ABSTRACT

Introduction  Severe neonatal hyperbilirubinaemia can place a neonate at risk for acute bilirubin encephalopathy and kernicterus spectrum disorder. Early diagnosis is essential to prevent these deleterious sequela. Currently, screening by visual inspection followed by laboratory-based bilirubin (LBB) quantification is used to identify hyperbilirubinaemia in neonates cared for at home in the Netherlands. However, the reliability of visual inspection is limited. We aimed to evaluate the effectiveness of universal transcutaneous bilirubin (TcB) screening as compared with visual inspection to: (1) increase the detection of hyperbilirubinaemia necessitating treatment, and (2) reduce the need for heel pricks to quantify bilirubin levels. In parallel, we will evaluate a smartphone app (Picterus), and a point-of-care device for quantifying total bilirubin (Bilistick) as compared with LBB.

Methods and analysis  We will undertake a multicentre prospective cohort study in nine midwifery practices across the Netherlands. Neonates born at a gestational age of 35 weeks or more are eligible if they: (1) are at home at any time between days 2 and 8 of life; (2) have their first midwife visit prior to postnatal day 6 and (3) did not previously receive phototherapy. TcB and the Picterus app will be used after visual inspection. When LBB is deemed necessary based on visual inspection and/or TcB reading, Bilistick will be used in parallel. The coprimary endpoints of the study are: (1) hyperbilirubinaemia necessitating treatment; (2) the number of heel pricks performed to quantify LBB. We aim to include 2310 neonates in a 2-year period. Using a decision tree model, a cost-effectiveness analysis will be performed.

Ethics and dissemination  This study has been approved by the Medical Research Ethical Committee of the Erasmus MC Rotterdam, Netherlands (MEC-2020-0618). Parents will provide written informed consent. The results of this study will be published in peer-reviewed journals.

Trial registration number  Dutch Trial Register (NL9545).

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study is the first prospective multicentre cohort study to assess the diagnostic accuracy, user convenience and (cost-)effectiveness of universal transcutaneous bilirubin screening in neonates cared for at home.

⇒ In parallel, we evaluate a novel smartphone app (Picterus) providing an optical bilirubin estimate, and a recently developed point-of-care device (Bilistick) for measurement of measure whole blood bilirubin.

⇒ A limitation is that the study is observational and not blinded.

INTRODUCTION

Neonatal jaundice, caused by elevated levels of unconjugated bilirubin, is a common phenomenon during the first few days of life. This unconjugated hyperbilirubinaemia is generally transient and considered benign. However, when very high bilirubin levels are left untreated, non-protein bound bilirubin may pass the blood–brain barrier and damage the neonate’s developing brain. Hence, severe neonatal hyperbilirubinaemia may result in acute bilirubin encephalopathy. Neonates with acute bilirubin encephalopathy urgently require treatment as they are prone to develop a chronic phase with irreversible brain damage, known as kernicterus spectrum disorder. Kernicterus spectrum disorder consists of a variety of severe pathologic conditions including motor, cognitive and auditory disorders with life-long sequelae. Phototherapy is often successful in reducing bilirubin to non-hazardous levels, and may also revert early stages of acute bilirubin encephalopathy. When phototherapy is unsuccessful in treating severe neonatal hyperbilirubinaemia, one or more exchange transfusions are necessary to lower circulating bilirubin levels. However, exchange transfusions are invasive, high-risk procedures that should be avoided if possible. Thus, timely
recognition of neonates with imminent severe hyperbilirubinaemia is essential.

In the Netherlands, the majority of neonates born at a gestational age of 35 weeks or more are eventually cared for at home, irrespective of the place of birth. Daily postpartum care at home is provided by maternity care assistants during the first week after birth, who are supervised by community midwives. The community midwife visits mother and neonate at home usually at least three times in the first week after birth. As per the current multidisciplinary Dutch national guideline for identification and treatment of neonatal jaundice, visual inspection is used as first-line screening in neonates cared for at home. If potentially severe hyperbilirubinaemia is suspected based on visual inspection, the community midwife may decide to have bilirubin levels checked in the neonate’s blood, taken by the midwife or a specialised laboratory home service. Based on this laboratory-based bilirubin (LBB) level, the need for treatment of hyperbilirubinaemia is assessed using the nomogram of the Dutch national guideline, adapted from the American Academy of Pediatrics 2004 guideline. If the LBB level indicates the need for treatment, this is usually applied in-hospital.

Several studies have shown that visual inspection is not a reliable screening tool for neonatal hyperbilirubinaemia. In a significant proportion of neonates admitted from home with severe hyperbilirubinaemia and/or acute bilirubin encephalopathy, neonatal jaundice either went unnoticed or was misclassified by maternity care assistants, midwives and/or parents. Moreover, neonatal hyperbilirubinaemia is one of the most common indications for hospital (re)admission in the neonatal period. Hence, there is an urgent need for more effective approaches towards timely recognition of clinically relevant jaundice in neonates cared for at home.

Hospital-based screening programmes have been shown to be effective in preventing severe hyperbilirubinaemia via timely quantification of LBB or transcutaneous bilirubin (TcB). A Dutch randomised controlled trial in the hospital setting showed that selective TcB screening (ie, when visually jaundiced) reduced the need for heel pricks to quantify LBB by 38% as compared with only visual assessment. The effectiveness of universal TcB screening compared with only visual inspection in preventing severe hyperbilirubinaemia is currently being investigated in the STARSHIP trial in seven Dutch primary
care birth centres. The BEAT Jaundice @home study extends this work by focusing on screening and diagnosis of neonatal hyperbilirubinaemia in the home setting, and adding additional screening (ie, Picterus) and diagnostic (ie, Bilistick) tools in an attempt to improve early recognition and diagnosis of potentially severe neonatal hyperbilirubinaemia. Picterus is a smartphone application that provides a bilirubin reading based on photographs of the neonate’s skin overlying the sternum, and Bilistick is a point-of-care (POC) test for total bilirubin in whole blood.

In this prospective study, our main aim is to evaluate the effectiveness of universal TcB screening in neonates cared for at home to increase the timely detection of hyperbilirubinaemia necessitating treatment while reducing the need for heel pricks to quantify bilirubin levels in blood. Using a decision tree model, an additional cost-effectiveness analysis (CEA) will be performed. In parallel, we will evaluate the diagnostic accuracy, user convenience and (cost-)effectiveness of the Picterus app and Bilistick in the same population.

METHODS AND ANALYSIS
Study design and setting
We will conduct a prospective multicentre cohort study in nine community midwifery practices across two regions in the Netherlands. Three approaches for screening (ie, TcB and Picterus) and diagnosis (ie, Bilistick) of neonatal hyperbilirubinaemia in the home setting will be evaluated in parallel with standard care (ie, visual inspection followed by LBB quantification in case of jaundice). The study inclusions have started in July 2021, with an anticipated inclusion period of 2 years.

Participant eligibility
Neonates are considered eligible for inclusion if they: (1) are born at a gestational age of 35 weeks or more, (2) are at home at any time between days 2 and 8 of life and (3) have their first midwife visit at home prior to postnatal day 6. Neonates are excluded if they: (1) previously received phototherapy, or (2) have parents who are unable to understand the patient information sheet due to insufficient understanding of the Dutch language.

Recruitment
Participants will be recruited in the community midwifery practices by the midwives. Parents will be informed about the study at the regular antenatal care by the community midwife and/or on the first midwife visit following delivery. Written parental informed consent is obtained by the community midwife during the first postnatal visit at home.

Interventions
We will evaluate three universal screening tools in parallel: (1) visual inspection (standard care); (2) TcB (Draeger JM-105, Lübeck, Germany) and (3) a smartphone-based mobile health application (Picterus app, Trondheim, Norway). TcB and Picterus app will be applied in all included neonates directly following visual inspection (standard care) during each midwife home visits in the
following order: (1) visual inspection, (2) TcB and (3) Picterus app. The Picterus app is a novel screening tool for neonatal hyperbilirubinemia, and may be a cheaper alternative to the TcB tool depending on its performance. Moreover, we will evaluate two diagnostic tools: (1) LBB quantification (standard care) and (2) a hand-held POC device for quantifying total bilirubin in whole blood (Bilistick, Trieste, Italy). Bilistick will only be used when LBB quantification is indicated by the presence of visual jaundice and/or an elevated TcB reading. Its place is distinct from that of Picterus, which is a screening tool, whereas Bilistick is evaluated as a diagnostic tool. Figure 1 displays a flow chart with an overview of the timeline per participant. During each midwife visit, this participant flow is followed, except when treatment for neonatal hyperbilirubinemia is started.

Transcutaneous bilirubinometer (TcB screening)

Draeger JM-105 is a CE certified, transcutaneous bilirubinometer and will be applied at every midwife visit at home. Three measurements will be taken on the neonate’s sternum. The highest of the three TcB readings will be plotted on a customised version of the Dutch bilirubin nomogram (figure 2). The original nomogram (figure 3) consists of three different curves indicating the need for phototherapy based on LBB level, postnatal age and a set of risk factors for hyperbilirubinemia and neurotoxicity.7 If the TcB reading is above the phototherapy threshold or higher than the corresponding TcB level on this TcB nomogram,7 15 20 Participating midwives have experience in using the original LBB nomograms in everyday practice in the Netherlands and will be trained to use the customised TcB nomograms.

Smartphone-based mobile health application (Picterus app)

The Picterus reading involves three components: (1) Samsung S7 smartphone with the Picterus app, (2) a unique colour calibration card and (3) the Picterus server. A Picterus reading will also be taken at each midwife visit at home. For this purpose, the colour calibration card is placed on the neonate’s sternum and a collection of six photographs (three with and three without flash) is automatically taken of the card on the sternum. Only the colour calibration card and a small part of the skin of the neonate’s sternum will be visible on the photographs.17 The role of the colour calibration card is to have a fixed colour reference for the digital images and to calibrate the images from variations in illumination and optical variations between different smartphones. Images taken with the Picterus app are sent to the Picterus server, checked for quality and colour calibrated. Thereafter, a bilirubin level estimation will be provided based on a large database of simulated colours of neonate’s skin. The simulation with the best matching colour will be automatically identified through an algorithm. Bilirubin estimates are calculated for each image and an average value of the six images is calculated to get a final result. The Picterus app is currently under development and is expected to be CE certified and become commercially available.
in 2021. One clinical study has been published so far, mainly in neonates with lighter skin types. This study will be the first to evaluate Picterus in the home setting. To avoid treatment decisions being made based on the Picterus reading, community midwives will be blinded to its reading. Picterus diagnostic accuracy and user convenience will be analysed and determined in retrospect on finalisation of the project.

Hand-held point-of-care device for quantifying total bilirubin in whole blood (Bilistick)

Bilistick is a CE certified POC test for total bilirubin quantification in 35 µL of whole blood that has recently become commercially available. The Bilistick will be used by the community midwife in neonates requiring bilirubin quantification in blood, as indicated by visual inspection and/or elevated TcB reading (see figure 1). First, approximately 500 µL of blood will be taken and sent to a nearby laboratory for LBB quantification according to standard practice. Second, 35 µL (ie, one or two drops) of blood from the same heel prick is used to additionally quantify total bilirubin in whole blood using the Bilistick device. The result of the conventional LBB quantification will be used to determine the need for treatment, according to the current guidelines. From a safety point of view however, if the Bilistick bilirubin level—which will be available earlier than the LBB—exceeds the exchange transfusion threshold, the neonate will immediately be referred to a paediatrician at a nearby hospital pending the LBB result, in order to avoid delay in instituting treatment. No other decisions will be made based on the Bilistick reading. Bilistick might facilitate earlier treatment initiation. The device provides results rapidly and requires a smaller blood volume as compared with LBB.

**Outcome measures**

**Primary outcome**

The primary objective is to evaluate the effectiveness of universal TcB screening to increase the detection of neonates with hyperbilirubinaemia necessitating treatment compared with visual inspection, and at the same time decrease the number of heel pricks performed to quantify total bilirubin in blood, if TcB would replace visual inspection as a screening method. As such, there are two primary endpoints assessed at each time point for each neonate: (1) having a LBB above the treatment threshold, and (2) requiring a heel prick to determine LBB.

**Secondary outcomes**

Secondary outcomes are described in **table 1** and mainly relate to diagnostic accuracy parameters of the applied methods.

**Study procedures**

Community midwives visit the new family a number of times (usually three times) at home during the first week after delivery. The frequency of these home visits can also depend on the advice of the maternity care assistants.

| Table 1 | Secondary outcomes |
|---------|------------------|
| **Assessed in** | **Outcomes** |
| All participating neonates | Diagnostic accuracy of TcB:  
Correlation between TcB and LBB  
Bland-Altman plots of TcB vs LBB  
(including mean difference with SD and 95% CI of limits of agreement) |
| | Diagnostic accuracy of Picterus:  
Correlation between Picterus and TcB  
The proportion of failed readings using Picterus  
Added value of Picterus app compared with only visual inspection in picking up neonates with hyperbilirubinaemia requiring treatment, while reducing the need for heel pricks to quantify bilirubin in blood  
Bland-Altman plots of Picterus vs LBB  
(including mean difference with SD and 95% CI of limits of agreement)  
Picterus vs TcB  
Correlation between Picterus and TcB |
| Neonates requiring LBB | Diagnostic accuracy of Bilistick:  
Correlation between Bilistick and LBB  
Bland-Altman plots of Bilistick vs LBB (including mean difference with SD and 95% CI of limits of agreement)  
Number of true positives  
Number of true negatives  
Number of false positives  
Number of false negatives  
Sensitivity  
Specificity  
Positive predictive value  
Negative predictive value  
The proportion of failed readings using Bilistick  
Difference in time-to-test result between LBB and Bilistick (in min). |
| All participating neonates | Proportion of neonates having an LBB level above the phototherapy threshold  
Proportion of neonates who actually received phototherapy  
Proportion of neonates having an LBB level above the exchange transfusion threshold  
Proportion of neonates who actually received an exchange transfusion  
Potential cost-effectiveness of implementation of the novel methods in daily practice |

At each visit, the midwife will, according to usual practice, first assess and record whether or not the neonate is visually jaundiced. The duration of jaundice evaluation depends on the last home visit made by the midwife (which can be beyond day 8). Accordingly, the midwife will decide based on the visual inspection whether LBB quantification in blood is required. Within the current project, the midwife will subsequently quantify TcB, and apply the Picterus app in every participant. LBB quantification is then performed in any neonate who is either visually jaundiced, has a TcB above the threshold, or both. When blood is taken to quantify LBB, total bilirubin is also measured using the Bilistick. Importantly,
The midwife may never revert her initial decision to have LBB quantified in a visually jaundiced neonate based on a ‘negative’ TcB reading (ie, TcB reading below the threshold) to ensure safety of the participants, which is in accordance with the current Dutch guidelines.7

Midwives of all participating practices will be trained in the study procedures, data collection and use of the Draeger JM-105, the Picterus app, and the Bilistick device prior to the start of the inclusion to ensure safety and adequate use. The researchers will be available 24/7 by phone to address any issues that the midwives may come across while performing study procedures.

Data collection
Data collection will be performed digitally by the community midwife using the Picterus app. ABBR, Automated Auditory Brainstem Response.

The midwife can never make a decision to have LBB quantified on a ‘negative’ TcB reading (ie, TcB reading below the threshold) to ensure safety of the participants, which is in accordance with the current Dutch guidelines.7

Midwives of all participating practices will be trained in the study procedures, data collection and use of the Draeger JM-105, the Picterus app, and the Bilistick device prior to the start of the inclusion to ensure safety and adequate use. The researchers will be available 24/7 by phone to address any issues that the midwives may come across while performing study procedures.

**Table 2** Baseline characteristics

| Characteristics of (biologic) parents | Ethnicity (Caucasian, Asian, North African, African, Latin American, Hindu, Turkish, mixed) | Maternal blood group | Maternal rhesus factor |
|--------------------------------------|------------------------------------|---------------------|----------------------|
| Characteristics of pregnancy and birth | Gestational age at birth (weeks, days) | Parity | Gravity |
|                                       | Type of delivery (vaginal, caesarean section and instrumental) | Presence of irregular antibodies during pregnancy | Presence of risk factors for neonatal hyperbilirubinaemia (blood group antagonism, presence of other haemolytic disease and presence of birth asphyxia) |
| Characteristics of neonate            | Date of birth | Time of birth | Sex |
|                                       | Birth weight (in grams) | Blood group (if known) | Rhesus factor (if known) |
|                                       | Apgar score <5 at 5 min | |
| Characteristics of family members     | Sibling having received phototherapy or exchange transfusion for severe neonatal hyperbilirubinaemia | Presence of haemolytic disease | |

* At discretion of community midwives.

**Table 3** Data collection during home visits

| Assessed in | Collected data items |
|-------------|----------------------|
| All | Date and time of screening assessment |
| Participating neonates | Severity of jaundice based on visual inspection at the discretion of the community midwife and noted as either pink, slightly yellow, yellow (indicating reason for blood prick) |
| TcB reading: highest of three measurements |
| Comments/inconveniences TcB assessment |
| Encrypted code Picterus app reading |
| Comments/inconveniences Picterus app assessment |
| Decision made by community midwife—no extra measures taken, LBB+Bilistick quantification, contact with paediatrician |
| Neonates with an indication for LBB quantification | Decision for LBB quantification was based on—visual inspection, elevated TcB measurement, visual inspection and elevated TcB measurement, other |
| Date and time of Bilistick measurement |
| Result Bilistick (in µmol/L) |
| Comments/inconveniences/haemolytic Bilistick assessment |
| Person responsible for the blood transportation to the laboratory for LBB (partner, family member other than partner/friend, community midwife, other) |
| Hospital laboratory responsible for the LBB quantification |
| Date and time LBB quantification result communicated to midwife |
| Result LBB (in µmol/L) |

LBB, laboratory-based bilirubin; TcB, transcutaneous bilirubinometry.

**Table 4** Follow-up data collection

| Follow-up data collection | Hospital admission (yes/no) |
|---------------------------|----------------------------|
| Type of feeding of the neonate | More than 10% loss of birth weight within the first week of life (yes/no) |
| Number of (heel) pricks to determine bilirubin levels | Coombs test (if known) |
| Neonates admitted to the hospital | Duration of hospitalisation |
| Indicator for hospital admission | Phototherapy treatment (yes/no) |
| Characteristics of phototherapy (age of neonate at start of treatment, duration, number of lamps/blankets and intensity) | Exchange transfusion (yes/no) |
| Complications due to blood transfusion treatment (thrombocytopenia, hypercalcaemia, other electrolyte imbalance and necrotising enterocolitis) | Number of (heel) pricks to determine bilirubin levels |
| Maximum level of bilirubin measured (if known) | Result MRI cerebrum (if applicable) |
| Result AABR (if performed) | More than 10% loss of birth weight within the first week of life (yes/no) |
| Proportion of neonates diagnosed with acute bilirubin encephalopathy | Proportion of neonates diagnosed with kernicterus spectrum disorder |

AABR, automated auditory brainstem response.

**Study size**

For the sample size calculation, we estimated the proportion of participants at each assessment time point in four groups based on the combination of visual inspection and TcB quantification. The four groups are as follows: (A) no significant jaundice on visual inspection, and TcB below threshold; (B) significant jaundice on visual inspection, but TcB below threshold; (C) no significant jaundice on visual inspection, but TcB equal to or higher than threshold and (D) significant jaundice on visual inspection.
inspection, and TcB equal to or higher than threshold (figure 1). For this purpose, ‘significant jaundice’ is any degree of jaundice considered to be sufficiently severe to indicate the need for LBB quantification as per the midwife’s assessment. The probability matrix of neonates belonging to each of the four groups at various time points across the study period is based on published literature and preliminary data from the STARSHIP trial.\textsuperscript{15} We estimated the proportions of neonates falling in each of the four groups across the observations as follows: A: 83%; B: 10.5%; C: 2% and D: 4.5%.\textsuperscript{15} These proportions take into account the fact that multiple assessments may be performed for individual neonates, and as such, a neonate may be categorised in a different group at different time points.

To address the primary objective of assessing whether universal TcB screening can increase the timely detection of neonates with hyperbilirubinaemia necessitating treatment compared with using only visual inspection, and at the same time decrease the number of heel pricks performed to quantify total bilirubin in blood, we will test the following hypotheses:

(1) whether the absolute number of neonates with hyperbilirubinaemia necessitating treatment is higher in group C than group B, that is, more neonates with hyperbilirubinaemia necessitating treatment are detected with TcB screening than with visual inspection; and

(2) whether the absolute number of neonates without hyperbilirubinaemia necessitating treatment is higher in group B than group C, that is, fewer neonates without hyperbilirubinaemia necessitating treatment get a heel prick if TcB screening would replace visual inspection as the standard screening approach.

To test these hypotheses, we will use two McNemar tests, conditional on hyperbilirubinaemia necessitating treatment. Assuming a distribution of participants over the groups as indicated above, we expect 20% of the neonates who test negative on visual inspection and positive on the TcB test (ie, those in group C) to require phototherapy. Similarly, we expect 0% of the neonates who test positive on visual inspection and negative on the TcB test (ie, those in group B) to require phototherapy; that is: TcB screening can effectively rule out hyperbilirubinaemia above the phototherapy threshold among participants with visual jaundice. Using a p-value of 0.05, the sample size calculation for the exact McNemar test indicates that, for the first hypothesis, 38 neonates with hyperbilirubinaemia requiring treatment are required to obtain a power of 80%.\textsuperscript{22} With an anticipated event rate of hyperbilirubinaemia requiring treatment across all time points of 1.8% (20% of group C and 30% of group D combined), this would require 2100 neonates in total. The power is not affected by the repeated measures per subject since, by definition, there will be at most one measurement per neonate indicating the requirement for treatment: at this point, the neonate will be admitted to hospital and started on phototherapy and therefore will no longer contribute data points to the study.

For the second hypothesis, each neonate can have multiple outcome events (ie, require a heel prick at more than one time point). Assuming the distributions over the four groups and event rates indicated as above, group C is expected to contain 42 neonates (ie, 2% of 2100) of which we assume that 80% (ie, n=34) do not require phototherapy (figure 4). We would expect that 1.6% (34/2100×100=1.6%) of neonates not requiring phototherapy would have a negative visual inspection and a positive TcB test (group C). Similarly, group B is expected to contain 220 neonates (ie, 10.5% of 2100; figure 4). We would expect that 10.5% (220/2100×100=10.5%) of neonates not requiring phototherapy have a positive visual inspection and a negative TcB test (group B).

An analysis without repeated measures would require 125 neonates to obtain 80% power at a significance level
of 0.05. Taking into account the expected median of three repeated measures per neonate and an intraclass correlation of 0.05, we obtain a variance inflation factor (VIF) of

\[
\text{VIF} = 1 + (m - 1)\rho = 1 + (3 - 1) \times 0.05 = 2
\]

leading to a total of 250 neonates who do not necessitate treatment. With an event rate of neonates requiring phototherapy of 1.8%, this would lead to sample size of 260 neonates to address the second hypothesis.

We expect a loss to follow-up of maximum 10%. To account for this expected loss to follow-up, we aim to include 2310 neonates for the first hypothesis.

**Statistical analysis**

The primary analysis will involve two separate McNemar tests, conditional on LBB being above the phototherapy threshold or not. These tests will provide \( p \)-values for testing the following two \( H_0 \) hypotheses: (1) there is no difference in the detection of neonates with hyperbilirubinaemia necessitating treatment between TcB screening and visual inspection, and (2) there is no difference in the number of heel pricks needed to quantify total bilirubin in blood if TcB would replace visual inspection as standard care.

For the first hypothesis, we test whether, conditional on having hyperbilirubinaemia, the probability of a ‘positive’ TcB test (ie, TcB reading above the threshold) is different than that of a ‘positive’ visual inspection (ie, visual inspection indicating the need for LBB quantification at the discretion of the community midwife). As both tests are performed in the same neonate and hence are paired, we only need to test whether the probabilities of having discordant test results are different: \( (\text{TcB}^+, \text{Visual}^-) \) vs. \( \text{TcB}^-, \text{Visual}^+ \). In which \( + \) and \( - \) are displayed as positive and negative tests, respectively. Similarly, for the second hypothesis, we will perform an adjusted McNemar test, taking into account the repeated measurements in neonates, to test if the probability of having a positive TcB test is different than the probability of having a positive visual inspection, conditional on not having hyperbilirubinaemia: \( \text{P (TcB}^+, \text{Visual}^-) \) vs. \( \text{P (TcB}^-, \text{Visual}^+) \).

Sensitivity of the TcB and visual inspection is estimated by the proportion of positive tests divided by the number of babies requiring phototherapy. Agresti-Coull 95% CIs will be estimated for the sensitivities. If, as expected, the sensitivity of TcB is indeed 100%, we will use the ‘rule of three’ to derive the 95% CI for the first hypothesis. To quantify the specificity of TcB and visual inspection at the population level, we will use generalised estimating equations for both tests separately.

Furthermore, we will use 2×2 tables and Bland-Altman plots to assess the diagnostic accuracy of each novel screening method to obtain a first impression at each time point, without taking correlation between repeated measurements into account. The primary diagnostic comparison will involve TcB versus visual inspection. Other relevant diagnostic comparisons are displayed in table 1.

We will perform subgroup analyses to determine effectiveness of TcB screening and the diagnostic accuracy of the tools across different ethnicities and gestational age strata. Ethnicities will be defined according to the guideline of the Royal Dutch Organisation of Midwives.

**Cost-effectiveness analysis**

Following established methodologies, a model-based CEA will be performed to analyse the cost-effectiveness of the three screening and diagnostic tools compared with standard care, as described above. A decision tree model will be built, because screening for neonatal hyperbilirubinaemia can be represented by a relatively simple sequence of decisions. The decision tree will be constructed such that a hypothetical neonate would either receive screening by one of the novel tools or standard care. The model will consider the accuracy of the different screening tools. The analysis will be performed from a healthcare perspective, using a time horizon of the 8–14 day observation period of the study. The probabilities used in the model will be based on the current study.

All relevant healthcare use will be considered in the model, including screening for hyperbilirubinaemia by TcB and by the Picterus app, total bilirubin quantification using Bilistick, LBB testing, frequency of community midwife visits (including telephone calls), visits of a specialised laboratory home service that performs the heel prick (if relevant), transport costs for LBB samples to the local laboratory, consultations with a paediatrician, hospital admission days and treatments for hyperbilirubinaemia (phototherapy, exchange transfusion). To arrive at costs, resources used will be multiplied by integral cost prices. The effectiveness measure of the CEA will be the number of correctly diagnosed cases of hyperbilirubinaemia necessitating treatment and the number of heel pricks avoided.

For each of the three tools, the mean expected costs and effects will be calculated and compared with standard care. To this aim, incremental cost-effectiveness ratios will be calculated expressed as incremental costs per additional correctly diagnosed case of hyperbilirubinaemia necessitating treatment (regarding TcB and Picterus app) or incremental costs per heel prick avoided (regarding TcB and, Picterus app). Sensitivity analysis will be carried out to assess the robustness of the results to realistic variations in the levels of the underlying parameter values and key assumptions.

**Safety**

All adverse events reported spontaneously by the parents of the participant or observed by the study personnel, starting from the first application of the novel approaches until 24 hours after the last application of the novel approaches, will be recorded. Serious adverse events (SAEs) will be reported by the community midwives.
and communicated to the principal investigator within 1 week after identification of the event. The principal investigator, in collaboration with the study team, will record these events and report them to the sponsor and Medical Research Ethics Committee of the Erasmus MC Rotterdam. The SAEs are defined apriori and communicated with the study team. SAEs consisted of bilirubin levels above exchange transfusion threshold, exchange transfusion, disabilities related to severe hyperbilirubinaemia, hospital admission or death.

Monitoring will be conducted according to the Guideline for Good Clinical Practice by an independent, professional monitor once a year per community midwifery practice.31

Patient and public involvement

The study protocol and procedures were discussed with the patient panels of the Regional Consortia Pregnancy and Childbirth South-West and North Netherlands. Parents and student midwives from the Rotterdam University of Applied Sciences were involved to give feedback on the study design and recruitment into the study. Moreover, the Regional Consortia Pregnancy and Childbirth South-West and North Netherlands, multidisciplinary organisations that provide a platform for healthcare professionals to share and obtain knowledge regarding childbirth at the regional level, were actively involved in developing the study protocol. Prior to start of the inclusion, participating community midwives were invited to give feedback and discuss the study protocol and procedures thoroughly with the study team. Their input is incorporated to optimise study procedures.

Ethics and dissemination

The study has received ethical approval from the Medical Research Ethical Committee of the Erasmus MC Rotterdam, Netherlands (MEC-2020-0618), and was registered at the Netherlands Trial Register (NL9545; now available via the International Clinical Trials Registry Platform). Parents will provide written informed consent. Any protocol modifications will be reported to the medical research ethics committee and, if relevant, to the participating community midwifery practices.

Data will be pseudonymised. The identification list will only be accessible by the investigators and the study monitor. The data provided to Pictorus will only be used for providing bilirubin level estimates. Data of participants will not be used for any other purpose (eg, commercial purposes).

The results of this study will be published in peer-reviewed journals and presented at national and international scientific meetings. The results will be reported using the guideline for diagnostic studies of the EQUATOR Network.32 The statistical code will be published alongside the manuscript. We intend to provide an anonymised version of the study data on reasonable request after publication. The International Committee of Medical Journal Editors guideline will be used to determine authorship.33

Author affiliations

1Division of Neonatology, Department of Paediatrics, Erasmus MC Sophia Children Hospital, University Medical Centre Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
2Department of Obstetrics and Gynaecology, Erasmus MC Sophia Children’s Hospital, University Medical Centre Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
3Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
4Department of Epidemiology, University Medical Centre Groningen, Groningen, The Netherlands
5Department of Clinical Chemistry, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands
6Department of Pediatric Intensive Care, Division of Paediatric Surgery, Erasmus MC Sophia Children Hospital, University Medical Centre Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
7Nursing Science, Department of Internal Medicine, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
8Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands
9Division of Neonatology, Department of Paediatrics, Beatrix Children’s Hospital, University Medical Centre Groningen, Groningen, The Netherlands

Twitter Henk Groen @Groen62H and Jasper V Been @JasperBeen

Collaborators On behalf of the BEAT study group: Hanneke Bontenbal-Plezier (Midwifery practice Chailoos, Rotterdam, the Netherlands), Helene Bouma (Division of Neonatology, Department of Paediatrics, Beatrix Children's Hospital, University Medical Centre Groningen, Groningen, the Netherlands), Jan Jaap Erwich (Department of Obstetrics and Gynaecology, University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands), Esther I. Feijen-de Jong (Department of Midwifery Science, Amsterdam Public Health Research Institute, VU University Medical Centre, Amsterdam, the Netherlands and Department of General Practice & Eldery Medicine, University of Groningen, University Medical Centre Groningen, Groningen, the Netherlands), Dorien Kramia (Midwifery practice Veendam, Veendam, the Netherlands), Jenneke Kruidhof (Midwifery practice Fiere, Groningen, the Netherlands), Celine Leonard (Midwifery practice Langersland, Bergschenhoek, the Netherlands), Gea Reussing (Midwifery practice NOP/Lemstersland, Emmeloord, the Netherlands), Joyce Vaandragter (Midwifery practice Luselmonde, Rotterdam, the Netherlands), Ellen Vreugenhil (Midwifery practice Midden Groningen, Hoogezand, the Netherlands)

Contributors LEHW and JVB wrote the first version of the manuscript. JVB, LEHW, BAMvdG, CW, JV, EI, PHD, HG, HL, IKMR and EAPS designed the study. JVB, BAMvdG, CW, JV, HD, HJ, IKMR and EAPS secured funding. HL, DN and HG developed the methodological part of this study. MP developed the cost-effectiveness analyses. All authors critically revised the manuscript and gave final approval of the version to be published.

Funding This work was supported by The Netherlands Organisation for Health Research and Development (ZonMW), grant number 943002011.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Study is currently in the inclusion phase, no data are yet available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.
REFERENCES

1. Das S, van Laneghem FKH. Clinico-pathological spectrum of bilirubin Encephalopathy/Kernicterus. *Diagonstics* 2019;9:24.
2. Le Pichon J-B, Riordan SM, Watchko J, et al. The neurological sequelae of neonatal hyperbilirubinemia: definitions, diagnosis and treatment of the kernicterus spectrum disorders (KSDs). *Curr Pediatr Rev* 2017;13:199–209.
3. Bousema S, Govaert P, Dudink J. Kernicterus is preventable but still occurs] Kernicterus is verminjdbaar, maar komt nog steeds voor: [Ned Tijdschr Geneeskd] 2015;159:A8518.
4. Olsanaya BO, Kaplan M, Hansen TWR. Neonatal hyperbilirubinemia: a global perspective. *Lancet Child Adolesc Health* 2018;2:610–20.
5. The Netherlands Perinatal Registry. Yearbooks of perinatal care in the Netherlands. Available: www.perined.nl/onderwerpen/publicaties-perined/jaarboek-zorg
6. de Boer J, Zondag L. [Multidisciplinary guideline postnatal care] Multidisciplinaire richtlijn Postnatale Zorg - Verloskundige basisszorg voor moeder en kind: Koninklijke Nederlandse Organisatie van Verloskundigen, 2018. Available: https://www.knov.nl/serve/file/knov.nl/knov_downloads/2882/file/Postnatale_zorg_opgemaakt_versie_door_IB_md_10_aug_2018.pdf
7. Nederlandse Vereniging voor Kindergeneeskunde. Kwaliteitsinstituut voor de Gezondheidszorg CBO. [Guideline prevention, diagnosis and treatment of hyperbilirubinemia among newborns born at a gestational age of more than 35 weeks] Richtlijn preventie, diagnostiek en behandeling van hyperbilirubinemie bij de pasgeborene, geboren na een zwangerschapsduur van meer dan 35 weken, 2008. Available: https://www.nvk.nl/Portals/0/richtlijnen/hyperbilirubinemie.pdf
8. Maisels MJ, Clune S, Coleman K, et al. The natural history of jaundice in predominantly breastfed infants. *Pediatrics* 2014;134:e340–5.
9. Riskin A, Tamir A, Kugelman A, et al. Is visual assessment of jaundice reliable as a screening tool to detect significant neonatal hyperbilirubinemia? *J Pediatr* 2008;152:782–7.
10. Keren R, Tremont K, Luan X, et al. Visual assessment of jaundice in term and late preterm infants. *Arch Dis Child Fetal Neonatal Ed* 2009;94:F317–22.
11. Gotink MJ, Benders MJ, Lavrijsen SW, et al. Severe neonatal hyperbilirubinemia in the Netherlands. *Neonatology* 2013;104:137–42.
12. Eggert LD, Wiedmeier SE, Wilson J, et al. The effect of instituting a prehospital-discharge newborn bilirubin screening program in an 18-hospital health system. *Peditrics* 2006;117:e855–62.
13. Wainer S, Parmar SM, Allegro D, et al. Impact of a transcutaneous bilirubinometry program on resource utilization and severe hyperbilirubinemia. *Pediatrics* 2012;129:77–86.
14. van den Ever-Jonker B, den Boer L, Pepping RMC, et al. Transcutaneous Bilirubinometry in jaundiced neonates: a randomized controlled trial. *Pediatrics* 2016;138:e20162414.
15. van der Geest BAM, de Graaf JR, Bertens LCM, et al. Screening and treatment to reduce severe hyperbilirubinemia in infants in primary care (STARSHIP): a factorial stepped-wedge cluster randomised controlled trial protocol. *BMJ Open* 2019;9:e028270.
16. Coda Zabetta CD, Iskander IF, Greco C, et al. Bilistick: a low-cost point-of-care system to measure total plasma bilirubin. *Neonatology* 2010;12:176–8.
17. Aune A, Vartdal G, Bergseng H, et al. Bilirubin estimates from smartphone images of newborn infants’ skin correlated highly to serum bilirubin levels. *Acta Paediatr* 2020;109:2532–8.
18. Hulzebos CV, Vitek L, Coda Zabetta CD, et al. Screening methods for neonatal hyperbilirubinemia: benefits, limitations, requirements, and novel developments. *Pediatr Res* 2021;90:272–6.
19. Bhardwaj K, Locke T, Biringer A, et al. Newborn bilirubin screening for preventing severe hyperbilirubinemia and bilirubin encephalopathy: a rapid review. *Curr Pediatr Rev* 2017;13:67–90.
20. Nagar G, Vandemeyer B, Vanambeek L, et al. Reliability of transcutaneous bilirubin devices in preterm infants: a systematic review. *Pediatrics* 2013;132:871–81.
21. Greco C, Iskander IF, Hoouchi SZ, et al. Diagnostic performance analysis of the point-of-care Bilistick system in identifying severe neonatal hyperbilirubinemia by a Multi-Country approach. *ECLinicalMedicine* 2018;1:14–20.
22. Duffy SW. Asymptotic and exact power for the McNemar test and its analogue with R controls per case. *Biometrics* 1984;40:1005–15.
23. Wu Y. A robust adjustment to McNemar test when the data are clustered. *Commun Stat Theory Methods* 2021;50:1515–29.
24. Agresti A, Coull BA. Approximate Is Better than “Exact” for Interval Estimation of Binomial Proportions. *The American Statistician* 1998;52:119–26.
25. Ly G, Lü, Friston KJ. A practical introduction to Gaussian processes. *Int J Mach Learn Comput* 2012;2:56–58.
26. van den Esker W, van der Geest BAM, de Graaf JP. Transcutaneous Bilirubinometry in jaundiced neonates: a randomized controlled trial. *Pediatrics* 2016;138:e20162414.