bilirubin testing using commutable specimens: a progress report on a College of American Pathologists Study. Arch Pathol Lab Med 2008;132: 1781–5.