UK National Screening Committee

# SCREENING FOR BILIARY ATRESIA IN NEWBORNS

## External review against programme appraisal criteria for the UK National Screening Committee

Version: 3 (consultation)

Author: Solutions for Public Health

Date: February 2026

**The UK National Screening Committee secretariat is hosted by the Department of Health and Social Care**

# Contents

# About the UK National Screening Committee

The UK National Screening Committee (UK NSC) advises ministers and the NHS in the 4 UK countries about all aspects of population screening and targeted screening and supports implementation of screening programmes.

Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK National Screening Committee, Southside, 39 Victoria Street, London, SW1H 0EU

www.gov.uk/uknsc

Blog: https://nationalscreening.blog.gov.uk/

For queries relating to this document, please contact: uknsc@dhsc.gov.uk.



Published February 2026

# List of abbreviations

| | |
|---|---|
| CI | Confidence intervals |
| ELISA | Enzyme-linked immunosorbent assay |
| GP | General practitioner |
| GRADE | Grading of recommendations assessment, development and evaluation |
| IQR | Interquartile range |
| JBI | Joanna Briggs Institute |
| MMP-7 | Matrix metalloproteinase-7 |
| N/A | Not applicable |
| Ng/mL | Nanogram per millilitre |
| NICE | National Institute for Health and Care Excellence |
| NPV | Negative predictive value |
| PPV | Positive predictive value |
| QUADAS-2 | Quality Assessment of Diagnostic Accuracy Studies-2 |
| RCT | Randomised controlled trials |
| SD | Standard deviation |
| SPH | Solutions for Public Health |
| UK NSC | United Kingdom National Screening Committee |
| USA | United States of America |
| µmol/L | Micromole per litre |
| WHO | World Health Organisation |

# List of tables

# List of figures

# Plain English summary

Biliary atresia is a rare disease where the bile ducts become inflamed and blocked. Symptoms develop within a few weeks after birth. Without treatment, biliary atresia can cause death by the age of 2 years old. The only treatment for biliary atresia is a type of surgery called the Kasai procedure. This surgery is more likely to be successful if done before 45 to 60 days of age.

Newborn screening for biliary atresia is not currently recommended in the UK.

The use of newborn dried blood spots is one way to screen for biliary atresia. There are national screening programmes using stool colour cards in Japan and Taiwan.

This evidence summary considers new evidence published since 2012. It aimed to answer 3 questions to explore:

- the accuracy of screening tests to detect biliary atresia in newborns
- the age at surgery for biliary atresia in the UK
- whether screening using stool colour cards leads to better outcomes for infants

The conclusions of this evidence summary were that:

- there is not enough evidence to recommend the use of dried blood spots to screen for biliary atresia
- the accuracy of stool colour cards to identify biliary atresia is unclear because the available evidence is inconsistent
- screening using stool colour cards has improved outcomes for infants with biliary atresia in some countries
- recent evidence shows that babies in the UK usually have surgery at around 50 days old. This is about the same age as in countries that use stool colour cards to help spot the condition earlier

There is still not enough evidence to recommend newborn screening for biliary atresia.

# Executive summary

## Purpose of the review

This document reviews the evidence on screening for biliary atresia in newborns against UK NSC criteria to support decision making about whether a screening programme would be of benefit in the UK. Evidence relating to 4 UK NSC criteria was assessed including:

- the accuracy of screening tests
- the current clinical management of the condition
- the effectiveness of screening programmes
- the benefits and harms of screening programmes

## Background

Biliary atresia is a rare neonatal liver disease of unknown aetiology where inflamed and blocked bile ducts lead to a build-up of bile in the liver. Without treatment, the disease progresses rapidly and can cause death from liver disease within the first 2 years of life. There are approximately 37 cases of biliary atresia in the UK each year.

The symptoms of biliary atresia include persistent jaundice, pale stools and dark urine which develop within the first few weeks after birth. Currently, early detection of biliary atresia relies upon prompt recognition of prolonged jaundice lasting for more than 14 days in term infants, or more than 21 days in pre-term infants. However, jaundice has multiple causes and is common in early infancy. Identifying the infants who have jaundice due to biliary atresia is key for early diagnosis and treatment.

The only treatment for biliary atresia is a surgical procedure, the Kasai procedure or Kasai portoenterostomy, to re-establish bile flow from the liver into the intestine. Surgery can prevent or delay the need for liver transplant. Best practice guidelines suggest that surgery should ideally be performed by 45 to 60 days of life and that patients who have surgery after 100 days have significantly worse outcomes.

Screening for biliary atresia using newborn dried blood spots was identified in previous UK NSC reviews as an approach that could potentially detect disease very early and would be applicable to current UK screening practice. The 2021 UK NSC evidence map also considered the evidence around home-based screening using stool colour cards.

National screening programmes using stool colour cards have been implemented in Taiwan and Japan since 2004 and 2012 respectively. Regional or pilot screening programmes have also been implemented in several countries including China, Canada, Argentina and Germany. There are no existing screening programmes using dried blood spots.

## Focus of the review

This evidence summary about screening for biliary atresia in newborns addresses 3 key questions:

1. What is the accuracy of screening tests to detect biliary atresia in newborns? (UK NSC criterion 4)

2. What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK? (UK NSC criterion 15)
3. Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes? (UK NSC criteria 11 and 13)

The search for evidence for questions 1 and 3 included relevant studies published since January 2012. The search for evidence for question 2 included relevant studies published since November 2016. Searches were conducted in August 2025.

## Recommendation under review

The UK NSC does not currently recommend newborn screening for biliary atresia.

The last UK NSC evidence summary on newborn screening for biliary atresia was published in 2017 as an update of a previous evidence summary in 2012. The 2017 evidence summary did not identify any studies on testing for biliary atresia using dried blood spots in a general newborn population. Furthermore, no studies were found to update the median age at surgery in the UK reported in the 2012 evidence summary. The 2017 UK NSC evidence summary concluded that the age at surgery reported from countries with stool colour card screening programmes was comparable with the age at surgery in the UK reported in the 2012 evidence summary.

In 2021, a UK NSC evidence map explored the volume and type of evidence available on the diagnostic accuracy of screening tests using dried blood spots or stool colour cards, whether screening for biliary atresia using stool colour cards improves time to surgery and clinical outcomes and the reported age at surgery for biliary atresia in the UK. The evidence map concluded that there was sufficient new evidence available to justify commissioning an evidence summary.

## Findings

Seventeen publications were included in this evidence summary. The evidence for each question is summarised below:

Criterion 4 — '*There should be a simple, safe, precise and validated screening test'*

Question 1 — *What is the accuracy of screening tests to detect biliary atresia in newborns?*

The volume of evidence about the accuracy of screening tests using dried blood spots was limited to one meta-analysis, one retrospective study and 2 case control studies, all testing for different targets within dried blood spots. The details of test performance results reported in these studies was limited. Sensitivities and specificities of over 90% were reported, but confidence intervals were wide, increasing uncertainty in the results. Overall, the quantity and quality of the evidence available for dried blood spot tests is insufficient to draw any conclusions.

There was a greater volume and quality of evidence about the accuracy of screening tests using stool colour cards with data collected from consecutively enrolled populations as part of screening programmes in several non-UK countries. There was a lack of consistency in the sensitivity results reported across different studies and confidence intervals, where reported, were wide reflecting the low number of biliary atresia cases identified. There was also variation in the positive predictive values (PPV) reported reflecting the differing incidence of biliary atresia in different countries. Two studies reported PPV for a prevalence that is more applicable to that

found in the UK and this was between 4% and 6%. The specificity and negative predictive value results were more consistent across studies and were generally high.

Overall, limitations in the quantity, quality, consistency and applicability of the evidence available for the accuracy of screening tests to detect biliary atresia in newborns means that this criterion is currently not met.

Criterion 15 — '*Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme*'

Question 2 — *What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK?*

Previous UK NSC evidence summaries have concluded that the median age at surgery of 54 days (range 7 to 209) in the UK, based on data up to 2009, was comparable with the reported age at surgery in countries with stool colour card screening programmes.

The current evidence summary reported a median age at surgery in the UK of 51 days (interquartile range (IQR) 39 to 64) based on data up to 2019. There was also evidence that this has improved over time to 48 days (IQR 35 to 57) for the most recent time period available (2014 to 2018). No data were identified to update the proportion of infants that have a late (>60 days and >90 days) Kasai portoenterostomy in the UK.

This criterion has been considered met in previous UK NSC evidence summarises. Based on the more recent data identified for the current evidence summary, age at surgery in the UK is still comparable to the reported age at surgery in countries with stool colour card screening programmes. Therefore, this criterion is still met.

Criteria 11 and 13 — '*There should be evidence from high quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity*' and '*The benefit gained by individuals from the screening programme should outweigh any harms, for example from overdiagnosis, overtreatment, false positives, false reassurance, uncertain findings and complications'*

Question 3 — *Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes?*

The studies identified consistently reported that time to surgery improved for infants with biliary atresia after the introduction of screening using stool colour cards. There also appeared to be a reduction in the proportion of infants receiving late surgery and improvement in other clinical outcomes, although the statistical and clinical significance of these outcomes was less clear. There was no indication within the included studies that screening using stool colour cards is likely to lead to harm. However, none of the studies identified used a randomised controlled trial design. Therefore criterion 11 cannot be considered as met.

Criterion 13, about the benefits of screening outweighing the harms, could be considered as met within the context of the countries in which screening with stool colour cards has been introduced. For example, in countries with screening, mean age at surgery after introducing stool colour cards improved and ranged from 46 to 60 days. However, the applicability of the results of these studies to the UK context is uncertain as the age at surgery following the introduction of the screening programmes is similar to that reported in the UK without a screening programme being in place.

Overall, whilst improvements have been observed following the introduction of stool colour card screening in some countries, limitations in the study design, uncertainties about the clinical

relevance of the improvements observed and the applicability to the UK context means that these criteria cannot be considered as met.

## Recommendations on screening

The volume, quality and direction of the new evidence published since the last evidence summary suggests that the current recommendation not to introduce a UK systematic population screening programme for biliary atresia in newborns should be retained.

## Limitations

This evidence summary was conducted according to the UK NSC evidence review process over a condensed period of time. The review only looked for peer-reviewed scientific work and does not include work published elsewhere (grey literature). Studies not available in the English language, abstracts and poster presentations were not eligible for inclusion. Given that these are accepted methodological adjustments for a rapid review, and that the searches included relevant systematic reviews published since this topic was last considered by the UK NSC, these limitations should not have led to the exclusion of any pivotal studies.

## Evidence uncertainties and gaps to be addressed

As in previous evidence summaries, there is a lack of evidence considering the diagnostic accuracy of dried blood spot tests in consecutively enrolled or randomly assigned populations. The studies identified about the use of dried blood spots to screen for biliary atresia have considered a range of different targets to test for suggesting ongoing uncertainty about the potential of using dried blood spots to screen for biliary atresia. The potential of using serum blood samples to screen newborns for biliary atresia has been explored in the USA.

Data were available to update the median age at surgery in the UK reported in previous evidence summaries. However, data were only available up to 2018 and no data were available to update the proportion of infants that have a late Kasai portoenterostomy in the UK. Further peer reviewed publications reporting the latest data on age at surgery in the UK would be helpful, particularly as the years since 2018 include the years in which the COVID-19 pandemic impacted the delivery of many services.

# Introduction and approach

The UK NSC does not currently recommend screening for biliary atresia. The Committee based this recommendation on the evidence provided by the 2017 UK NSC review [1]. In 2020, the UK NSC commissioned an evidence map to examine if there was sufficient new evidence to justify the commissioning of an evidence summary to address the evidence gaps highlighted by the 2017 UK evidence summary. This evidence map was published in 2021 and concluded that there was sufficient evidence to justify commissioning an evidence summary [2].

This document reviews the evidence on screening for biliary atresia in newborns against the UK NSC criteria about test accuracy, the current clinical management of the condition, the effectiveness of screening programmes, and the benefits and harms of screening programmes.

## Background

Biliary atresia is a rare neonatal liver disease of unknown aetiology where the bile ducts become inflamed and blocked preventing the flow of bile from the gall bladder to the intestine. This leads to a build-up of bile in the liver. Untreated, the disease rapidly progresses to cholestasis, growth failure, cirrhosis and death from end-stage liver disease within the first 2 years of life [3]. In England and Wales, the estimated incidence of biliary atresia is 0.58 per 10,000 live births [4, 5]. There were 594,677 live births in England and Wales in 2024 which would equate to approximately 34 annual cases of biliary atresia [6]. The reported incidence of biliary atresia in Scotland is 1 in 14,745 live births (0.68 per 10,000) [7] which would equate to approximately 3 cases of biliary atresia annually [8].

Biliary atresia is classified according to the site of blockage:

- type 1: atresia restricted to the common bile duct (about 5% of cases)
- type 2: atresia of the common hepatic duct (about 2% of cases)
- type 3: atresia of the section of biliary tract closest to the liver, with atresia of both right and left hepatic ducts (the most common type, accounting for over 90% of cases) [5]

The symptoms of biliary atresia include persistent jaundice, pale stools and dark urine developing within the first few weeks after birth. Currently, early detection of the condition relies upon prompt recognition of prolonged jaundice, which the National Institute for Health and Care Excellence (NICE) define as jaundice lasting for more than 14 days in term infants, or more than 21 days in pre-term infants [9].

The NICE clinical guideline on jaundice in newborn babies under 28 days [9] recommends that babies should be examined for jaundice at every opportunity, especially within the first 72 hours after birth. A check for jaundice is part of the NHS newborn physical examination which is carried out within 72 hours of birth*. Parents/carers are also advised to check for symptoms of jaundice after this as symptoms can sometimes take longer to appear and to speak to their midwife, health visitor or GP as soon as possible if they detect any signs of jaundice.

When jaundice is suspected or obvious within the first 24 hours of life, the NICE recommendation is to measure and record the serum bilirubin level urgently (within 2 hours) with referral for an urgent medical review, ideally within 6 hours, to exclude pathological causes

* Newborn jaundice - Diagnosis - NHS

of jaundice. For babies more than 24 hour old, the recommendation is to measure and record the bilirubin level urgently (within 6 hours) [9].

For babies who are in the first 24 hours of life or who have a gestational age of less than 35 weeks, serum bilirubin measurement is recommended. For babies who are more than 24 hours old and have a gestational age of 35 weeks or more, a transcutaneous bilirubinometer† can be used to measure bilirubin level with a blood sample taken to measure serum bilirubin if a transcutaneous bilirubinometer is not available or if the transcutaneous bilirubinometer measurement indicates a bilirubin level of more than 250 µmol/L [9].

Jaundice is common in early infancy, affecting approximately 60% of term infants and 80% of pre-term infants during the first week of life, with about 10% of breastfed infants still being jaundiced at 28 days [10]. Most of these infants will have non-cholestatic jaundice with raised levels of unconjugated bilirubin. Whereas jaundice caused by cholestasis results in raised levels of conjugated (also referred to as direct) bilirubin [11]. Distinguishing jaundice caused by cholestasis from non-cholestatic conditions is critical in the early diagnosis of biliary atresia. Biliary atresia is the most common cause of cholestatic jaundice in the first months of live accounting for around 25% to 40% of cases. Other causes of cholestatic jaundice include diseases caused by single gene disorders and combinations of multiple factors, for example relating to parental nutrition. In some cases, the cause is never determined. Identifying the infants who have cholestatic jaundice due to biliary atresia is key for early diagnosis and treatment [11].

The only treatment for biliary atresia is a surgical procedure, a hepatoportoenterostomy (the Kasai procedure or Kasai portoenterostomy), to re-establish bile flow from the liver into the intestine. The damaged ducts outside the liver are removed and a loop of intestine is attached to the remaining healthy bile ducts to reinstate bile flow to the intestine [12]. A Kasai portoenterostomy can prevent or delay the need for liver transplant.
The younger the infant at the time of surgery the more likely the surgery will be successful with longer liver survival [13]. Previous UK NSC reviews indicated that the procedure should be performed before 90 days. The British Medical Journal best practice guidelines suggest that surgery should ideally be performed by 45 to 60 days of life and that patients who have surgery after 100 days have significantly worse outcomes [14].

Screening for biliary atresia using newborn dried blood spots was identified in previous UK NSC reviews as an approach that could potentially detect disease very early and would be applicable to current UK screening practice. The 2021 UK NSC evidence map [2] also considered the evidence around home-based screening using stool colour cards. Stool colour cards include photographs of infant stool (usually 6 to 9 images) depicting shades of abnormal stool colour and normal stool colour with information on the significance of the stool colour and advice to parents to monitor their infant's stool colour for the first month of life and contact health services if the colour is abnormal.

## Newborn screening for biliary atresia

Screening using stool colour cards has been implemented in some countries nationally or regionally. The first screening programme using stool colour cards was introduced in Tochigi Prefecture in Japan in 1994 and was gradually expanded before being rolled out nationally in 2012 [15, 16]. In 2004, a national screening programme was introduced in Taiwan after a successful pilot scheme in 2002 [17, 18]. These countries have a higher incidence of biliary

† A transcutaneous bilirubinometer shines light onto the baby's skin to calculate the level of bilirubin by analysing how light reflects off or is absorbed by the skin

atresia than the 0.58 per 10,000 live births reported in the UK. For Japan this is 1.1 per 10,000 live births [15]. For Taiwan this ranged between 1.2 and 1.9 per 10,000 live births in the years between 1997 and 2010 with no evidence of a specific trend over time [18]. In addition, regional screening programmes or pilot programmes using stool colour cards have been introduced in China [19], Canada [20, 21, 22], Argentina [23] and Germany [24].

Studies reporting outcomes from the stool colour card screening programmes in Japan, Taiwan, China and British Columbia, Canada are included in this evidence summary against the key questions. Studies regarding experiences of screening programmes in Argentina, Canada, and Germany were also identified by the searches for this evidence summary. These papers are not formally included in the evidence summary because they did not report either the diagnostic accuracy of stool colour cards or compare outcomes from screening using stool colour cards to no screening.

There are no existing screening programmes using dried blood spots. However, the searches for this evidence summary also identified studies from some states in the USA reporting the diagnostic accuracy of a 2-stage newborn screening strategy for biliary atresia based on the measurement of direct or conjugated bilirubin in serum blood samples [25, 26, 27].

Studies about the screening programmes in Argentina, Germany, Canada and the USA are briefly discussed in the evidence summary below.

## Current policy context and previous reviews

There is no newborn screening programme for biliary atresia in the UK. Currently, the early detection of biliary atresia relies on the prompt recognition of prolonged jaundice.

The UK NSC does not currently recommend screening for biliary atresia. The last UK NSC evidence summary on newborn screening for biliary atresia was published in 2017 [1]. This evidence summary, an update on the 2012 UK NSC evidence summary, looked for evidence published between 2012 and 2016 on screening for biliary atresia using dried blood spots and the reported mean age at surgery for biliary atresia (Kasai portoenterostomy) in the UK. The 2017 UK NSC evidence summary did not identify any studies that considered any form of testing for biliary atresia using dried blood spots in a general newborn population. Furthermore, no studies were found that provided the mean or median age at surgery for biliary atresia in the UK. It was therefore not possible to determine whether the age at surgery had changed from the median of 54 days reported in the previous 2012 UK NSC evidence summary. The 2017 review identified one study in Taiwan [28] and one study in Japan [15] that reported reductions in age at surgery after the introduction of screening programmes using stool colour cards to 46.0 ± 23.8 days and 58.5 days (range 18 to 109) respectively. The 2017 UK NSC evidence summary concluded that this age at surgery was comparable with the median age of 54 days (range 7 to 209) at surgery in the UK as reported in the 2012 UK NSC evidence summary.

In 2021, a UK NSC evidence map [2] explored the volume and type of evidence available on the diagnostic accuracy of screening tests using dried blood spots or stool colour cards, whether screening for biliary atresia using stool colour cards improves time to surgery and clinical outcomes and the reported age at surgery for biliary atresia in the UK. The evidence map concluded that there was sufficient new evidence available to justify commissioning an evidence summary.

# Objectives

The current evidence summary aims to look at the accuracy of screening tests to detect biliary atresia in newborns, the reported age at surgery/time to surgery for biliary atresia in the UK and whether screening for biliary atresia using stool colour cards improves time to surgery and clinical outcomes.

Table 1 provides the number of studies included for each key question. Seventeen studies were included overall with 2 studies included in 2 questions.

Table 1: Key review questions for the evidence summary and relationship to the UK NSC screening criteria for a population

| | Criterion | Key questions | Studies included |
|---|---|---|---|
| The test | | | |
| 4 | There should be a simple, safe, precise and validated screening test. | What is the accuracy of screening tests to detect biliary atresia in newborns? | 10 |
| The screening programme | | | |
| 11 | There should be evidence from high quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (e.g. Down's syndrome or cystic fibrosis carrier screening), there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened. | Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes? | 6 |
| 13 | The benefit gained by individuals from the screening programme should outweigh any harms, for example from overdiagnosis, overtreatment, false positives, false reassurance, uncertain findings and complications. | | |

| | Criterion | Key questions | Studies included |
|---|---|---|---|
| Implementation criteria | | | |
| 15 | Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme. | What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK? | 3 |

# Methods

The current evidence summary was conducted by Solutions for Public Health (SPH), in keeping with the UK NSC evidence review process. Database searches were conducted on 20 August 2025 to identify studies relevant to the questions detailed in Table 1.

# Eligibility for inclusion in the review

The following review process was followed:

1. Each title and abstract was reviewed against the inclusion/exclusion criteria by one reviewer. Where the applicability of the inclusion criteria was unclear, the article was included at this stage to ensure that all potentially relevant studies were captured
2. Full-text articles required for the full-text review stage were acquired
3. Each full-text article was reviewed against the inclusion/exclusion criteria by one reviewer, who determined whether the article was relevant to one or more of the review questions
4. Any queries at the abstract or full-text stage were resolved through discussion with a second reviewer
5. The review was quality assured by a second senior reviewer, not involved with the writing of the review

Eligibility criteria for each question are presented in Table 2 below. Further details relating to the eligibility of studies are provided in Appendix 2.

A total of 1,007 unique references were identified and initially sifted by an information scientist by title and abstract for potential relevance. A reviewer assessed 169 titles and abstracts for further appraisal and possible inclusion in the evidence summary. Overall, 56 studies were identified as possibly relevant during title and abstract sifting and further assessed at full text. Appendix 2 contains a full PRISMA flow diagram (Figure 1), along with a table of the included publications and details of which questions these publications were identified as being relevant to (Table 18).

Table 2: Inclusion and exclusion criteria for the key questions

| Key question | Inclusion criteria | | | | | | | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|
| | Population | Target condition | Intervention/ index test | Reference standard | Comparator | Outcome | Study type | |
| What is the accuracy of screening tests to detect biliary atresia in newborns? | Newborns | Biliary atresia | Any tests used to detect biliary atresia using dried blood spots or stool colour cards | Any specific "gold standard", as determined by the study itself | None or any | Sensitivity<br>Specificity<br>PPV and NPV<br>Likelihood ratios<br>Area under the curve<br>False positives<br>Incidental findings/byproducts | Studies from UK populations to be prioritised. Studies from other countries, including those known to screen using stool colour cards, for example Japan and Taiwan, can also be reported<br>Studies in randomly assigned or consecutively enrolled populations and | Case reports, conference abstracts, comment/ editorials/ letters<br>Case-control studies, if randomly assigned or consecutively enrolled populations are identified |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | systematic reviews of these to be prioritised<br>Studies in the English language published since January 2012 | |
| What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK?<br>Sub-questions:<br>Is there any data to suggest that age at surgery has changed since 2016?<br>What proportion of affected infants have a late (>60 and >90 days) Kasai | Newborns /infants with biliary atresia | Biliary atresia | Kasai portoenterostomy | N/A | Reported mean age at procedure in screening programmes internationally or none | Median/mean age at which Kasai procedure is performed<br>Time to surgery | Cross-sectional studies, cohort studies, systematic reviews and other relevant surveillance reports<br>Studies in the English language published since November 2016 | Case reports, conference abstracts, comment/ editorials/ letters |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| portoenterostomy in the UK? | | | | | | | | |
| Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes?<br><br>Sub-question:<br><br>Does screening for biliary atresia using stool colour cards lead to harms? | Newborns /infants with biliary atresia | Biliary atresia | Screening for biliary atresia using stool colour cards | N/A | No screening<br><br>Screening using dried blood spot or any other screening pathway | Mean/median time to surgery<br><br>Age at surgery<br><br>Morbidity<br><br>Mortality<br><br>Number needing liver transplant following surgery<br><br>Misclassification<br><br>Increased monitoring<br><br>Overdiagnosis<br><br>Overtreatment<br><br>Stress and anxiety for carers | RCTs and systematic reviews of these to be prioritised<br><br>Studies in the English language published since January 2012 | Case reports, conference abstracts, comment/ editorials/ letters<br><br>Observational studies (e.g. case series, cohort, registry, survey data), if RCTs are identified |

## Appraisal for quality/risk of bias tool

The following tools were used to assess the quality and risk of bias of each study included in the review:

- Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for diagnostic studies
- JBI checklist for systematic reviews
- JBI checklist for cohort studies
- JBI checklist for case series

Results of the quality assessments are presented in the summary and appraisal of individual studies in Appendix 3.

## Methods of analysis/synthesis

A narrative synthesis of results is presented, structured by the UK NSC criteria and key questions. No meta-analyses were conducted.

## Databases/sources searched

Systematic searches of 4 databases (Medline, Embase, Cochrane Library and TRIPdatabse) were conducted to identify studies relevant to the questions detailed in Table 1.

The searches were conducted on 20 August 2025. Checks for additional relevant papers were made by checking the reference lists of relevant systematic reviews.

The search strategies are presented in Appendix 1.

# Question level synthesis

## Criterion 4 — Test accuracy

*There should be a simple, safe, precise and validated screening test*

Question 1 — What is the accuracy of screening tests to detect biliary atresia in newborns?

The 2017 UK NSC evidence summary [1] did not identify any studies that considered any form of screening for biliary atresia using dried bloodspots in a general newborn population. Studies about the effectiveness of screening using stool colour cards were identified and reported a sensitivity and specificity of 89% and 99.9% respectively.

The 2021 UK NSC evidence map [2] identified 2 studies about the accuracy of screening tests using dried blood spots (Harpavat et al. 2020 [26][‡] and Gong et al. 2020 [29]) and 5 studies about the accuracy of stool colour cards (Chiu et al. 2015 [30], Gu et al. 2020 [16], Gu et al. 2015 [15], Kong et al. 2016 [31], Woolfson et al. 2018 [21]). None of these studies were conducted in the UK.

### Eligibility for inclusion in the review

For this evidence summary, studies in randomly assigned or consecutively enrolled populations that reported the accuracy of dried blood spots or stool colour cards to detect biliary atresia in newborns were prioritised for inclusion. Case control studies on the accuracy of screening tests using dried blood spots were also considered as few studies from randomly assigned or consecutively enrolled populations were identified. Studies about the development of new potential screening tests and case control studies on the accuracy of stool colour cards were not considered due to the availability of higher level evidence.

### Description of the evidence

Of the 56 papers reviewed at full text, 27 related to this question. Ten studies met the criteria for inclusion. Appendix 2 contains a full PRISMA diagram (Figure 1), with a table of the included publications for each question (Table 18). The excluded studies are listed in Table 19 in Appendix 2.

Four studies for this question concerned dried blood spot tests:

- a systematic review and meta-analysis by Arshad et al. 2023 [32] searched for studies published up to September 2022 and included 4 studies reporting the diagnostic accuracy of bile acids of dried blood spot tests
- a retrospective study by Gong et al. 2020 [29] reported the diagnostic accuracy of free carnitine in dried blood spots. Some results for unconjugated and conjugated bilirubin in dried blood spots were also reported. The study population included 52,862 infants born in Shanghai, China between 2015 and 2017

[‡] On further review of the full text, and after considering additional information about this study obtained from studies published after the completion of the evidence map, this study by Harpavat et al. 2020 has been reclassified as concerning testing using serum blood rather than dried blood spots. The results of this and other similar studies do not meet the criteria for formal inclusion in this evidence summary but are discussed at the end of this question section

- a case control study by Lee et al. 2023 [33] reported the diagnostic accuracy of matrix metalloproteinase-7 (MMP-7) as a screening tool for biliary atresia using stored dried blood spots. The study included patients who attended a national hospital in Taiwan between 2018 and 2021, including 25 infants with biliary atresia and 107 controls with other congenital or perinatal conditions or who attended the hospital for a well-child visit
- a case control study by Xiao et al. 2022 [34] reported the diagnostic accuracy of metabolites as a screening tool for biliary atresia using stored dried blood spots. The study included patients born at one hospital in Shanghai, China between 2013 and 2020, including 21 infants with biliary atresia and 100 healthy controls

Two systematic reviews with meta-analysis and 5 individual studies concerned stool colour cards. These individual studies were included in one or both systematic reviews but are also separately included as they provide additional information about the study designs, populations and results:

- a systematic review and meta-analysis by Arshad et al. 2023 [32] searched for studies published up to September 2022 and included 5 studies reporting the diagnostic accuracy of stool colour cards
- a systematic review and meta-analysis by Gopal et al. 2024 [35] searched for studies published up to February 2023 and included 7 studies reporting the diagnostic accuracy of stool colour cards
- a screening study by Chiu et al. 2013 [30] reported the diagnostic accuracy of stool colour cards for pre-term (n=27) and term (n=170) infants with biliary atresia born in Taiwan between 2004 and 2010 after the implementation of a universal national screening programme
- a screening study by Gu et al. 2020 [15] reported the diagnostic accuracy of stool colour cards for 37,478 infants born in Sapporo, Japan between 2012 and 2015 and 27,561 infants born in Beijing, China between 2013 and 2014. Additional detail for the same cohort of infants born in Beijing was also available from Kong et al. 2016 [31]
- a screening study by Gu et al. 2015 [15] reported the diagnostic accuracy of stool colour cards for 264,071 infants born in Tochigi Prefecture, Japan between 1994 and 2011
- a screening study by Woolfson et al. 2018 [22] reported the diagnostic accuracy of stool colour cards for 87,583 infants born in British Columbia, Canada between 2014 and 2016

In addition, 3 studies were identified about the diagnostic accuracy of other tests for biliary atresia used within a screening context [25, 26, 27]. These studies are not formally included in this evidence summary as they tested for direct or conjugated bilirubin in serum blood rather than using dried blood spots. However, as this test was used to screen newborns for biliary atresia these studies are also discussed below.

## Discussion of findings

A study-level summary of data extracted from each included publication is presented in Appendix 3. In Appendix 3, publications are stratified by question and are presented in alphabetical order.

The diagnostic accuracy of dried blood spot tests is discussed first followed by the diagnostic accuracy of stool colour cards.

## Dried blood spot tests

### Sensitivity and specificity

The sensitivity and specificity of dried blood spot tests was reported in one systematic review with meta-analysis, one retrospective study and 2 case control studies. The tests reported considered bile acid, free carnitine levels, MMP-7 and for a combination of 3 metabolites (taurohyocholic acid, 2-hydroxyglutaric acid, and indoleacetic acid). The results are summarised in Table 3. Confidence intervals are provided where reported.

Table 3: Sensitivity and specificity of dried blood spot tests

| | Test | Sensitivity | Specificity |
|---|---|---|---|
| Arshad et al. 2023 [32] | Bile acids | 93.2% (95%CI 34.8 to 99.7) | 95.5% (95%CI 65.8 to 99.5) |
| Gong et al. 2020 [29] | Free carnitine | 85% | 85% |
| Lee et al. 2023 [33] | MMP-7 | 92.0% (95%CI 75.0 to 98.6) | 92.5% (95%CI 85.9 to 96.1) |
| Xiao et al. 2022 [34] | 3 metabolites | 90.48% (95%CI 69.62 to 98.83) | 92% (95%CI 84.84 to 96.48) |

### Positive predictive value (PPV) and negative predictive value (NPV)

The PPV and NPV of dried blood spot tests was reported in one systematic review with meta-analysis and one case control study.

The meta-analysis by Arshad et al. 2023 [32] reported an extremely low PPV (rounded to 0.0%) and NPV of 100% for bile acids based on a prevalence of 1 in 15,000. No confidence intervals were reported.

Lee et al. 2023 [33] reported a PPV of 71.9% (95%CI 56.3 to 87.5) and NPV of 98.0% (95%CI 95.3 to 100) respectively for MMP-7 in stored dried blood spots. The prevalence of biliary atresia used for the calculation was not clearly stated but appears to have been based on the fact that 18.9% of the 132 patients in this case control study had biliary atresia.

### Area under the curve

Outcomes relating to area under the curve for dried blood spot tests were reported in one retrospective study and 2 case control studies. The tests reported considered free carnitine levels, conjugated bilirubin, the ratio of conjugated bilirubin and total bilirubin, MMP-7 and a combination of 3 metabolites. The results are summarised in Table 4. Confidence intervals are provided where reported.

Table 4: Area under the curve of dried blood spot tests

| | Test | Area under the curve |
|---|---|---|
| Gong et al. 2020 [29] | Free carnitine | 0.92 |
| Gong et al. 2020 [29] | Conjugated bilirubin | 0.57 |
| Gong et al. 2020 [29] | Ratio of conjugated bilirubin and total bilirubin | 0.73 |
| Lee et al. 2023 [33] | MMP-7 | 93.7% (95%CI 87.7 to 99.7) |
| Xiao et al. 2022 [34] | 3 metabolites | 0.938 (95%CI 0.87 to 1.00) |

### Incidental findings

Outcomes relating to false positives were reported in one case control study.

In Lee et al. 2023 [33], 9 patients had an MMP-7 level greater than the cut-off used for a positive test but did not have biliary atresia. Diagnoses included choledochal cysts (n=2) and haemangioma (n=2) and single cases of parental nutrition-related cholestasis and hypothyroidism related to maternal Grave disease. For 3 patients there was no documented hepatobiliary disease.

### Summary of the results for dried blood spot tests

Limited information about the diagnostic accuracy of screening tests using dried blood spots was identified. The sensitivity and specificity for bile acids in dried blood spots was high in one meta-analysis at over 90% but the confidence intervals were wide and the PPV was extremely low (rounded to 0.0%). Two case control studies assessed the potential of 2 different tests using stored dried blood spots. These studies both reported sensitivities and specificities of over 90% but again the confidence intervals reported were wide. Some information on the accuracy of free carnitine levels, conjugated bilirubin and the ratio of conjugated bilirubin and total bilirubin in dried blood spots was available from one retrospective study but this was insufficient to draw any conclusions.

## Stool colour cards

### Sensitivity and specificity

The sensitivity and specificity of stool colour cards was reported in 2 systematic reviews with meta-analysis and 5 individual studies. Confidence intervals are provided where reported.

The 2 meta-analyses reported the sensitivity of stool colour cards as 87.9% (95%CI 80.4 to 92.8) (Arshad et al. 2023 [32]) and 79.6% (95%CI 70.6 to 86.4) (Gopal et al. 2024 [35]) respectively. In both studies the specificity was 99.9% (95%CI 99.9 to 99.9).

Chiu et al. 2013 [30] reported sensitivity to detect biliary atresia before 60 days of age for 170 term (92.8%) and 27 pre-term (96.3%) infants who had been diagnosed with biliary atresia in Taiwan. The difference between the groups was not statistically significant (p=0.798). Specificity was not reported.

Three studies reported sensitivity and specificity to detect biliary atresia one month after birth for infants screened using stool colour cards as part of screening programmes. These results are summarised in Table 5.

Table 5: Sensitivity and specificity of stool colour cards at one month

| | Location | Sensitivity | Specificity |
|---|---|---|---|
| Gu et al. 2020 [16] | Sapporo, Japan | 0% | 99.9% |
| Gu et al. 2020 [16] | Beijing, China | 100% | 99.9% |
| Gu et al. 2015 [15] | Tochigi Prefecture | 76.5% (95%CI 62.2 to 90.7) | 99.9% (95%CI 99.9 to 100) |
| Woolfson et al. 2018 [22] | British Columbia, Canada | 50% | 83% |

Gu et al. 2020 [16] and Kong et al. 2016 [31] also reported test performance at 4 months after birth for the infants born in Beijing, China. Sensitivity was 100% and specificity was 99.9% (95%CI 99.9 to 99.9).

### Positive predictive value (PPV) and negative predictive value (NPV)

The PPV and NPV of stool colour cards was reported in one systematic review with meta-analysis and 3 screening studies. Two studies reported PPV and NPV one month after birth in Tochigi Prefecture, Japan (Gu et al. 2015 [15]) and British Columbia, Canada (Woolfson et al. 2018 [22]) respectively. One study reported PPV at 4 months after birth in Beijing, China (Kong et al. 2016 [31]). These results are summarised in Table 6. Confidence intervals are provided where reported.

Table 6: PPV and NPV of stool colour cards

| | Prevalence | PPV | NPV |
|---|---|---|---|
| Arshad et al. 2023 [32] | 1 in 15,000 | 5.6% | 100% |
| Gu et al. 2015 [15] | 1.1 in 10,000 | 12.7%<br>(95%CI 8.2 to 17.3) | 99.9%<br>(95%CI 99.9 to 99.9) |
| Kong et al. 2016 [31] | 1.3 in 10,000 | 8.3%<br>(95%CI 2.7 to 19.4) | Not reported |
| Woolfson et al. 2018 [22] | 1 in 14,597 | 4% | 99% |

### False positives and false negatives

Outcomes relating to false positives were reported in one systematic review with meta-analysis and 2 screening studies. No confidence intervals were reported.

In the systematic review by Gopal et al. 2024 [35], the authors calculated that for a hypothetical cohort of 100,000 newborn infants with a biliary atresia prevalence of 1 in 15,000, one patient with biliary atresia would be missed and there would be 1,000 false positives.

Two screening studies reported false positive rates one month after birth for infants screened in Sapporo, Japan and Beijing, China (Gu et al. 2020 [16] and British Columbia, Canada (Woolfson et al. 2018 [22]) respectively. The false positive rate was 0.03% in Sapporo, 0.01% in Beijing and 0.09% in British Columbia. The false negative rate was 0% in Sapporo and Beijing but not reported in British Columbia.

Gu et al. 2020 [16] also reported results 4 months after birth for infants in Beijing. The false positive rate was 0.08% and the false negative rate was 0%.

Woolfson et al. 2018 [22] also reported that the biliary atresia screening centre in British Columbia received 75 phone calls from parents during the 2-year study period. One of these calls was for an infant who had biliary atresia. The authors described the other 74 calls as false-positive stool colour card readings. These included 9 calls for infants who were more than 6 months of age, 25 calls with normal stool colour or other non-related stool issues, 29 calls with transient abnormal stool that resolved without further investigation, 2 calls relating to alternative diagnoses and 9 cases with no cause identified.

### Incidental findings

Outcomes relating to incidental findings were reported in 3 screening studies.

Gu et al. 2020 [16] reported incidental findings for infants screened in Sapporo, Japan and Beijing, China. In Sapporo there were single patients diagnosed with neonatal intrahepatic cholestasis caused by citrin deficiency, neonatal hepatitis, suspected neonatal hepatitis and transient cholestasis. In Beijing there were single patients diagnosed with Alagille syndrome, cytomegalovirus hepatitis and Duchenne muscular dystrophy and one patient was receiving parenteral nutrition.

In Gu et al. 2015 [15] one infant in Tochigi Prefecture, Japan was diagnosed with Alagille syndrome.

In Woolfson et al. 2018 [22] there were 3 individual cases of Alagille syndrome, a urinary tract infection and thyroid disease for infants screened in British Columbia, Canada.

### Biliary atresia cases missed by screening programmes

Outcomes relating to missed cases were reported in 3 screening studies.

Gu et al. 2020 [16] reported that one of 3 cases of biliary atresia in Sapporo, Japan, and 2 of 4 cases in Beijing, China were diagnosed by clinical findings other than stool colour card screening. No further details were reported.

Gu et al. 2015 [15] reported details of 8 cases of biliary atresia that were not detected by a screening programme in Tochigi Prefecture, Japan. Of these, 2 infants were in a neonatal intensive care unit for more than one month after birth and their abnormal stool colour was overlooked and for 3 infants, their pale-pigmented stool colour was reported by their guardians at the one-month check-up but no further examination was conducted by their paediatricians. For the remaining 3 infants, 1 infant did not have abnormal stool colour at 1 month but the guardian did report pale stool at 1.5 months, the guardians of one infant did not use the stool colour card, and 1 infant was missed by the screening programme and identified via a national database.

Woolfson et al. 2018 [22] reported that 6 cases of biliary atresia were diagnosed after the introduction of the screening programme with 3 of the cases identified by the screening programme. For 2 of the 3 ‘missed’ cases, the family identified abnormal stool colour using the stool colour card but the infants were not referred by their physician in a timely manner. In one case, the abnormal stool colour was not detected by the family.

### Summary of the results for stool colour cards

Overall, there was variation in the sensitivity and PPV reported across the studies. Sensitivity ranged from 0% to 100%, however, 7 of the 9 sensitivity results reported were over 75%. PPV ranged from 4% to 12.7% in the 4 studies reporting this, with the lower figures reported being based on a prevalence rate that is more similar to that found in the UK (0.58 per 10,000 live births). There was less variation in specificity and NPV across studies. Specificity was 99.9% in 6 of the 7 studies reporting this, with the remaining result being 83%. In the 3 studies reporting NPV this was 99% or 100%. Confidence intervals, when reported, confirmed the variation in sensitivity and PPV results. False positive rates ranged from 0.01% to 0.09% with the most common incidental finding being Alagille syndrome. Commentary from the study authors about missed cases of biliary atresia suggest that in several cases, concerns were raised by the families based on the stool colour card but the infants were not referred for further investigation by physicians.

## Quality of the evidence base

The studies were either assessed using the JBI checklist for systematic reviews or the QUADAS-2 checklist.

The main limitations identified for the systematic review by Arshad et al 2023 [32] was a lack of detail about the reference standard used in the included studies, the thresholds applied or the processes used for the interpretation of results. Neither systematic review conducted a formal assessment of heterogeneity for the meta-analysis results. The systematic review authors conducted critical appraisal of the individual included studies and stated that most were at low risk of bias.

The results of the QUADAS-2 checklists for individual studies reporting the diagnostic accuracy of dried blood spot tests and stool colour cards are presented in Table 7 and Table 8 respectively. The main risk of bias across the studies related to the applicability of the populations, with most of the studies conducted in countries with a higher incidence of biliary atresia than the UK. An area of uncertainty in many of the studies related to the reference standard. Intervals between the index test and reference standard was generally not reported and only infants who screened positive received the reference standard. In most studies, the risk of bias from missed cases is likely to be low given the nature of the condition and fact that data about biliary atresia cases were taken from regional and/or national databases. Additional areas of risk of bias identified for the 3 individual studies assessing dried blood spots concerned the retrospective and/or case control study design and the lack of a pre-specified threshold for a positive screening test.

Table 7: QUADAS-2 scores summary indicating the areas of low, unclear or high risk of bias for dried blood spot tests

| Risk of bias | Gong et al. 2020 [29] | Lee et al. 2023 [33] | Xiao et al. 2022 [34] |
|---|---|---|---|
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Low | Low | Low |
| Case-control design avoided? | Unclear | High | High |
| Inappropriate exclusions avoided? | Low | Low | Low |
| Domain II: Index Test | | | |
| Index test results interpreted without knowledge of reference standard results? | High | High | High |
| Threshold pre-specified? | High | High | High |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Low | Unclear | Unclear |
| Reference standard results interpreted without knowledge of index test results? | Unclear | Unclear | Unclear |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Unclear |
| Did all participants receive same reference standard? | Unclear | Unclear | Unclear |
| All patients included in analysis? | Unclear | Low | Low |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | High | High | High |
| Applicable to UK screening test of interest? | Low | Low | Low |
| Target condition measured by reference test applicable to UK screening condition of interest? | Low | Unclear | Unclear |

Table 8: QUADAS-2 scores summary indicating the areas of low, unclear or high risk of bias for stool colour cards

| Risk of bias | Chiu et al. 2013 [30] | Gu et al. 2020 [16] & Kong et al. 2016 [31] | Gu et al. 2015 [15] | Woolfson et al. 2018 [22] |
|---|---|---|---|---|
| Domain I: Patient selection | | | | |
| Consecutive or random sample of population enrolled? | Low | Low | Low | Low |
| Case-control design avoided? | Low | Low | Low | Low |
| Inappropriate exclusions avoided? | Low | Low | Low | Low |
| Domain II: Index Test | | | | |
| Index test results interpreted without knowledge of reference standard results? | Low | Low | Low | Low |
| Threshold pre-specified? | Low | Low | Low | Low |
| Domain III: Reference standard | | | | |
| Reference standard likely to correctly classify condition? | Low | Unclear | Low | Low |
| Reference standard results interpreted without knowledge of index test results? | Unclear | Unclear | Unclear | Unclear |
| Domain IV: Test strategy flow and timing | | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Unclear | Unclear |
| Did all participants receive same reference standard? | Unclear | Unclear | Unclear | Unclear |
| All patients included in analysis? | Low | Low | Low | Low |
| Domain V: Applicability | | | | |
| Applicable to UK screening population of interest? | High | High | High | Low |
| Applicable to UK screening test of interest? | Low | Low | Low | Low |
| Target condition measured by reference test applicable to UK screening condition of interest? | Low | Unclear | Low | Low |

### Other screening tests for biliary atresia

Three studies from the USA reported the diagnostic accuracy of a 2-stage newborn screening strategy for biliary atresia based on bilirubin measurements in serum blood samples in centres that already routinely tested newborns for hyperbilirubinemia. The timing of the initial test varied, but all infants with a positive initial screening result underwent repeat testing at 2 weeks of age.

Harpavat et al. 2020 [26] measured direct or conjugated bilirubin within the first 60 hours of life in 123,279 infants born over a 3-year period at 14 hospitals in Texas. Seven infants with biliary atresia were identified with a sensitivity of 100% (95%CI 56.1 to 100.0) and a specificity of 99.9% (95% CI 99.9 to 99.9). Using a prevalence of 0.6 per 10,000, the authors reported PPV as 5.9% (95% CI 2.6 to 12.2) and NPV as 100.0% (95% CI 100 to 100). The authors also reported that 112 infants screened positive on initial and repeat testing but did not have biliary atresia.

In 2 further studies, Guthery et al. 2024 [25] measured direct bilirubin before discharge from hospital in 12,055 infants born in a 15-month period in 33 hospitals in Utah and Rabbani et al. 2025 [27] measured direct bilirubin at 24 to 48 hours of life in 3,880 infants born in a 15-month period at one centre in Texas. No cases of biliary atresia were identified in either study. Guthery et al. 2024 [25] reported a false positive rate of 0.12% and Rabbani et al. 2025 [27] reported a specificity of 99.7% and a false positive rate of 0.3%.

### Studies in progress

Horizon scanning-style searches were conducted on 8th September 2025 and updated on 23rd October 2025. These searched for any studies in progress that would meet the criteria for inclusion in this evidence summary. Four registries were searched (National Institutes of Health ClinicalTrials.gov, European Union Clinical Trials Register, WHO International Clinical Trials Registry Platform (ICTRP) and ScanMedicine).

The searches identified 2 studies further assessing the diagnostic accuracy of MMP-7 for detecting biliary atresia in newborns in China (ChiCTR2000034127, ChiCTR2400084093). The current status of these studies was not clear from the information available on the WHO ICTRP.

The searches identified one relevant study in progress on stool colour cards. This prospective study will evaluate the effectiveness of using a stool colour card for the early detection of biliary atresia within the first 6 months of life in Egypt (NCT07139717). This 12-month study, starting in September 2025 has an estimated enrolment of 100 infants.

## Summary of findings relevant to criterion 4: not met

The volume of evidence about the accuracy of screening tests using dried blood spots was limited to one meta-analysis, one retrospective study and 2 case control studies, all testing for different targets within dried blood spots. The details of test performance results reported in these studies was also limited. Sensitivities and specificities of over 90% were reported, but confidence intervals were wide increasing uncertainty in the results. Overall, the quantity and quality of the evidence available for dried blood spot tests is insufficient to draw any conclusions. Two studies in progress in China further assessing the diagnostic accuracy of MMP-7 in dried blood spots were identified. The current status of these studies was not clear. A study due to complete in September 2026 was identified assessing the diagnostic accuracy of stool colour cards in Egypt.

There was a greater volume and quality of evidence about the accuracy of screening tests using stool colour cards with data collected from consecutively enrolled populations as part of

screening programmes in several non-UK countries. There was a lack of consistency in the sensitivity results reported across different studies and confidence intervals, where reported, were wide reflecting the low number of biliary cases identified. There was also variation in the PPV's reported reflecting the differing incidence of biliary atresia in different countries. Two studies reported PPV for a prevalence that is more applicable to that found in the UK and this was between 4% and 6%. The specificity and NPV results were more consistent across studies and were generally high.

Overall, limitations in the quantity, quality, consistency and applicability of the evidence available for the accuracy of screening tests to detect biliary atresia in newborns means that this criterion is not met.

# Criterion 15 – Clinical management of the condition

*Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme*

Question 2 — What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK?

Sub-questions — Is there any data to suggest that age at surgery has changed since 2016?

What proportion of affected infants have a late (>60 and >90 days) Kasai portoenterostomy in the UK?

The 2017 UK NSC evidence summary [1] did not identify any studies to update the median time to surgery in the UK. This was previously reported in a 2012 UK NSC evidence summary [36] as a median of 54 days (range 7 to 209), with 10% of infants (n=44) 90 days or older at the time of the procedure. These figures were based on data for all infants with confirmed biliary atresia born in England and Wales between 1999 and 2009 and were considered comparable with performance in Taiwan which has had a screening programme using stool colour cards since 2004 [36].

The 2021 UK NSC evidence map identified 2 studies about reported age at surgery/time to surgery for biliary atresia in the UK (Durkin et al. 2017 [37], Williams et al. 2018 [38]).

## Eligibility for inclusion in the review

UK studies published since November 2016 were eligible for inclusion. Studies reporting age at surgery/time to surgery in other countries were not eligible for this question but were considered for Question 3.

## Description of the evidence

Of the 56 papers reviewed at full text, 10 related to this question. Three studies met the criteria for inclusion. Appendix 2 contains a full PRISMA diagram (Figure 1), along with a table of the included publications and details of which questions these publications were identified as being relevant to (Table 18). The excluded studies are listed in Table 19 in Appendix 2.

The 3 included studies comprised:

- a retrospective analysis of a prospectively maintained database by Davenport et al. 2025 [39] which reported data on age at surgery for 831 infants with biliary atresia treated at one of the 3 national paediatric liver units in England between 1999 and 2019
- a retrospective analysis of a prospectively maintained database by Durkin et al. 2017 [37] which reported data on age at surgery and time to surgery for 20 premature infants and 42 term infants diagnosed with biliary atresia at a single centre in the UK between 1988 and 2016
- a Lancet Standing Commission report on Liver Disease in the UK by Williams et al. 2018 [38], which presented data on age at referral to the 3 national paediatric liver units between 2012 and 2017 for 258 children diagnosed with biliary atresia

## Discussion of findings

A study-level summary of data extracted from each included publication is presented in Appendix 3. In Appendix 3, publications are stratified by question and are presented in alphabetical order.

### Age at referral

The paper by the Lancet Standing Commission on Liver Disease in the UK (Williams et al. 2018 [38]), reported that a total of 2,117 infants (<6 months old) with persistent conjugated jaundice were referred to the 3 national paediatric liver units between 2012 and 2017. Of these, 258 were subsequently diagnosed with extrahepatic biliary atresia. Amongst those diagnosed with extrahepatic biliary atresia, the median age of referral (no further details given) was 45 days (range 0 to 242).

### Age at surgery/time to surgery

Davenport et al. 2025 [39] reported that the median age at Kasai portoenterostomy for 831 infants treated at one of the 3 national UK paediatric liver units was 51 days (interquartile range (IQR) 39 to 64[§]). The improvement over time from 53 days (IQR 42 to 68) for the period 1999 to 2003 to 48 days (IQR 35 to 57) for the period 2014 to 2018 was statistically significant (p=0.0001).

Durkin et al. 2017 [37] reported that the median age at Kasai portoenterostomy for 42 term infants in one UK centre was 56 days (range 27 to 141) and for 20 premature infants was 65 days (range 16 to 323). The deferral time (time from diagnosis with liver biopsy to surgery) was 7 days in term infants and 11 days in premature infants (range not reported).

No data were identified to specifically address the sub-question about whether age at surgery has changed since 2016. However, the study by Davenport et al. 2025 [39] reported a statistically significant improvement in age at surgery over time up to 2018.

### Proportion of infants with a late (>60 and >90 days) Kasai portoenterostomy

No data were identified on the proportion of infants who had a Kasai portoenterostomy at more than 60 or 90 days old.

Davenport et al. 2025 [39] reported that 49 infants (5.6%) presented at >100 days old. This improved over time from 7.3% for the period 1999 to 2003 to 4.1% for the period 2014 to 2018. This difference was not statistically significant (p=0.1).

Durkin et al. 2017 [37] reported that 5 premature infants (25%) were operated at >100 days old. This figure was not provided for term infants.

The paper by the Lancet Standing Commission on Liver Disease in the UK (Williams et al. 2018 [38]), reported that 22% (56 out of 258 patients) were referred after 56 days of age. No information was provided on the interventions received by infants.

## Quality of the evidence base

The studies were assessed using the JBI checklist for case series. The main limitation identified for all 3 studies was an absence on information about how biliary atresia was detected. Additional limitations included a lack of information about study participants and uncertainty in the study by Durkin et al. 2017 [37] about whether the inclusion of participants was consecutive and complete.

[§] At one point in the paper this IQR was given as 39 to 67

## Summary of findings relevant to Criterion 15: met

Previous UK NSC evidence summaries have concluded that biliary atresia care in England and Wales has been optimised through the centralisation of care since 1999. Previous UK NSC evidence summaries have also concluded that the median age at surgery of 54 days (range 7 to 209) in the UK, based on data up to 2009, was comparable with the reported age at surgery in countries with stool colour card screening programmes.

The current evidence summary reported a median age at surgery in the UK of 51 days (IQR 39 to 64) based on data up to 2019. There was also evidence that this has improved over time and was 48 days (IQR 35 to 57) for the most recent time period available (2014 to 2018). No data were identified to update the proportion of infants that have a late (>60 days and >90 days) Kasai portoenterostomy in the UK. The study by Davenport et al. 2025 [39] did report that the proportion of infants presenting at >100 days old was 4.1% in the most recent data available (2014 to 2018).

Mean age at surgery after screening using stool colour cards was introduced is considered in Question 3 and ranged from 46 to 60 days.

This criterion has been met in previous UK NSC evidence summarises. Based on the more recent data identified for the current evidence summary, age at surgery in the UK is still comparable to the reported age at surgery in countries with stool colour card screening programmes. Therefore, this criterion is still met.

# Criterion 11 – Effectiveness of the screening programme

*There should be evidence from high quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity*

# Criterion 13 – Benefit and harms from screening programmes

*The benefit gained by individuals from the screening programme should outweigh any harms, for example from overdiagnosis, overtreatment, false positives, false reassurance, uncertain findings and complications*

Question 3 — Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes?

Sub-question — Does screening for biliary atresia using stool colour cards lead to harms?

The 2017 UK NSC evidence summary concluded that the benefit of screening using stool colour cards was unclear because although median time to surgery improved in countries with screening programmes it was not certain how much of the improvement was due to screening or whether the difference was clinically significant [1].

The 2021 UK NSC evidence map identified 5 studies about outcomes following screening with stool colour cards (Gu et al. 2015 [15], Lin et al. 2015 [28], Lee et al. 2016 [18], Gu & Matsui 2017 [40], Zheng et al. (2020) [19]). None of these studies were conducted in the UK.

## Eligibility for inclusion in the review

Studies in UK populations were of particular interest. Studies from any other countries that have implemented screening for biliary atresia using stool colour cards were also eligible for inclusion.

Studies reporting outcomes for infants with biliary atresia where there was no indication that they were detected through screening using stool colour cards were excluded. Studies reporting experiences of screening using stool colour cards, without a comparison to no screening, were not selected for inclusion as comparative studies were available. However, as these studies are about screening newborns for biliary atresia using stool colour cards, the results are briefly discussed below.

## Description of the evidence

Of the 56 papers reviewed at full text, 19 related to this question. Six studies met the criteria for inclusion. Appendix 2 contains a full PRISMA diagram (Figure 1), along with a table of the included publications and details of which questions these publications were identified as being relevant to (Table 18). The excluded studies are listed in Table 19 in Appendix 2.

The 6 studies for this question comprised:

- a screening study in one district (Tochigi Prefecture) in Japan by Gu et al. 2015 [15] which compared outcomes for 34 infants with biliary atresia diagnosed after the implementation of a stool colour card screening programme in 1994 with infants with biliary atresia born between 1987 and 1992 in the same district (n not stated)

- a retrospective cohort study by Gu & Matsui 2017 [40] set in one district (Tochigi Prefecture) in Japan which compared outcomes for 34 infants with biliary atresia diagnosed between 1994 and 2011 (after the implementation of a stool colour card screening programme in 1994) with 114 infants with biliary atresia born during the same time period in areas of Japan where screening had not been implemented
- a retrospective cohort study by Lee et al. 2016 [18] which used data from a national database in Taiwan to compare outcomes for infants with biliary atresia who underwent Kasai portoenterostomy before the implementation of a national screening programme using stool colour cards in 2004 (n=301) with infants diagnosed after screening was introduced (n=156)
- a retrospective cohort study by Lin et al. 2015 [28] which used data from a national database in Taiwan to compare outcomes for infants with biliary atresia born in 3 time periods: 1997 to 2001 (n=208), 2002 to 2006 (n=200) and 2007 to 2011 (n=132). 188 (90.4%), 174 (87.0%) and 122 (92.4%) infants had Kasai portoenterostomy in each of the time periods respectively. A national screening programme using stool colour cards was introduced in 2004
- a cohort study by Zheng et al. 2020 [19] which compared outcomes for infants with biliary atresia born in one hospital in China before stool colour cards were added to a handbook given to all pregnant women in 2015 (n=50) with infants born after screening was introduced (n=68). Before screening, 34 (68.0%) infants with biliary atresia had Kasai portoenterostomy and 57 (83.8%) had surgery after screening was introduced
- a screening study by Woolfson et al. 2018 [22] primarily focused on the diagnostic accuracy of stool colour cards as a screening test (see Question 1) but also reported results relating to misclassification from a province-wide screening programme using stool colour cards in British Columbia, Canada for infants born between 2014 to 2016

No studies were identified comparing outcomes following screening using stool colour cards and screening using dried blood spots or any other screening pathway. One study was identified about change in age at surgery within a screening context [26]. This study was not formally included in this evidence summary as it tested for direct or conjugated bilirubin in serum blood rather than using dried blood spots or stool colour cards. However, as the study relates to screening newborns for biliary atresia the results are also briefly discussed below.

## Discussion of findings

A study-level summary of data extracted from each included publication is presented in Appendix 3. In Appendix 3 publications are stratified by question and are presented in alphabetical order.

### Age at surgery

Age at Kasai portoenterostomy was reported in 5 studies.

Gu et al. 2015 [15] and Gu & Matsui 2017 [40] reported age at surgery after screening was implemented in one district in Japan compared to before screening was implemented or compared to an area with no screening respectively. Lee et al. 2016 [18] and Lin et al. 2015 [28] both reported age at surgery before and after the introduction of a national screening programme in Taiwan in 2004, using data from the same national database. Zheng et al. 2020 [19] reported age at surgery before and after the addition of stool colour cards to handbooks distributed to pregnant women at one hospital in China in 2015. The results are summarised in Table 9. Standard deviation is provided where reported.

Table 9: Mean (SD) age at Kasai portoenterostomy (days)

| | Before screening | No screening | After screening | p value |
|---|---|---|---|---|
| Gu et al. 2015 [15] | 70.3 | N/A | 59.7 (19.4) | p=0.003 |
| Gu & Matsui 2017 [40] | N/A | 68.1 (25.6) | 59.7 (19.4) | p=0.003 |
| Lee et al. 2016 [18] | 59.9 (76.4) | N/A | 48.2 (24.5) | p=0.064 |
| Zheng et al. 2020 [19] | 81 (± 12) | N/A | 56 (± 15) | p<0.05 |

In addition, Lin et al. 2015 [28] reported a statistically significant improvement in mean age at surgery over 3 time periods that encompassed the introduction of screening in 2004. This improved from 58.2 days (SD 42.0) in 1997 to 2001, 50.5 days (SD 30.8) in 2002 to 2006, and 46.0 days (SD 23.8) in 2007 to 2011 (p=0.006).

### Proportion of infants who received early or late Kasai portoenterostomy

Results relating to early or late Kasai portoenterostomy was reported in 5 studies. Studies differed in their reporting of this outcome, with some reporting the proportion who had surgery late and others the proportion who received surgery early. In Table 10 and Table 11, the results are presented as the proportion who received late surgery for ease of reference. The full study results, with confidence intervals where reported, are available in Appendix 3.

Table 10: Proportion of infants who received surgery ≥ 60 days of age

| | Before screening | No screening | After screening | Statistical significance |
|---|---|---|---|---|
| Gu et al. 2015 [15] | 66.0% | N/A | 44.1% | Not reported |
| Gu & Matsui 2017 [40] | N/A | 59.6% | 44.1% | Not significant |
| Lee et al. 2016 [18] | 31.6% | N/A | 26.3% | Not significant |

In addition, Lin et al. 2015 [28] reported that the proportion of infants that received Kasai portoenterostomy within 60 days of age was 76.6% in 1997 to 2001 (n=188), 88.5% in 2002 to 2006 (n=174) and 81.1% in 2007 to 2011 (n=122), p=0.285.

Zheng et al. 2020 [19] reported a statistically significant improvement in the proportion of infants diagnosed with biliary atresia before 60 days of age before (35.3%) and after (64.9%) the distribution of stool colour cards (p<0.05) but did not report the proportion who received surgery before or after 60 days of age.

Table 11: Proportion of infants who received surgery ≥ 90 days of age

| | Before screening | No screening | After screening | p value |
|---|---|---|---|---|
| Gu et al. 2015 [15] | 13.0% | N/A | 5.9% | p<0.05 |
| Gu & Matsui 2017 [40] | N/A | 20.2% | 5.9% | p=0.067 |

### Morbidity

Outcomes relating to morbidity was reported in 2 studies.

Lee et al. 2016 [18] reported a statistically significant reduction in the mean number of hospitalisations per case up to 2 years of age, for infants in Taiwan who had surgery after the implementation of screening (6.4 (SD 4.1, n=301) vs 5.0 (SD 3.2, n=156), p<0.001).

Zheng et al. 2020 [19] reported a statistically significant reduction in post-operative complications before (58.8%, n=34) and after (52.6%, n=57) the distribution of stool colour cards (p<0.05). Zheng et al. 2020 [19] also reported a statistically significant improvement in post-operative jaundice-free rate (47.1%, n=34 vs 54.4%, n=57, p<0.05).

### Mortality

Mortality was reported in 2 studies.

Lee et al. 2016 [18] reported a statistically significant reduction in mortality for infants with biliary atresia in Taiwan after the implementation of a national screening programme (47.8% (n=301) vs 21.2% (n=156), p<0.001).

Zheng et al. 2020 [19] reported a statistically significant reduction in mortality for infants with biliary atresia who received Kasai portoenterostomy in one hospital in China after the distribution of stool colour cards (20.6%% (n=34) vs 10.5% (n=57), p<0.05).

### Native liver survival

Results relating to native liver survival was reported in 5 studies.

Gu et al. 2015 [15] reported 5-year, 10-year and 15-year native liver survival rate as 87.6% (standard error (SE) 0.06), 76.9% (SE 0.08) and 48.5% (SE 0.11) respectively for the 34 infants diagnosed with biliary atresia after the introduction of screening in one district in Japan. These data were compared to historical rates in other countries from published studies but these data were not extracted as comparative evidence for this population was available from Gu & Matsui 2017 [40]. These results are presented in Table 12 (no SE reported).

Table 12: Probability of native liver survival reported by Gu & Matsui 2017 [40]

| | No screening | After screening | p value |
|---|---|---|---|
| 5-year | 53.1% | 87.6% | Not reported |
| 10-year | 43.9% | 76.9% | Not reported |
| 12.5-year | 36.6% | 48.5% | Not reported |

Gu & Matsui 2017 [40] also reported risk factors relating to native liver survival. The risk of the native liver not surviving was statistically significantly higher without screening (hazard ratio 2.61, 95%CI 1.20 to 5.70, p=0.016). There was no statistically significant difference if Kasai portoenterostomy was performed at more than 90 days old (hazard ratio 1.37, 95%CI 0.74 to 2.52, p=0.318).

Lee et al. 2016 [18] and Lin et al. 2015 [28] reported no statistically significant difference in the proportion of infants in Taiwan that had a liver transplant. In Lee et al. 2016 [18] this was 28.6% for infants who received Kasai portoenterostomy before the implementation of screening and 28.2% after screening was introduced, (p=0.934). In Lin et al. 2015 [28], the proportion of infants with biliary atresia needing liver transplantation was 32.7% in 1997 to 2001, 41.0% in 2002 to 2006 and 29.6% in 2007 to 2011, p=0.782. Lin et al. 2015 [28] also reported no

statistically significant difference in the liver transplantation rate for infants who received Kasai portoenterostomy at less than 60 days of age (25.6%) compared to infants who were more than 60 days of age (32.3%), p=0.133.

Zheng et al. 2020 [19] reported a statistically significant improvement in 2-year liver survival rate for infants who received Kasai portoenterostomy at one hospital in China after the distribution of stool colour cards (44.4% vs 52.6%, p<0.05). Zheng et al. 2020 [19] reported no statistically significant difference in liver transplantation rate (38.2% vs 40.4%, p>0.05). For infants who were 60 days old or less at Kasai portoenterostomy the liver transplantation rate was 30.1%. For infants who were aged more than 60 days this was 50.1%. No significance test was reported.

### Misclassification

Outcomes relating to misclassification were reported in 2 studies.

Gu et al. 2020 [16] stated that no infants with a false positive screening result underwent any invasive procedures.

Woolfson et al. 2018 [22] reported one false positive case where an infant identified by a screening programme in Canada had a Kasai portoenterostomy at 79 days old "*due to uncertain diagnosis, despite careful post-referral subspeciality evaluation".* This infant was ultimately diagnosed with Alagille syndrome.

### Summary of the results for screening using stool colour cards

Overall, age at surgery did improve following the introduction of screening across the studies with this improvement reaching statistical significance in most cases. The proportion of infants receiving late surgery and native liver survival also appeared to improve following the introduction of screening. There were fewer statistically significant improvements for these outcomes, possibly reflecting the small number of patients included in the analyses. There was some evidence from a small number of studies for improvements in morbidity and mortality following the introduction of screening.

## Quality of the evidence base

The studies were assessed using the JBI checklist for cohort studies. A limitation identified across most of the studies related to potential confounding factors. Whilst there was some discussion of confounding factors, the studies generally did not report any strategies to deal with these. Whilst outcomes were reported for infants diagnosed with biliary atresia after screening programmes were introduced, not all of the infants with biliary atresia included in the analyses were detected by the screening programmes. The attribution of the improvements reported to the screening programmes is therefore uncertain. In one study (Gu & Matsui 2017 [40]), outcome data were not available for deceased patients in the control group and in several studies outcomes were only reported for infants who underwent surgery and the follow-up period for more recently diagnosed patients was short. These factors introduce risk of bias to the reporting of longer-term outcomes such as native liver survival.

### Other screening studies for biliary atresia

Four studies reporting experiences of regional or pilot stool colour card screening programmes in Argentina, Canada, and Germany were identified by the searches for this evidence summary but did not meet the criteria for inclusion. These are briefly described below for information.

- Ramonet et al. 2013 [23] explored the use of a stool colour card by healthcare professionals at the one-month newborn check in one hospital in Argentina over a 3-year period. The authors reported that 4,239 of 12,484 infants were examined using the stool colour cards with 18 screening positive. There were no cases of biliary atresia
- Schreiber et al. 2014 [21] assessed the practicability of initiating a stool card screening programme in British Columbia in Canada. 6,187 families were asked to observe their infant's stool colour card daily for 4 weeks after birth. One infant was diagnosed with biliary atresia in this study. However, although a stool colour card was returned to the programme for this infant, this was returned early at 5 days after birth and was recorded as 'normal'. The authors described this infant as being missed by the screening programme
- Morinville et al. 2016 [20] assessed the feasibility of introducing a stool colour card programme in Quebec, Canada. Cards were distributed to the parents of 2,246 infants born at one hospital and the authors estimated stool colour card utilisation rates as between 82% and 100%. There were no cases of biliary atresia
- Madadi-Sanjani et al. 2021 [24] described their experiences of introducing a voluntary stool card screening programme in Lower-Saxony, Germany. Stool colour cards were provided to maternity wards over a 3-year period. Of the 13 infants diagnosed with biliary atresia during this 3-year period, 7 (54%) had received a stool colour card and one infant was referred for further assessment by their parents based on the stool colour card result. The authors stated that one of the 13 (7.7%) infants with biliary atresia was identified through the screening programme. This infant underwent surgery at 52 days of life and was described as jaundice-free with its native liver at 18 months-follow-up

In addition, one study compared age at Kasai portoenterostomy before and after the introduction of 2-stage screening for biliary atresia in Texas, USA using a screening test based on the analysis of serum blood (Harpavat et al. 2020 [26]). Infants in the post screening implementation group included those detected by hospitals in the screening study (6 infants), non-study hospitals that implemented a similar screening approach (7 infants) and non-study hospitals based on clinical symptoms without screening (6 infants). Infants undergoing surgery after screening was introduced in some hospitals in the region were statistically significantly younger compared to those undergoing surgery before screening was introduced (mean age 36 days (SD 22) vs 56 days (SD 19), between group difference 19 days (95%CI 7 to 32), p=0.004). Infants were also statistically significantly more likely to undergo surgery by 30 days (12.5% vs 57.9%, risk ratio 4.6 (95%CI 1.7 to 14.0), p=0.003).

## Summary of findings relevant to criteria 11 and 13: not met

The studies identified consistently reported that time to surgery improved for infants with biliary atresia after the introduction of screening using stool colour cards. There also appeared to be an improvement in the proportion of infants receiving late surgery and in other clinical outcomes although the statistical and clinical significance of these outcomes was less clear. There was no indication within the included studies that screening using stool colour cards is likely to lead to harm. However, none of the studies identified used a randomised controlled trial design. Therefore criterion 11 cannot be considered as met. Criterion 13, about the benefits of

screening outweighing the harms, could be considered as met within the context of the countries in which screening with stool colour cards has been introduced. However, the applicability of the results of these studies to the UK context is uncertain as the age at surgery achieved following the introduction of the screening programmes is similar to that reported in the UK without a screening programme being in place.

Overall, whilst improvements have been observed following the introduction of stool colour card screening in some countries, limitations in the study design, uncertainties about the clinical relevance of the improvements observed and the applicability to the UK context means that these criteria cannot be considered as met.

# Review summary

## Conclusions and implications for policy

This review focussed on systematically identifying new evidence to determine if reconsideration of the current recommendation for newborn screening for biliary atresia in the UK is required. The key questions focused on the accuracy of screening tests to detect biliary atresia in newborns, the reported age at surgery/time to surgery in the UK, and whether screening using stool colour cards improves time to surgery and clinical outcomes.

Evidence about the diagnostic accuracy of stool colour cards was available from consecutively enrolled populations from non-UK countries. There was considerable variation in the sensitivity and PPV values reported across the studies although specificity and NPV results were more consistent and were generally high. There was a lack of evidence to draw conclusions about the diagnostic accuracy of dried blood spots due to limitations in the quantity and quality of the studies available which considered different targets to test for within dried blood spots.

Previous UK evidence summaries have concluded that age at surgery in the UK is comparable to that achieved by countries with screening programmes in place. More recent data relating to age at surgery for biliary atresia in the UK were identified and suggest that this has improved further. However, no data were available to update the proportion of infants that have a late Kasai portoenterostomy in the UK.

Studies reporting outcomes from stool colour card screening programmes in other countries consistently report that time to surgery improved for infants with biliary atresia after the introduction of screening with no indication that screening is likely to cause harm. However, the applicability of the results to a UK context is uncertain as age at surgery after screening remains comparable to that reported in the UK without a screening programme being in place.

The volume, quality and direction of the new evidence published since the last evidence summary suggests that the current recommendation not to introduce a UK systematic population screening programme for biliary atresia in newborns should be retained.

## Limitations

This evidence summary was conducted according to the UK NSC evidence review process over a condensed period of time. The review only looked for peer-reviewed scientific work and does not include work published elsewhere (grey literature). Studies not available in the English language, abstracts and poster presentations were not eligible for inclusion. Given that these are accepted methodological adjustments for a rapid review, and that the searches included relevant systematic reviews published since this topic was last considered by the UK NSC, these limitations should not have led to the exclusion of any pivotal studies.

## Evidence uncertainties and gaps to be addressed

As in previous evidence summaries, there is a lack of evidence considering the diagnostic accuracy of dried blood spot tests in consecutively enrolled or randomly assigned populations. The studies identified about the use of dried blood spots to screen for biliary atresia have considered a range of different targets to test for suggesting ongoing uncertainty about the potential of using dried blood spots to screen for biliary atresia. The potential of using serum blood samples to screen newborns for biliary atresia has been explored in the USA.

Data were available to update the median age at surgery in the UK reported in previous evidence summaries. However, data were only available up to 2018 and no data were available to update the proportion of infants that have a late Kasai portoenterostomy in the UK. Further peer reviewed publications reporting the latest data on age at surgery in the UK would be helpful, particularly as the years since 2018 includes the years in which the COVID-19 pandemic impacted the delivery of many services.

# Appendix 1 — Search strategies

## Electronic databases

The search strategy included searches of the databases shown in Table 13.

Table 13: Databases searched

| Database | Platform | Searched on date | Date range of search |
|---|---|---|---|
| MEDLINE | Ovid SP | 20.08.2025 | 2012 to present |
| Embase | Ovid SP | 20.08.2025 | 2012 to present |
| The Cochrane Library, including:<br>Cochrane database of systematic reviews (CDSR)<br>Cochrane Central Register of Controlled Trials (CENTRAL) | Wiley Online | 20.08.2025 | 2012 to present |
| TRIP database | | 20.08.2025 | 2012 to present |

## Search Terms

Search terms for MEDLINE and Embase are shown in Table 14 and Table 15. Search terms for The Cochrane Library and TRIP databases are shown in Table 16 and Table 17.

Table 14: Search strategy for MEDLINE

| # | Search terms | Results |
|---|---|---|
| Question 1 | | |
| 1 | exp Infant, Newborn/ | 707972 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 926153 |
| 3 | 1 or 2 | 1247827 |
| 4 | Biliary Atresia/ | 3925 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 5894 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 170 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2334 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 2229 |
| 9 | 4 or 5 or 6 or 7 or 8 | 10963 |
| 10 | mass screening/ | 122313 |
| 11 | Dried Blood Spot Testing/ | 2406 |
| 12 | (screen* or detect*3 or test*3).ti,ab,kw,kf. | 7268465 |
| 13 | (dried blood spot* or finger prick* or fingerprick*).ti,ab,kw,kf. | 8708 |
| 14 | Feces/an, di | 11153 |
| 15 | ((stool? adj2 (colour* or color)) or scc).ti,ab,kw,kf. | 28901 |
| 16 | ((f?eces or f?ecal) adj2 (colour* or color)).ti,ab,kw,kf. | 51 |

| | | |
|---|---|---|
| 17 | (stool? adj3 (chart? or card? or test* or exam* or investigat*)).ti,ab,kw,kf. | 8571 |
| 18 | ((f?ces or f?ecal) adj3 (chart? or card? or test* or exam* or investigat*)).ti,ab,kw,kf. | 15342 |
| 19 | or/10-18 | 7324519 |
| 20 | 3 and 9 and 19 | 1003 |
| 21 | neonatal screening/ | 12872 |
| 22 | 9 and 21 | 65 |
| 23 | 20 or 22 | 1013 |
| 24 | exp animals/ not human/ | 5364786 |
| 25 | 23 not 24 | 995 |
| 26 | limit 25 to (english language and yr="2012 -Current") | 556 |
| Question 2 | | |
| 1 | exp Infant, Newborn/ | 707972 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 926153 |
| 3 | 1 or 2 | 1247827 |
| 4 | Biliary Atresia/ | 3925 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 5894 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 170 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2334 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 2229 |
| 9 | 4 or 5 or 6 or 7 or 8 | 10963 |
| 10 | Biliary Atresia/su [Surgery] | 2030 |
| 11 | Portoenterostomy, Hepatic/ | 990 |
| 12 | (portoenterostom* or hepatoportoenterostom* or kasai).ti,ab,kw,kf. | 2019 |
| 13 | bile drainage.ti,ab,kw,kf. | 557 |
| 14 | (surg* or operat* or procedure*).ti. | 1093273 |
| 15 | 10 or 11 or 12 or 13 or 14 | 1096325 |
| 16 | 3 and 9 and 15 | 1347 |
| 17 | exp animals/ not human/ | 5364786 |
| 18 | 16 not 17 | 1342 |
| 19 | limit 18 to (english language and yr="2016 -Current") | 526 |
| Question 3 | | |
| 1 | Newborn, Infant/ | 0 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 926153 |
| 3 | 1 or 2 | 926153 |
| 4 | biliary atresia/ | 3925 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 5894 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 170 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2334 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 2229 |
| 9 | 4 or 5 or 6 or 7 or 8 | 10963 |
| 10 | exp Time-to-Treatment/ | 11610 |
| 11 | ((time or wait* or age) adj2 surg*).ti,ab,kw,kf. | 60196 |
| 12 | 10 or 11 | 71033 |
| 13 | 3 and 9 and 12 | 73 |
| 14 | limit 13 to (english language and yr="2012 -Current") | 42 |

Table 15: Search strategy for Embase

| # | Search terms | Results |
|---|---|---|
| Question 1 | | |
| 1 | Newborn/ | 657765 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 1093075 |
| 3 | 1 or 2 | 1340926 |
| 4 | *bile duct atresia/ | 5187 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 8985 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 288 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2978 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 3206 |
| 9 | 4 or 5 or 6 or 7 or 8 | 15546 |
| 10 | screening/ or mass screening/ or screening test/ | 359648 |
| 11 | dried blood spot testing/ | 8475 |
| 12 | (screen* or detect*3 or test*3).ti,ab,kw,kf. | 9784081 |
| 13 | (dried blood spot* or finger prick* or fingerprick*).ti,ab,kw,kf. | 14006 |
| 14 | feces color/ | 722 |
| 15 | ((stool? adj2 (colour* or color)) or scc).ti,ab,kw,kf. | 43884 |
| 16 | ((f?eces or f?ecal) adj2 (colour* or color)).ti,ab,kw,kf. | 73 |
| 17 | (stool? adj3 (chart? or card? or test* or exam* or investigat*)).ti,ab,kw,kf. | 14327 |
| 18 | ((f?ces or f?ecal) adj3 (chart? or card? or test* or exam* or investigat*)).ti,ab,kw,kf. | 22232 |
| 19 | or/10-18 | 9871206 |
| 20 | 3 and 9 and 19 | 1584 |
| 21 | newborn screening/ | 25864 |
| 22 | 9 and 21 | 102 |
| 23 | 20 or 22 | 1599 |
| 24 | limit 23 to (english language and yr="2012 -Current") | 982 |
| 25 | limit 24 to ("remove clinical trial (clinicaltrials.gov) records" and "remove preprint records") | 963 |
| 26 | conference*.pt. or conference abstract/ | 6367237 |
| 27 | 25 not 26 | 578 |
| Question 2 | | |
| 1 | Newborn/ | 657765 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 1093075 |
| 3 | 1 or 2 | 1340926 |
| 4 | bile duct atresia/ | 10422 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 8985 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 288 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2978 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 3206 |
| 9 | 4 or 5 or 6 or 7 or 8 | 17526 |

| | | |
|---|---|---|
| 10 | bile duct atresia/su [Surgery] | 2969 |
| 11 | portoenterostomy/ | 2728 |
| 12 | (portoenterostom* or hepatoportoenterostom* or kasai).ti,ab,kw,kf. | 2944 |
| 13 | bile drainage.ti,ab,kw,kf. | 773 |
| 14 | (surg* or operat* or procedure*).ti. | 1279131 |
| 15 | 10 or 11 or 12 or 13 or 14 | 1284160 |
| 16 | 3 and 9 and 15 | 1915 |
| 17 | limit 16 to (english language and yr="2016 -Current") | 869 |
| 18 | limit 17 to ("remove clinical trial (clinicaltrials.gov) records" and "remove preprint records") | 842 |
| 19 | conference*.pt. or conference abstract/ | 6367237 |
| 20 | 18 not 19 | 624 |
| Question 3 | | |
| 1 | Newborn/ | 657765 |
| 2 | (newborn? or infan* or neonat*).ti,ab,kw,kf. | 1093075 |
| 3 | 1 or 2 | 1340926 |
| 4 | bile duct atresia/ | 10422 |
| 5 | ((biliary or bile duct?) adj atresia).ti,ab,kw,kf. | 8985 |
| 6 | ((bile duct? or hepatic duct?) adj5 (block* or scar*)).ti,ab,kw,kf. | 288 |
| 7 | ((hepatobiliary or hepato-biliary) adj disease?).ti,ab,kw,kf. | 2978 |
| 8 | (conjugated adj (bilirubin or bile acid*)).ti,ab,kw,kf. | 3206 |
| 9 | 4 or 5 or 6 or 7 or 8 | 17526 |
| 10 | time interval/ or diagnosis to treatment interval/ or time to treatment/ | 37572 |
| 11 | ((time or wait* or age) adj2 surg*).ti,ab,kw,kf. | 97365 |
| 12 | 10 or 11 | 133445 |
| 13 | 3 and 9 and 12 | 108 |
| 14 | limit 13 to (english language and yr="2012 -Current") | 58 |
| 15 | limit 14 to ("remove clinical trial (clinicaltrials.gov) records" and "remove preprint records") | 56 |
| 16 | conference*.pt. or conference abstract/ | 6367237 |
| 17 | 15 not 16 | 49 |

Table 16: Search strategy for The Cochrane Library databases

| # | Search terms | Results |
|---|---|---|
| #1 | MeSH descriptor: [Biliary Atresia] explode all trees | 71 |
| #2 | (((biliary OR (bile NEXT duct)) NEXT atresia)):ti,ab,kw OR (((bile OR hepatic) NEXT duct*)):ti,ab,kw OR ((hepatobiliary NEXT disease) OR (hepato-biliary NEXT disease)):ti,ab,kw OR ((conjucated NEXT (bilirubin OR (bile NEXT acid*)))):ti,ab,kw | 4451 |
| #3 | #1 OR #2 | 4451 |
| #4 | MeSH descriptor: [Infant, Newborn] explode all trees | 23817 |
| #5 | (newborn* or infan* or neonat*):ti,ab,kw | 98057 |

| | | |
|---|---|---|
| #6 | #4 OR #5 | 98057 |
| #7 | #3 AND #6 with Cochrane Library publication date Between Jan 2012 and Dec 2025, in Cochrane Protocols | 0 |
| #8 | #3 AND #6 with Publication Year from 2012 to 2025, in Trials | 78 |

Table 17: Search strategy for The TRIP database

| Search terms | Term group | Results |
|---|---|---|
| newborn* OR neonat* OR infan*, "biliary atresia" from_date:2012 to_date:2025 | Secondary Evidence | 8 |

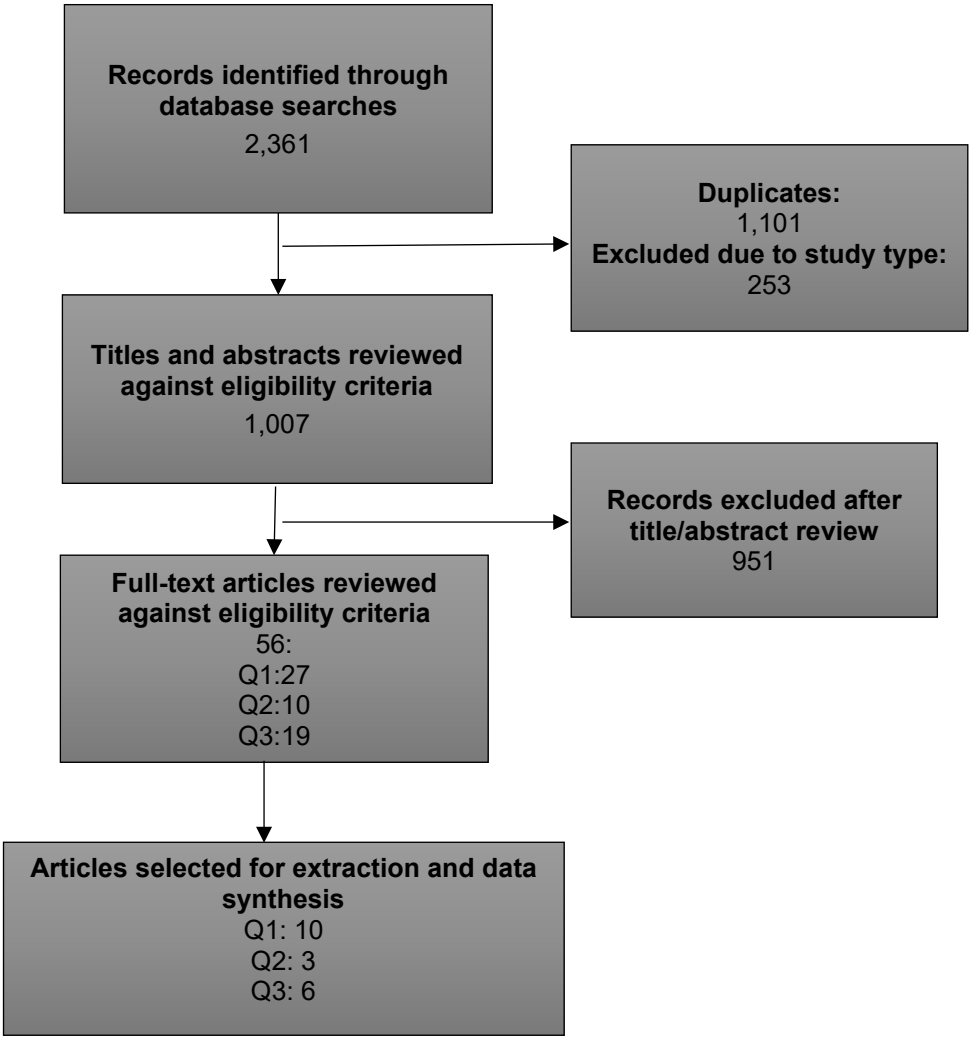
Results were imported into Endnote and duplicates removed.

# Appendix 2 — Included and excluded studies

## PRISMA flowchart

Figure 1 summarises the volume of publications included and excluded at each stage of the review. Seventeen publications were ultimately judged to be relevant to one or more review questions and were considered for extraction. Publications that were included or excluded after the review of full-text articles are detailed below.

Figure 1: Summary of publications included and excluded at each stage of the review




Two studies were included in both Questions 1 and 3.

# Publications included after review of full text articles

The 17 publications included after review of full texts are summarised in Table 18 below.

Studies were prioritised for extraction and data synthesis. It was planned *a priori* that the following approach would be taken to prioritise studies for extraction:

For question 1 about the accuracy of screening tests to detect biliary atresia in newborns, studies in randomly assigned or consecutively enrolled populations and systematic reviews of these were prioritised.

For question 3 about whether screening for biliary atresia using stool colour cards improves time to surgery and clinical outcomes, RCTs and systematic reviews of these were prioritised.

Publications not selected for extraction and data synthesis are clearly detailed in Table 19.

Table 18: Summary of publications included after review of full-text articles, and the questions each publication was identified as being relevant to

| Study | The test | The screening programme | Implementation criteria | Comments |
|---|---|---|---|---|
| Arshad et al. 2023 [32] | x | | | |
| Chiu et al. 2013 [30] | x | | | |
| Gong et al. 2020 [29] | x | | | |
| Gopal et al. 2024 [35] | X | | | |
| Gu et al. 2020 [16] | x | | | |
| Gu et al. 2015 [15] | x | | x | |
| Kong et al. 2016 [31] | x | | | |
| Lee et al. 2023 [33] | x | | | |
| Woolfson et al. 2018 [22] | x | | x | |
| Xiao et al. 2022 [34] | x | | | |
| Davenport et al. 2025 [39] | | x | | |
| Durkin et al. 2017 [37] | | x | | |
| Williams et al. 2018 [38] | | x | | |
| Gu & Matsui 2017 [40] | | | x | |
| Lee et al. 2016 [18] | | | x | |
| Lin et al. 2015 [28] | | | x | |
| Zheng et al. 2020 [19] | | | x | |

# Publications excluded after review of full text articles

Of the 56 publications considered at full text after the review of titles and abstracts, 32 were ultimately judged not to be relevant to this evidence summary. A further 7 studies were mentioned in the evidence summary as they provided information relating to newborn screening programmes for biliary atresia but were not selected for extraction and data synthesis. These 39 publications, along with reasons for exclusion, are listed in Table 19.

Table 19: Publications excluded after review of full text articles

| Reference | Reason for exclusion |
|---|---|
| Alam R, Nahid KL, Faruk MO, Rasna EH, Rukunuzzaman M. Use of stool color card as screening tool for biliary atresia in resource-constraint country. Gastroenterol. 2024;17(2):146-50. | Case control study about stool colour card test performance in infants who presented for assessment. Studies in consecutively enrolled populations available |
| Anonymous. Prediction of Native Liver survival in patients with biliary atresia 20 years after the Kasai procedure. Journal of Pediatric Surgery. 2025;60(3):162158. | No indication that the infants were screen detected. Does not address any of the review questions |
| Chan KWE, Lee KH, Wong HYV, Tsui SYB, Mou JWC, Tam YH. Impact of age of patient and experience of surgeon on the outcome after Kasai portoenterostomy: Can we delay the surgery? European Journal of Pediatric Surgery. 2020;06:06. | Not about the effectiveness of screening. Does not address any of the review questions |
| Franciscovich A, Vaidya D, Doyle J, Bolinger J, Capdevila M, Rice M, et al. PoopMD, a mobile health application, accurately identifies infant acholic stools. PLoS ONE [Electronic Resource]. 2015;10(7):e0132270. | Validation study assessing the performance of an app. Does not address any review questions |
| Fujishiro J, Suzuki K, Watanabe M, Uotani C, Takezoe T, Takamoto N, et al. Outcomes of Alagille syndrome following the Kasai operation: a systematic review and meta-analysis. Pediatric Surgery International. 2018;34(10):1073-7. | No patients had biliary atresia |
| Gong Z, Zheng L, Wang Y, Wu Y, Tian G, Lv Z. Quantification of bilirubin from dry blood spots using tandem mass spectrometry. New Journal of Chemistry. 2018;42(24):19701-6. | This paper is about the development of a potential screening test. A later paper about the use of the test in a screening context is included in the review |
| Gu YH, Yokoyama K, Mizuta K, Tsuchioka T, Kudo T, Sasaki H, et al. Stool color card screening for early detection of biliary atresia and long-term native liver survival: a 19-year cohort study in Japan. Journal of Pediatrics. 2015;166(4):897-902.e1. | Study was referenced in the 2017 UK NSC evidence summary. Does not provide new evidence |

| | |
|---|---|
| Guthery SL, Kyle Jensen M, Sean Esplin M, O'Brien E, Krong J, Srivastava R. Feasibility of biliary atresia newborn screening in an integrated health network. Journal of Pediatric Gastroenterology & Nutrition. 2024;79(5):954-61. | Mentioned in relation to Question 1. Screening test used serum blood rather than dried blood spot |
| Harpavat S, Garcia-Prats JA, Anaya C, Brandt ML, Lupo PJ, Finegold MJ, et al. Diagnostic yield of newborn screening for biliary atresia using direct or conjugated bilirubin measurements. JAMA. 2020;323(12):1141-50. | Mentioned in relation to Question 1. Screening test used serum blood rather than dried blood spot |
| Harpavat S, Ramraj R, Finegold MJ, Brandt ML, Hertel PM, Fallon SC, et al. Newborn direct or conjugated bilirubin measurements as a potential screen for biliary atresia. Journal of Pediatric Gastroenterology & Nutrition. 2016;62(6):799-803. | Study was referenced in the 2017 UK NSC evidence summary. Does not provide new evidence |
| Hoshino E, Moriwaki K, Morimoto K, Sakai K, Shimohata N, Konomura K, et al. Cost-effectiveness analysis of universal screening for biliary atresia in Japan. Journal of Pediatrics. 2023;253:101-6.e2. | Cost-effectiveness not considered in this evidence summary. Studies used to inform model were all in the search results and separately considered for inclusion |
| Hoshino E, Muto Y, Sakai K, Shimohata N, Urayama KY, Suzuki M. Age at surgery and native liver survival in biliary atresia: a systematic review and meta-analysis. European Journal of Pediatrics. 2023;182(6):2693-704. | Review is not about the impact of screening on outcomes. Individual potentially in scope studies separately considered |
| Huang SY, Yeh CM, Chen HC, Chou CM. Reconsideration of laparoscopic Kasai operation for biliary atresia. Journal of Laparoendoscopic & Advanced Surgical Techniques Part A. 2018;28(2):229-34. | No indication that the infants were screen detected. Does not address any of the review questions |
| Jimenez-Rivera C, Jolin-Dahel KS, Fortinsky KJ, Gozdyra P, Benchimol EI. International incidence and outcomes of biliary atresia. Journal of Pediatric Gastroenterology & Nutrition. 2013;56(4):344-54. | Study was referenced in the 2017 UK NSC evidence summary. Does not provide new evidence |
| Kastenberg ZJ, Deneau MR, O'Brien EA, Huynh K, Book LS, Srivastava R, et al. Fractionated bilirubin among 252 892 Utah newborns with and without biliary atresia: A 15-year historical birth cohort study. Journal of Pediatrics. 2023;257:113339. | Case control study assessing a test using serum blood |
| Lacaille F, Nicastro E, Czubkowski P, Goncalves CC, Le Thi TG, Koletzko S. Awareness, referral and age at Kasai surgery for biliary atresia in Europe: A survey of the Quality-of-Care Task Force of ESPGHAN. Journal of Pediatric Gastroenterology & Nutrition. 2024;78(6):1374-82. | No separate data on UK outcomes. Does not address any review questions |
| Lam L, Musaad S, Kyle C, Mouat S. Utilization of reflex testing for direct bilirubin in the early recognition of biliary atresia. Clinical Chemistry. 2017;63(5):973-9. | Focus on performance of diagnostic tests for biliary atresia rather than screening tests |

| | |
|---|---|
| Liao FM, Chang KC, Wu JF, Chen HL, Ni YH, Chang MH. Direct bilirubin and risk of biliary atresia. Pediatrics. 2022;149(6):01. | Diagnostic accuracy of a test using serum blood and not within the context of a screening programme |
| Madadi-Sanjani O, Blaser J, Voigt G, Kuebler JF, Petersen C. Home-based color card screening for biliary atresia: the first steps for implementation of a nationwide newborn screening in Germany. Pediatric Surgery International. 2019;35(11):1217-22. | Assessing acceptability of stool colour cards amongst maternity hospital staff. Does not address any of the review questions |
| Madadi-Sanjani O, Kuebler JF, Uecker M, Pfister ED, Baumann U, Kunze-Hullmann B, et al. Province-wide stool color card screening for biliary atresia in Lower-Saxony: Experiences with passive distribution strategies and results. Int. 2021;7(4):04. | Experiences of screening in Germany. No comparison to no screening. Mentioned in the introduction |
| Madsen SS, Kvist N, Thorup J. Increased conjugated bilirubin is sufficient to initiate screening for biliary atresia. Dan Med J. 2015;62(8):A5114. | Retrospective analysis of biliary atresia cases to inform recommendations for diagnostic testing. Does not address any review questions |
| Matcovici M, Stoica I, Smith K, Davenport M. What makes a "successful" Kasai portoenterostomy "unsuccessful"? Journal of Pediatric Gastroenterology & Nutrition. 2023;76(1):66-71. | This paper reports data from a UK centre that featured in 2 papers that have already been included (Durkin et al. 2017 [37] and Davenport et al. 2025 [39]). This paper focuses on whether age at Kasai portoenterostomy is a predictor of need for subsequent liver transplant. It does not report any outcomes or more recent data compared to the studies already included |
| Morinville V, Ahmed N, Ibberson C, Kovacs L, Kaczorowski J, Bryan S, et al. Home-based screening for biliary atresia using infant stool color cards in Canada: Quebec feasibility study. Journal of Pediatric Gastroenterology & Nutrition. 2016;62(4):536-41. | Experiences of screening in Canada. No comparison to no screening. Mentioned in the introduction |
| Noorulla F, Dedon R, Maisels MJ. Association of early direct bilirubin levels and biliary atresia among neonates. JAMA Network Open. 2019;2(10):e1913321. | Retrospective review of patients diagnosed with biliary atresia, exploring the possibility of using direct bilirubin as a marker for screening |
| Okubo R, Nio M, Sasaki H. Impacts of early Kasai portoenterostomy on short-term and long-term outcomes of biliary atresia. Hepatology Communications. 2021;5(2):234-43. | No indication that the infants were screen detected. Does not address any of the review questions |

| | |
|---|---|
| Rabbani T, Shah J. Newborn screening for biliary atresia using direct bilirubin: An implementation science study. Journal of Medical Screening. 2025;32(2):61-6. | Mentioned in relation to Question 1. Screening test used serum blood rather than dried blood spot |
| Ramonet MD, Gomez S, Morise S, Parga L, Caglio P, De Micheli M, et al. Early detection of neonatal cholestasis using a stool color card screening. [Spanish, English]. Archivos Argentinos de Pediatria. 2013;111(2):135-8. | Experiences of screening in Argentina. No comparison to no screening. Mentioned in the introduction |
| Schreiber RA, Masucci L, Kaczorowski J, Collet JP, Lutley P, Espinosa V, et al. Home-based screening for biliary atresia using infant stool colour cards: a large-scale prospective cohort study and cost-effectiveness analysis. Journal of Medical Screening. 2014;21(3):126-32. | Experiences of screening in Canada. No comparison to no screening. Mentioned in the introduction |
| Shen Z, Zheng S, Dong R, Chen G. Saturation of stool color in HSV color model is a promising objective parameter for screening biliary atresia. Journal of Pediatric Surgery. 2016;51(12):2091-4. | Assessing potential for apps to read stool colour cards. Does not address any review questions |
| Sun S, Zheng S, Lu X, Chen G, Ma Y, Chen L, et al. Clinical and pathological features of patients with biliary atresia who survived for more than 5 years with native liver. Pediatric Surgery International. 2018;34(4):381-6. | Outcomes following surgery for biliary atresia. Does not address any review questions |
| Tomita H, Shimojima N, Sasaki H, Shimotakahara A, Yamada Y, Kuroda T, et al. Predicting cirrhosis and poor outcomes of bile drainage surgery for biliary atresia: A multicentric observational study in Japan. Ann Surg. 2024;279(4):692-8. | No indication that the infants were screen detected. Does not address any review questions |
| Uecker M, Prehn C, Janzen N, Adamski J, Vieten G, Petersen C, et al. Infants with biliary atresia exhibit an altered amino acid profile in their newborn screening. Metabolomics. 2024;20(5):109. | Exploring potential screening tests. Does not report diagnostic accuracy |
| Wang L, Yang Y, Chen Y, Zhan J. Early differential diagnosis methods of biliary atresia: a meta-analysis. Pediatric Surgery International. 2018;34(4):363-80. | Focus on performance of diagnostic tests in children with cholestasis rather than screening tests |
| Xie C, Wang P, Wang D, Jin Y, Li S, Zhao Y, et al. Is performing the Kasai portoenterostomy in the neonatal period associated with a better prognosis? A single-center, retrospective cohort study from China. BMC Pediatrics. 2025;25(1):454. | No indication that the infants were screen detected. Does not address any review questions |
| Yachha SK, Das MC, Kumar P, Sharma L, Singh SK, Sen Sarma M, et al. Development of integrated neonatal cholestasis card for early recognition and referral of neonatal cholestasis. Indian Journal of Gastroenterology. 2020;39(6):584-90. | Case control study about the development of an integrated card with urine and stool colour identification. Test out of scope |

| | |
|---|---|
| Ye C, Gao W. Predictors of outcome among children with biliary atresia: a single centre trial. Peerj. 2025;13:e19001. | Infants were not detected as part of a screening programme. Does not address any review questions |
| Zhang Y, Li T, Wang T, Ji Q, Zhan J. Comparison for the diagnostic performance of early diagnostic methods for biliary atresia: a systematic review and network meta-analysis. Pediatric Surgery International. 2024;40(1):146. | Focus on performance of diagnostic tests in children with cholestasis rather than screening tests |
| Zheng Q, Zhang S, Ge L, Jia J, Gou Q, Zhao J, et al. Investigation into multi-centre diagnosis and treatment strategies of biliary atresia in mainland China. Pediatric Surgery International. 2020;36(7):827-33. | About the management of biliary atresia in China. Does not address any review questions |
| Zhou K, Lin N, Xiao Y, Wang Y, Wen J, Zou GM, et al. Elevated bile acids in newborns with Biliary Atresia (BA). PLoS ONE [Electronic Resource]. 2012;7(11):e49270. | Study was referenced in the 2017 UK NSC evidence summary. Does not provide new evidence |

# Appendix 3 — Summary and appraisal of individual studies

## Data extraction and appraisal for quality and risk of bias

### Question 1: What is the accuracy of screening tests to detect biliary atresia in newborns?

#### Dried blood spot

##### Systematic reviews

##### Table 20: Arshad et al. 2023 [32]

| | |
|---|---|
| Publication | Arshad A, Gardiner J, Ho C, Rees P, Chadda K, Baker A, et al. Population-based screening methods in biliary atresia: a systematic review and meta-analysis. Archives of Disease in Childhood. 2023;108(6):468-73. |
| Study details | Systematic review and meta-analysis |
| Study objectives | To investigate tested methods of population-based biliary atresia screening |
| Study setting | The systematic review was conducted in the UK |
| Inclusions | Observational studies, published between January 1975 and 12 September 2022, reporting outcomes of a biliary atresia screening method<br>Included studies were prospective cohort studies, cross-sectional studies, retrospective cohort studies and case control-studies |
| Exclusions | Studies that were opinions, reviews or non-peer-reviewed letters. Studies not published in English |
| Population | Newborns with biliary atresia<br>The systematic review considered various aspects of screening for biliary atresia. Only data that meet the inclusion criteria for this evidence summary have been extracted |
| Test | Bile acids in blood spots (4 studies published between 1993 and 2020) |
| Comparator/reference standard | Not stated |
| Outcomes | Meta-analysis results for sensitivity and specificity |
| Sensitivity | 93.2% (95%CI 34.8 to 99.7) |
| Specificity | 95.5% (95%CI 65.8 to 99.5) |
| PPV | Rounded to 0.0% |
| NPV | 100% |

Quality appraisal using the JBI checklist for systematic reviews

| Question | Assess-ment | Supporting Information |
|---|---|---|
| Is the review question clearly and explicitly stated? | Yes | |
| Were the inclusion criteria appropriate for the review question? | Yes | |
| Was the search strategy appropriate? | Yes | |
| Were the sources and resources used to search for studies adequate? | Yes | |

| | | |
|---|---|---|
| Were the criteria for appraising studies appropriate? | Yes | Two reviewers independently screened titles and abstracts |
| Was critical appraisal conducted by 2 or more reviewers independently? | Unclear | Not stated |
| Were there methods to minimize errors in data extraction? | Yes | Two reviewers independently extracted data |
| Were the methods used to combine studies appropriate? | Unclear | No formal assessment of heterogeneity |
| Was the likelihood of publication bias assessed? | Yes | |
| Were recommendations for policy and/or practice supported by the reported data? | Yes | |
| Were the specific directives for new research appropriate? | Yes | |
| Other comments | The 4 studies reporting bile acid in dried blood spots were from Japan, China and the UK. The UK study was published in 1999, outside the timeframe for the search for this evidence summary.<br><br>Risk of bias for the included studies was assessed using the Newcastle-Ottawa Scale. The included studies mostly had a low risk of bias.<br><br>Limited details were provided about the individual studies and the screening tests assessed. No details were provided about the comparator/reference standard used, the thresholds applied or the processes used for the interpretation of the results.<br><br>Case-control studies were in scope for the systematic review. However, most of the included studies were prospective cohort studies. It is not clear if any of the included studies were conducted in randomly assigned or consecutively enrolled populations.<br><br>No assessment of heterogeneity was conducted. The authors stated that this could not be conducted but did not provide any further detail.<br><br>The authors calculated PPV and NPV for differing prevalence of biliary atresia. The results for a prevalence of 1 in 15,000 were extracted as this is similar to the prevalence of biliary atresia in the UK. For information, the PPV and NPV for a prevalence of 1 in 10,000 was 0.0% and 100% respectively.<br><br>The systematic review also included one study testing blood carnitine levels in dried blood spots (Gong et al. 2020 [29]). This study is separately included in this evidence summary as no pooled data for this test were available from the systematic review.<br><br>The assessment of the diagnostic accuracy of stool colour cards reported in this systematic review is presented in Table 24. | |

## Individual studies

### Table 21: Gong et al. 2020 [29]

| | | | |
|---|---|---|---|
| Publication | Gong Z, Wu Y, Zheng L, Chen L, Lv Z. Can free carnitine or bilirubin in blood be used in neonatal screening for biliary atresia? European Journal of Pediatric Surgery. 2020;30(5):459-64. | | |
| Study details | Retrospective study | | |
| Study objectives | To investigate the efficiency of free carnitine, conjugated bilirubin and unconjugated bilirubin in dry blood spots measured using tandem mass spectrometry for screening for biliary atresia | | |
| Study setting | Four maternity hospitals in Shanghai, China | | |
| Inclusions | Newborns born between January 2015 and June 2017 | | |
| Exclusions | None stated | | |
| Population | 52,862 newborns screened using tandem mass spectrometry<br>The study also included 26 patients with biliary atresia admitted to one of the 4 centres who were not screened using tandem mass spectrometry | | |
| Test | Dried blood spot samples collected by heel prick (aged 3 to 14 days) and analysed for free carnitine (using the 95% percentile (30 µmol/L) as a cut-off), unconjugated bilirubin and conjugated bilirubin using tandem mass spectrometry | | |
| Comparator/reference standard | Liver biopsy and/or intraoperative cholangiogram | | |
| Outcomes | 7 cases of biliary atresia were detected using tandem mass spectrometry. No confidence intervals reported | | |
| | Free carnitine | Conjugated bilirubin | Ratio of conjugated bilirubin and total bilirubin |
| Area under the receiving operating characteristic curve | 0.92 | 0.57 | 0.73 |
| Sensitivity | 85% | Not reported | Not reported |
| Specificity | 85% | Not reported | Not reported |
| Quality appraisal using the QUADAS-2 tool | | | |
| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Newborns born at 4 centres in a 30-month period |
| Case-control design avoided? | Unclear | Unclear | The test performance results appear to use data from the neonates who were screened and other neonates with biliary atresia who were admitted to one of the 4 centres |
| Inappropriate exclusions avoided? | Yes | Low | No exclusions stated |
| Domain II: Index Test | | | |

| | | | |
|---|---|---|---|
| Index test results interpreted without knowledge of reference standard results? | No | High | Retrospective analysis of patients after biliary atresia status was known |
| Threshold pre-specified? | No | High | Different thresholds considered |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Yes | Low | |
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Retrospective analysis of patients after biliary atresia status was known |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | No | Unclear | Retrospective analysis of patients after biliary atresia status was known |
| All patients included in analysis? | Unclear | Unclear | The infants included in the analyses reported was not clear |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher prevalence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Yes | Low | |
| Other comments | This retrospective study considered several potential screening tests for biliary atresia. Dried blood spots were collected from infants born at the study centres. Samples from children admitted to the study centres were also used for some of the analyses reported. It is not clear which infants were included in each of the analyses reported in the study.<br><br>The authors stated that free carnitine, conjugated bilirubin and unconjugated bilirubin levels are elevated in infants with biliary atresia and were higher than the normal range one month after birth. The authors stated that these tests should not be used to screen newborns for biliary atresia using dried blood spots as levels may not be elevated in newborns resulting in high rates of false positives and false negatives.<br><br>The report included infants screened over a 2-year period up to 2017 in a county with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear.<br><br>This study also investigated the performance of direct bilirubin and total bilirubin measured in a control sample of 400 paediatric patients | | |

aged 0 to 15 years with various diseases. These data are out of scope for this evidence summary.

## Table 22: Lee et al. 2023 [33]

| | |
|---|---|
| Publication | Lee CS, Ni YH, Chen HL, Wu JF, Hsu HY, Chien YH, et al. A pilot study of biliary atresia newborn screening using dried blood spot matrix metalloproteinase-7. Journal of Pediatric Gastroenterology & Nutrition. 2023;76(4):418-23. |
| Study details | Case control study |
| Study objectives | To evaluate matrix metalloproteinase-7 (MMP-7) as a potential screening tool for biliary atresia using dried blood spots |
| Study setting | One national hospital in Taiwan |
| Inclusions | Patients with stored dried blood spots attending one hospital between September 2018 and May 2021 |
| Exclusions | Patients with no stored dried blood spot or a dried blood spot that had been stored for more than 5 years |
| Population | 25 newborns with biliary atresia<br>107 controls with other congenital or perinatal conditions or who attended the hospital for a well-child visit |
| Test | Stored dried blood spot samples from newborn screening centres collected within 3 days of birth tested for MMP-7 using a sensitive enzyme-linked immunosorbent assay (ELISA) |
| Comparator/reference standard | Diagnosis of biliary atresia. No further details provided |
| Outcomes | The test performance of MMP-7 was assessed using a cut-off value of 8.0 ng/mL |
| Area under the curve | 93.7% (95%CI 87.7 to 99.7) |
| Sensitivity | 92.0% (95%CI 75.0 to 98.6) |
| Specificity | 92.5% (95%CI 85.9 to 96.1) |
| PPV | 71.9% (95%CI 56.3 to 87.5) |
| NPV | 98.0% (95%CI 95.3 to 100) |
| Incidental findings | 9 patients had an MMP-7 level greater than 8.0 ng/mL but did not have biliary atresia. These infants were diagnosed with:<br>• choledochal cysts (n=2)<br>• haemangioma (n=2)<br>• parental nutrition-related cholestasis (n=1)<br>• hypothyroidism related to maternal Grave disease (n=1)<br>• no documented hepatobiliary disease (n=3) |

Quality appraisal using the QUADAS-2 tool

| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
|---|---|---|---|
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Newborns born in Taiwan with stored dried blood spots |
| Case-control design avoided? | No | High | Test performance assessed using known cases and controls |
| Inappropriate exclusions avoided? | Yes | Low | |
| Domain II: Index Test | | | |

| | | | |
|---|---|---|---|
| Index test results interpreted without knowledge of reference standard results? | No | High | Retrospective analysis of patients after biliary atresia status was known |
| Threshold pre-specified? | No | High | Different thresholds considered |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Unclear | Unclear | No details provided |
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Retrospective analysis of patients after biliary atresia status was known |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | Unclear | Unclear | No details provided |
| All patients included in analysis? | Yes | Low | |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher incidence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Unclear | Unclear | No details provided |
| Other comments | This retrospective study considered several potential cut-off levels for MMP-7. Stored dried blood spots were analysed for infants who attended the study hospital was a variety of reasons. The analysis was based on samples from infants who were known to have biliary atresia and controls who either had other congenital or perinatal conditions or who attended the hospital for a well-child visit.<br><br>The prevalence of biliary atresia used to calculate PPV and NPV was not clearly stated but the authors noted that 18.9% of the patients in their study had biliary atresia.<br><br>The authors concluded that MMP-7 dried blood spot analysis can be used to distinguish biliary atresia from other conditions as early as 3 days of age.<br><br>The study analysis used dried blood spots for patients who attended a national hospital over a 3-year period up to 2021 in a county with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear. | | |

This study also reported test performance for additional cut-off levels. These data were not extracted as they were not highlighted by the authors as the best performing level.

## Table 23: Xiao et al. 2022 [34]

| | | | |
|---|---|---|---|
| Publication | Xiao Y, Zhou Y, Zhou K, Cai W. Targeted metabolomics reveals birth screening biomarkers for biliary atresia in dried blood spots. J Proteome Res. 2022;21(3):721-6. | | |
| Study details | Case control study | | |
| Study objectives | To identify potential biliary atresia biomarkers in dried blood spots using ultraperformance liquid chromatography-triple quadrupole mass spectrometry-based targeted metabolomics profiling | | |
| Study setting | One hospital in Shanghai, China | | |
| Inclusions | Dried blood spots from infants born between September 2013 and December 2020 | | |
| Exclusions | None stated | | |
| Population | 21 infants with biliary atresia<br>100 healthy controls | | |
| Test | Unused dried blood spots originally collected for phenylketonuria screening | | |
| Comparator/reference standard | Diagnosis of biliary atresia. No further details provided | | |
| Outcomes | The test performance of a combination of 3 metabolites (taurohyocholic acid, 2-hydroxyglutaric acid, and indoleacetic acid) was modelled using a cut-off value of -0.336 | | |
| Area under the receiving operating curve | 0.938 (95%CI 0.874 to 1.000) | | |
| Sensitivity | 90.48% (95%CI 69.62 to 98.83) | | |
| Specificity | 92% (95%CI 84.84 to 96.48) | | |
| Quality appraisal using the QUADAS-2 tool | | | |
| **Question** | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Newborns born at one hospital over a 7-year period |
| Case-control design avoided? | No | High | Test performance was assessed using known cases and controls |
| Inappropriate exclusions avoided? | Yes | Low | No exclusions stated |
| Domain II: Index Test | | | |
| Index test results interpreted without knowledge of reference standard results? | No | High | Retrospective analysis of patients after biliary atresia status was known |
| Threshold pre-specified? | No | High | Different thresholds considered |
| Domain III: Reference standard | | | |

| | | | |
|---|---|---|---|
| Reference standard likely to correctly classify condition? | Unclear | Unclear | No details provided |
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Retrospective analysis of patients after biliary atresia status was known |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | Unclear | Unclear | No details provided |
| All patients included in analysis? | Yes | Low | |
| **Domain V: Applicability** | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher incidence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Unclear | Unclear | No details provided |
| Other comments | This retrospective study considered several potential biomarkers for biliary atresia. Dried blood spots were collected from infants born at the study hospital. The analysis was based on samples from infants who were known to have biliary atresia and healthy controls.<br>The authors concluded that metabolic markers in dried blood spots obtained from newborns have a great potential for biliary atresia screening.<br>The report included infants screened over a 7-year period up to 2020 in a county with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear.<br>This study also reported test performance for additional combinations of metabolites. These data were not extracted as they were not highlighted by the authors as the best performing combinations. | | |

# Stool colour cards

## Systematic reviews

### Table 24: Arshad et al. 2023 [32]

| | |
|---|---|
| Publication | Arshad A, Gardiner J, Ho C, Rees P, Chadda K, Baker A, et al. Population-based screening methods in biliary atresia: a systematic review and meta-analysis. Archives of Disease in Childhood. 2023;108(6):468-73. |
| Study details | Systematic review and meta-analysis |
| Study objectives | To investigate tested methods of population-based biliary atresia screening |
| Study setting | The systematic review was conducted in the UK |
| Inclusions | Observational studies, published between January 1975 and 12 September 2022, reporting outcomes of a biliary atresia screening method<br>Included studies were prospective cohort studies, cross-sectional studies, retrospective cohort studies and case control-studies |
| Exclusions | Studies that were opinions, reviews or non-peer-reviewed letters. Studies not published in English |
| Population | Newborns with biliary atresia<br>The systematic review considered various aspects of screening for biliary atresia. Only data that meet the inclusion criteria for this evidence summary have been extracted |
| Test | Stool colour cards (5 studies published between 2006 and 2016) |
| Comparator/reference standard | Not stated |
| Outcomes | Meta-analysis reported for sensitivity and specificity |
| Sensitivity | 87.9% (95%CI 80.4 to 92.8) |
| Specificity | 99.9% (95%CI 99.9 to 99.9) |
| PPV | 5.6% (CI not reported) |
| NPV | 100% (CI not reported) |

See Table 20 for the critical appraisal of Arshad et al. 2023

| | |
|---|---|
| Other comments | The 5 studies reporting stool colour cards were from Canada, China, Japan, and Taiwan.<br>Three studies published since 2012 (Chiu et al. 2013 [30], Gu et al. 2015 [15] and Kong et al. 2016 [31]) are also separately included in the Tables below.<br>The authors also calculated PPV and NPV for differing prevalence of biliary atresia. The results for a prevalence of 1 in 15,000 were extracted as this is similar to the prevalence of biliary atresia in the UK. For information, the PPV and NPV for a prevalence of 1 in 10,000 was 8.1% and 100% respectively.<br>The assessment of the diagnostic accuracy of dried blood spots also reported in this systematic review is presented in Table 20. |

## Table 25: Gopal et al. 2024 [35]

| | |
|---|---|
| Publication | Gopal SH, Zebda R, Mohan A, Borovsky K, Takwoingi Y, Scandrett K, et al. Population-based screening strategies for biliary atresia in the newborn: A systematic review and meta-analysis. PLoS ONE 2024;19(8):e0307837. |
| Study details | Systematic review and meta-analysis |
| Study objectives | To systematically review the accuracy of population-based screening strategies for biliary atresia in newborns |
| Study setting | The systematic review was conducted in Canada |
| Inclusions | Prospective or retrospective cohort or cross-sectional studies, including conference abstracts<br>Studies published up to February 2023, reporting outcomes of a biliary atresia screening test |
| Exclusions | Studies with a case-control design and studies not in the clinical context of population screening for biliary atresia<br>Reviews, research letters and ongoing clinical trials |
| Population | Newborns with biliary atresia<br>The systematic review considered various tests for screening for biliary atresia. Only data that meet the inclusion criteria for this evidence summary have been extracted |
| Test | Stool colour cards (7 studies published between 2006 and 2020) |
| Comparator/reference standard | When stated, this was biliary atresia diagnosed by intraoperative cholangiogram and/or tissue pathology |
| Outcomes | Meta-analysis reported for sensitivity and specificity |
| Sensitivity | 79.6% (95%CI 70.6 to 86.4) |
| Specificity | 99.9% (95%CI 99.9 to 99.9) |
| False positives | The authors calculated that for a hypothetical cohort of 100,000 newborn infants with a biliary atresia prevalence of 1 in 15,000, one patient with biliary atresia would be missed and there would be 1,000 false positives |

Quality appraisal using the JBI checklist for systematic reviews

| Question | Assessment | Supporting information |
|---|---|---|
| Is the review question clearly and explicitly stated? | Yes | |
| Were the inclusion criteria appropriate for the review question? | Yes | |
| Was the search strategy appropriate? | Yes | |
| Were the sources and resources used to search for studies adequate? | Yes | |
| Were the criteria for appraising studies appropriate? | Yes | Two reviewers independently screened titles and abstracts and a further 3 reviewers also independently checked selected full texts against the inclusion criteria |
| Was critical appraisal conducted by 2 or more reviewers independently? | Yes | Three reviewers independently conducted critical appraisal |

| | | |
|---|---|---|
| Were there methods to minimize errors in data extraction? | Yes | Three reviewers independently extracted data |
| Were the methods used to combine studies appropriate? | Unclear | No formal assessment of heterogeneity |
| Was the likelihood of publication bias assessed? | Yes | |
| Were recommendations for policy and/or practice supported by the reported data? | Yes | |
| Were the specific directives for new research appropriate? | Yes | |

Other comments

The 7 studies reporting stool colour cards were from Canada, China, Japan, and Taiwan.

Risk of bias for the included studies was assessed using QUADAS-2 and the certainty of the evidence was assessed using GRADE. The included studies mostly had a low risk of bias. The authors stated that many studies excluded infants born prematurely and/or with another serious medical condition and that most studies had a predefined threshold for a positive screen for biliary atresia. The authors stated that the evidence for screening using stool colour cards was of moderate uncertainty following downgrading for imprecision and inconsistency.

Case-control studies were not in scope for the systematic review.

No formal assessment of heterogeneity was conducted. The authors recognised that heterogeneity existed in the included studies due to the timing of screening using the different tests and differences in the populations screened.

The calculations of number of false positives and false negatives used a prevalence value that is similar to the prevalence of biliary atresia in the UK.

Four studies published since 2012 (Gu et al. 2015 [15], Gu et al. 2020 [16], Kong et al. 2016 [31] and Woolfson et al. 2018 [22]) are also separately included in this evidence summary

The systematic review also included one study testing blood carnitine levels in dried blood spots (Gong et al 2020 [29]). This study is separately included in this evidence summary as no pooled data for this test were available from the systematic review

## Individual studies

### Table 26: Chiu et al. 2013 [30]

| | | | | |
|---|---|---|---|---|
| Publication | Chiu CY, Chen PH, Chan CF, Chang MH, Wu TC. Taiwan infant stool color card study G. Biliary atresia in preterm infants in Taiwan: a nationwide survey. Journal of Pediatrics. 2013;163(1):100-3.e1. | | | |
| Study details | Screening study | | | |
| Study objectives | To compare the outcomes of biliary atresia in term and pre-term infants | | | |
| Study setting | Taiwan | | | |
| Inclusions | Infants born between January 2004 and June 2010, after a national universal stool colour screening programme was implemented | | | |
| Exclusions | None stated | | | |
| Population | 197 newborns diagnosed with biliary atresia<br>• 27 infants were pre-term (gestational age < 37 weeks)<br>• 170 infants were term | | | |
| Test | Stool colour cards placed within the Children's Health Handbook that is provided for every newborn<br>Care givers were asked to notify the stool card registry centre within 24 hours when abnormal stool colour was suspected. Paediatricians also routinely evaluated stool colour during hospital visits and were asked to report back to the centre if abnormal findings were observed | | | |
| Comparator/reference standard | Intraoperative cholangiography during Kasai portoenterostomy or histologic and operative findings during liver transplantation | | | |
| Outcomes | | | | |
| | | Term infants | Pre-term infants | p value |
| Sensitivity to detect biliary atresia before 60 days of age (CI not reported) | | 92.8% | 96.3% | p=0.798 |
| Quality appraisal using the QUADAS-2 tool | | | | |
| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information | |
| Domain I: Patient selection | | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Infants born over a 6-year period | |
| Case-control design avoided? | Yes | Low | | |
| Inappropriate exclusions avoided? | Yes | Low | No exclusions stated | |
| Domain II: Index Test | | | | |
| Index test results interpreted without knowledge of reference standard results? | Yes | Low | | |
| Threshold prespecified? | Yes | Low | Abnormal stool colour against the stool colour card | |
| Domain III: Reference standard | | | | |

| | | | |
|---|---|---|---|
| Reference standard likely to correctly classify condition? | Yes | Low | |
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Infants diagnosed with biliary atresia received the reference standard. The process for the assessment of any infants who screened positive but did not have biliary atresia, if any, was not stated |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | Unclear | Unclear | The process for the assessment of any infants who screened positive but did not have biliary atresia, if any, was not stated |
| All patients included in analysis? | Yes | Low | All infants with biliary atresia born during the study period |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher incidence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Yes | Low | |
| Other comments | Infants were identified after the implementation of a national universal screening programme.<br><br>The focus of this study was to compare outcomes for pre-term and term infants who had been diagnosed with biliary atresia. Although the sensitivity of detecting biliary atresia before 60 days of age was reported this was not a study that sought to fully assess the diagnostic accuracy of stool colour cards. No information was provided about whether any infants screened positive on the stool colour card but were not ultimately diagnosed with biliary atresia.<br><br>The study also compared outcomes for pre-term and term infants. These data have not been extracted as this comparison is not within the scope of this evidence summary.<br><br>The authors concluded that the stool colour card is sensitive to detect biliary atresia in pre-term infants.<br><br>The report included infants screened over a 6-year period up to 2010 in a county with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear. | | |

## Table 27: Gu et al. 2020 [16] and Kong et al. 2016 [31]

| | | |
|---|---|---|
| Publication | Gu YH, Zhao JQ, Kong YY, Yang HH, Diao M, Li L, et al. Repeatability and reliability of home-based stool color card screening for biliary atresia based on results in China and Japan. Tohoku Journal of Experimental Medicine. 2020;252(4):365-72.<br>Kong YY, Zhao JQ, Wang J, Qiu L, Yang HH, Diao M, et al. Modified stool color card with digital images was efficient and feasible for early detection of biliary atresia-a pilot study in Beijing, China. World Journal of Pediatrics. 2016;12(4):415-20. (For additional results on the patients in China reported in Gu et al. 2020) | |
| Study details | Screening study | |
| Study objectives | To assess the performance of a stool colour card universal screening programme | |
| Study setting | Childbirth facilities in one district (Sapporo) in Japan and one district (Beijing) in China | |
| Inclusions | Infants born in Sapporo between April 2012 and July 2015<br>Infants born in Beijing between December 2013 and October 2014 | |
| Exclusions | Infants in Sapporo who had received and used a previous edition of the stool colour card were excluded<br>Infants in Beijing who were born between August and October 2014 were excluded because they could not be follow-up up until 4 months after birth within the study timeframe | |
| Population | 37,478 of 48,770 newborns in Sapporo (94.3%)<br>27,561 of 29,799 newborns in Beijing (92.5%) | |
| Test | Stool colour cards distributed to guardians via a Maternal and Child Health Handbook (Sapporo) or by trained nurses in maternal facilities during information sessions on neonatal screening for phenylketonuria and congenital hypothyroidism (Beijing)<br>Stool colour card results were collected at the one-month health check-up (Sapporo) or at 2 weeks, one month and 1-4 months after birth (Beijing) | |
| Comparator/reference standard | Diagnosis of biliary atresia on the database of the Medical Aid Program for Chronic Paediatric Diseases of Specified Categories of Japan (Sapporo) or a database of congenital abnormalities (Beijing) | |
| Outcomes | 3 cases of biliary atresia were diagnosed in Sapporo. 4 cases of biliary atresia were diagnosed in Beijing<br>1 patient in Sapporo and 2 patients in Beijing were diagnosed by clinical findings other than stool colour card screening | |
| | Test performance at one month after birth (CI not reported): | |
| | Sapporo | Beijing |
| Sensitivity | 0% (0 confirmed cases at one month) | 100% |
| Specificity | 99.9% | 99.9% |
| False positive rate | 0.03% | 0.01% |
| False negative rate | 0% | 0% |
| | Test performance at one to 4 months after birth (for Beijing only): | |
| Sensitivity | 100% (CI not reported) | |
| Specificity | 99.9% (95CI 99.9 to 99.9) | |
| False positive rate | 0.08% (CI not reported) | |
| False negative rate | 0% (CI not reported) | |
| PPV | 8.3% (95% CI 2.7 to 19.4) | |
| Incidental findings | In Sapporo there were single patients diagnosed with neonatal intrahepatic cholestasis caused by citrin deficiency, neonatal hepatitis, suspected neonatal hepatitis and transient cholestasis | |

| | | | |
|---|---|---|---|
| | In Beijing there were single patients diagnosed with Alagille syndrome, cytomegalovirus hepatitis and Duchenne muscular dystrophy and one patient was receiving parenteral nutrition | | |
| Quality appraisal using the QUADAS-2 tool | | | |
| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Infants born in 2 districts for whom a stool colour card was returned |
| Case-control design avoided? | Yes | Low | |
| Inappropriate exclusions avoided? | Yes | Low | No inappropriate exclusions |
| Domain II: Index Test | | | |
| Index test results interpreted without knowledge of reference standard results? | Yes | Low | |
| Threshold pre-specified? | Yes | Low | A positive result was a stool colour that matched either image 1, 2 or 3 |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Unclear | Unclear | Limited details provided |
| Reference standard results interpreted without knowledge of index test results? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | Unclear | Unclear | Limited details provided |
| All patients included in analysis? | Yes | Low | |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher incidence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Unclear | Unclear | Limited details provided |

| | |
|---|---|
| Other comments | The paper by Gong et al. 2020 was used as the main source of information. Additional results (confidence intervals and PPV) for infants in Beijing were extracted from the paper by Kong et al. 2016.<br><br>2 of the biliary atresia cases diagnosed in Sapporo and one of the cases in Beijing were detected by guardians using the stool colour card. However, these infants were not positive on the stool colour card at one month old.<br><br>Limited details were reported about the reference standard. Biliary atresia cases were identified from regional or national databases. Limited details were provided about processes for diagnosing biliary atresia. The risk of bias from potential missed cases is low given the nature of the condition and the data sources used.<br><br>The PPV reported by Kong et al. 2016 was based on a biliary atresia incidence of 1.3 in 10,000.<br><br>The authors concluded that the repeatability and reliability of the home-based stool colour card has been demonstrated in both Japan and China and that their results demonstrate that observation of stool colour is useful for early detection of biliary atresia, especially for biliary atresia patients without obvious clinical manifestations. They recommended that the observation period should be up to 4 months after birth.<br><br>The report included infants screened over a 2- or 3-year period up to 2015 in 2 countries with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear. |

## Table 28: Gu et al. 2015 [15]

| | |
|---|---|
| Publication | Gu YH, Yokoyama K, Mizuta K, Tsuchioka T, Kudo T, Sasaki H, et al. Stool color card screening for early detection of biliary atresia and long-term native liver survival: a 19-year cohort study in Japan. Journal of Pediatrics. 2015;166(4):897-902.e1. |
| Study details | Screening study |
| Study objectives | To assess the performance of a stool colour card screening programme |
| Study setting | One district (Tochigi Prefecture) in Japan |
| Inclusions | Infants born to mothers living in the district between August 1994 and March 2011 |
| Exclusions | Infants born to mothers who lived outside of the prefecture before giving birth |
| Population | 264,071 of the 313,230 infants born during the study period (84.3%) |
| Test | Stool colour cards distributed to all pregnant women via a Maternal and Child Health Handbook<br><br>Stool colour cards were collected at the one-month health check-up |
| Comparator/reference standard | Laparotomy or operative cholangiography prior to Kasai portoenterostomy for high-risk cases by a paediatric hepatologist or surgeon. The database of the Medical Aid Program for Chronic Paediatric Diseases of Specified Categories of Japan was also checked |
| Outcomes | 2,014 infants screened positive and 26 were diagnosed with biliary atresia. A further 8 cases of biliary atresia were diagnosed but were missed at the one-month health check-up (34 biliary atresia cases in total) |
| Sensitivity | 76.5% (95%CI 62.2 to 90.7) |

| | |
|---|---|
| Specificity | 99.9% (95%CI 99.9 to 100) |
| PPV | 12.7% (95%CI 8.2 to 17.3) |
| NPV | 99.9% (95%CI 99.9 to 99.9) |
| Incidental findings | One patient was diagnosed with Alagille syndrome |
| Missed cases | 8 cases of biliary atresia were not detected by the screening programme. Of these:<br>• 2 patients were in a neonatal intensive care unit for more than one month after birth and their abnormal stool colour was overlooked<br>• 3 patients had their pale-pigmented stool colour reported by their guardians at the one-month check-up but no further examination was conducted by their paediatricians because they did not present with visible jaundice<br>• for one patient their guardians did not use the stool colour card<br>• one patient did not have abnormal stool colour at the one-month check-up but the patient's guardian reported pale-stool and jaundice at 1.5 months of age<br>• one patient was not on the study's patient list and was identified via the medical aid list |

Quality appraisal using the QUADAS-2 tool

| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
|---|---|---|---|
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Infants born over a 17-year period |
| Case-control design avoided? | Yes | Low | |
| Inappropriate exclusions avoided? | Yes | Low | |
| Domain II: Index Test | | | |
| Index test results interpreted without knowledge of reference standard results? | Yes | Low | |
| Threshold pre-specified? | Yes | Low | A positive result was a stool colour that matched either image 1, 2 or 3 |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Yes | Low | |
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Only infants who screened positive received the reference standard |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | No | Unclear | Only infants who screened positive received the reference standard |

| | | | |
|---|---|---|---|
| All patients included in analysis? | Yes | Low | |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | No | High | Population with higher incidence of biliary atresia than found in UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Yes | Low | |
| Other comments | Only infants who tested positive received the reference standard. However, biliary atresia is one of 514 diseases that are supported by a medical aid programme in Japan. The authors checked all patients included in the study against the medical aid programme to ensure that no patients with biliary atresia in the district were overlooked. This reduced the possibility of missing cases and given the nature of the condition, there is unlikely to have been any serious risk of bias. It is not clear if cases for patients who had moved out of the area could have been missed.<br><br>The authors reported a biliary atresia incidence of 1.1 in 10,000 for their population.<br><br>The authors concluded that the sensitivity and specificity of the stool colour card have been demonstrated by this 19-year cohort study.<br><br>The report included infants screened over a 19-year period up to 2011 in a country with a higher incidence of biliary atresia than the UK. The applicability of the outcomes reported to current UK practice is not clear.<br><br>This paper also reported outcomes following the implementation of the screening programme. These outcomes are reported under Question 3 below. | | |

## Table 29: Woolfson et al. 2018 [22]

| | |
|---|---|
| Publication | Woolfson JP, Schreiber RA, Butler AE, MacFarlane J, Kaczorowski J, Masucci L, et al. Province-wide biliary atresia home screening program in British Columbia: Evaluation of first 2 years. Journal of Pediatric Gastroenterology & Nutrition. 2018;66(6):845-9. |
| Study details | Screening study |
| Study objectives | To assess the performance of a province-wide stool colour card screening programme |
| Study setting | 126 maternity units in British Columbia, Canada |
| Inclusions | Infants born between April 2014 and March 2016 |
| Exclusions | Babies admitted to neonatal intensive care units were excluded |
| Population | 87,583 newborns |
| Test | Stool colour cards given to families at maternity discharge. Families were instructed to monitor their infant's stool colour for one month and contact the screening centre with concerns. An online tool to remind parents to check their newborn's stool daily for one month was also made available to families through sign up online or screening a QR code on the stool colour card |

| | | |
|---|---|---|
| | From June 2015, a statement was added to the stool colour card instructing physicians to order a bilirubin tests for infants jaundiced at more than 2 weeks of age | |
| Comparator/reference standard | Diagnosis of biliary atresia following referral to the British Columbia Children's Hospital | |
| Outcomes | 6 cases of biliary atresia were identified. 5 of these 6 cases were considered screen successes (acholic stools identified by families). 3 of these 5 screen successes were considered programme successes (acholic stools identified by the family and case referral by the physician in a timely manner leading to surgery before 90 days of age) | |
| | Test performance of the screening programme (CI not reported) | Test performance to detect acholic stools (CI not reported) |
| Sensitivity | 50% | 83% |
| Specificity | 99% | 99% |
| PPV | 4% | 6% |
| NPV | 99% | 99% |
| False positive rate | 0.09% | Not reported |
| False positives | The biliary atresia screening centre received 75 phone calls from parents during the study period. One of these calls had a confirmed diagnosis of biliary atresia<br><br>Of the 74 calls with a false-positive stool colour card reading:<br>• 9 calls were for cases more than 6 months of age<br>• 25 calls had normal stool colour or other non-related stool issues<br>• 29 had transient acholic stool that had resolved by the phone call follow-up without requiring further investigation<br>• 2 had identified alternative diagnoses<br>• 9 cases had no cause identified | |
| Incidental findings | There were single patients diagnosed with Alagille syndrome, a urinary tract infection and thyroid disease | |

Quality appraisal using the QUADAS-2 tool

| Question | Assessment (yes, no, unclear) | Risk of Bias (low, high, unclear) | Supporting information |
|---|---|---|---|
| Domain I: Patient selection | | | |
| Consecutive or random sample of population enrolled? | Yes | Low | Infants born over a 2-year period in one area |
| Case-control design avoided? | Yes | Low | |
| Inappropriate exclusions avoided? | Yes | Low | No inappropriate exclusions |
| Domain II: Index Test | | | |
| Index test results interpreted without knowledge of reference standard results? | Yes | Low | |
| Threshold pre-specified? | Yes | Low | Abnormal stool colour against the stool colour card |
| Domain III: Reference standard | | | |
| Reference standard likely to correctly classify condition? | Yes | Low | |

| | | | |
|---|---|---|---|
| Reference standard results interpreted without knowledge of index test results? | No | Unclear | Only patients who screened positive received further assessment |
| Domain IV: Test strategy flow and timing | | | |
| Appropriate interval between index test and reference standard? | Unclear | Unclear | Interval between index test and reference standard not stated |
| Did all participants receive same reference standard? | No | Unclear | Only patients who screened positive received further assessment |
| All patients included in analysis? | Yes | Low | |
| Domain V: Applicability | | | |
| Applicable to UK screening population of interest? | Yes | Low | Population with similar incidence of biliary atresia to the UK |
| Applicable to UK screening test of interest? | Yes | Low | |
| Target condition measured by reference test applicable to UK screening condition of interest? | Yes | Low | |
| Other comments | Infants were identified after the implementation of a regional screening programme.<br><br>Only infants who screened positive received the reference standard. Given the nature of the condition this is unlikely to have introduced serious risk of bias as it is unlikely that many cases of biliary atresia would have been missed within the timeframe of the pilot.<br><br>The authors reported a biliary atresia incidence of 1 in 14,597 for their population.<br><br>The authors concluded that the screening programme has high specificity and distribution with low cost, and successful programme case identification led to earlier age at surgery.<br><br>The report included infants screened over a 2-year period up to 2016 in a county with a similar incidence of biliary atresia to the UK. However, the applicability of the outcomes reported to current UK practice is not clear.<br><br>The study also reported age at surgery for cases detected after the implementation of screening. These data have not been extracted as studies reporting this outcome for screening vs no screening were identified and prioritised.<br><br>The study also reported misclassification. This is discussed under Question 3 below. | | |

# Question 2: What is the reported age at surgery/time to surgery for biliary atresia (Kasai portoenterostomy) in the UK?

## Sub-questions – Is there any data to suggest that age at surgery has changed since 2016?

## What proportion of affected infants have a late (>60 and >90 days) Kasai portoenterostomy in the UK?

### Table 30: Davenport et al. 2025 [39]

| | | | | | |
|---|---|---|---|---|---|
| Publication | Davenport M, Makin E, Ong EG, Sharif K, Dawrant M, Alizai N. The outcome of a centralization program in biliary atresia: Twenty years and beyond. Ann Surg. 2025;281(4):608-14. | | | | |
| Study details | Case series | | | | |
| Study objectives | To report outcomes and trends 20 years after the start of a national centralisation programme for the treatment of biliary atresia involving 3 English units | | | | |
| Study setting | Prospective registry and database of all infants with biliary atresia managed at one of the 3 national paediatric liver units in England | | | | |
| Inclusions | Infants with biliary atresia who had surgery between January 1999 and December 2019 | | | | |
| Exclusions | None stated | | | | |
| Population | 831 infants with biliary atresia treated with Kasai portoenterostomy at one of 3 national paediatric liver units (Birmingham, London, Leeds) | | | | |
| Intervention | Kasai portoenterostomy | | | | |
| Comparator | None | | | | |
| Outcomes | | | | | |
| Median (interquartile range (IQR) age at Kasai portoenterostomy (days) | 51 (39 to 64)** | | | | |
| Mean (standard deviation (SD)) age at Kasai portoenterostomy (days) | 53.8 (23.08)†† | | | | |
| Proportion of infants presenting at >100 days old | 49/831 (5.6%) | | | | |
| | 1999 to 2003 | 2004 to 2008 | 2009 to 2013 | 2014 to 2018 | p-value |
| Median (IQR) age at Kasai portoenterostomy (days) by time period | 53 (42 to 68) | 55 (44 to 70) | 48.5 (37 to 62) | 48 (35 to 57) | p=0.0001 |
| Proportion of infants presenting at >100 days old by time period | 14/192 (7.3%) | 14/204 (6.8%) | 10/234 (4.2%) | 8/191 (4.1%) | p=0.1 |
| Quality appraisal using the JBI checklist for case series | | | | | |

| Question | Assessment | Supporting information |
|---|---|---|
| Were there clear criteria for inclusion in the case series? | Yes | All infants referred to a UK national paediatric liver unit |

** At one point in the paper this IQR was given as 39 to 67
†† At one point in the paper this mean (SD) was given as 54.6 (± 0.9)

| | | |
|---|---|---|
| Was the condition measured in a standard, reliable way for all participants included in the case series? | Yes | |
| Were valid methods used for the identification of the condition for all participants included in the case series? | Yes | |
| Did the case series have consecutive inclusion of participants? | Yes | |
| Did the case series have complete inclusion of participants? | Yes | |
| Was there clear reporting of the demographics of the participants in the study? | No | Not reported |
| Was there clear reporting of clinical information of the participants? | Yes | |
| Were the outcomes or follow up results of cases clearly reported? | N/A | |
| Was there clear reporting of the presenting site(s)/clinic(s) demographic information? | Yes | |
| Was statistical analysis appropriate? | N/A | |
| Other comments | The data reported was for all infants with biliary atresia managed at the 3 national paediatric liver units in England. Only infants born within England or Wales were included.<br><br>No information was provided about how infants were detected or the processes for referral to the centres. There was no demographic information about the infants.<br><br>The report included infants with biliary atresia treated over a 20-year period up to 2019. The applicability of the outcomes reported to current UK practice is not clear. | |

## Table 31: Durkin et al. 2017 [37]

| | |
|---|---|
| Publication | Durkin N, Deheragoda M, Davenport M. Prematurity and biliary atresia: a 30-year observational study. Pediatric Surgery International. 2017;33(12):1355-61. |
| Study details | Case series |
| Study objectives | To assess whether premature infants with biliary atresia have a delayed time to surgery and worse outcomes |
| Study setting | Retrospective review of a prospectively maintained database at a single centre in the UK (London) |
| Inclusions | Infants diagnosed with biliary atresia between January 1988 and December 2016 |
| Exclusions | None stated |
| Population | 692 infants were diagnosed with biliary atresia during the study period. Relevant outcomes were reported for 21 premature infants (delivery <37/40 weeks gestation) and 42 term infants. The term infants included were contemporaneous to the premature infants to negate the effect of changes in practice over time |
| Intervention | Kasai portoenterostomy<br><br>This was the primary procedure for 20 of the premature infants and all 42 of the term infants |
| Comparator | None (outcomes for term and premature infants were compared. This comparison is not within the scope of this evidence summary) |

| Outcomes | Premature infants | Term infants |
|---|---|---|
| Median (range) age at diagnostic biopsy (days) | 57 (4 to 266) | 47 (24 to 129) |
| Median (range) age at Kasai portoenterostomy (days) | 65 (16 to 323) | 56 (27 to 141) |
| Proportion of infants operated at >100 days old | 5/20 (25%) | Not reported |
| Deferral time (time from diagnosis with liver biopsy to operation) (days) | 11 (range not reported) | 7 (range not reported) |

Quality appraisal using the JBI checklist for case series (this checklist was used as the comparison reported in the study was not within the scope of this evidence summary)

| Question | Assessment | Supporting information |
|---|---|---|
| Were there clear criteria for inclusion in the case series? | Yes | Infants diagnosed with biliary atresia at one UK centre |
| Was the condition measured in a standard, reliable way for all participants included in the case series? | Yes | |
| Were valid methods used for the identification of the condition for all participants included in the case series? | Yes | |
| Did the case series have consecutive inclusion of participants? | Unclear | Data taken from a database |
| Did the case series have complete inclusion of participants? | Unclear | Data taken from a database |
| Was there clear reporting of the demographics of the participants in the study? | Yes | |
| Was there clear reporting of clinical information of the participants? | No | Not reported |
| Were the outcomes or follow up results of cases clearly reported? | N/A | |
| Was there clear reporting of the presenting site(s)/clinic(s) demographic information? | No | Not reported |
| Was statistical analysis appropriate? | N/A | |

| | |
|---|---|
| Other comments | The data were taken from a prospectively maintained database but it is not clear if the inclusion of participants was consecutive and complete. Patients were treated at a single UK centre, but no further information was provided. The proportion of patients who received surgery at an older age was reported for premature infants but not for term infants.<br><br>No information was provided about how infants were detected or the processes for referral to the centre.<br><br>The main aim of this study was to assess whether premature infants with biliary atresia have worse outcomes than term infants. This comparison is not within the scope of this evidence summary.<br><br>The study included a small number of patients from a single centre treated over a 28 year period up to 2016. The applicability of the outcomes reported to current UK practice is not clear. |

## Table 32: Williams et al. 2018 [38]

| | |
|---|---|
| Publication | Williams R, Alexander G, Armstrong I, Baker A, Bhala N et al. Disease burden and costs from excess alcohol consumption, obesity, and viral hepatitis: fourth report of the Lancet Standing Commission on Liver Disease in the UK. Lancet. 2018, 391(10125):1097-1107. |
| Study details | Case series |
| Study objectives | To provide new and follow-up metric data relating to the 8 main recommendations of the Lancet Standing Commission on Liver Disease in the UK |
| Study setting | Under recommendation 4: specialist paediatric services and continuity of care in transition arrangements for children with liver disease reaching adult life, the paper reports number of referrals of infants with persistent conjugated jaundice to the 3 national paediatric liver units, number subsequently diagnosed with extrahepatic biliary atresia and age at referral |
| Inclusions | Infants (aged <6 months) born with persistent conjugated jaundice between 2012 and 2017 |
| Exclusions | None stated |
| Population | 2,117 infants with persistent conjugated jaundice referred to 3 national paediatric liver units (Birmingham, London, Leeds). 258 were diagnosed with extrahepatic biliary atresia |
| Intervention | Not reported |
| Comparator | None |
| Outcomes | |
| Median (range) age at referral (days) | 45 (0 to 242) |
| Referral after 56 days of age | 56/258 (22%) |

Quality appraisal using the JBI checklist for case series

| Question | Assessment | Supporting information |
|---|---|---|
| Were there clear criteria for inclusion in the case series? | Yes | All infants referred to a UK national paediatric liver unit |
| Was the condition measured in a standard, reliable way for all participants included in the case series? | Yes | |
| Were valid methods used for the identification of the condition for all participants included in the case series? | Yes | |
| Did the case series have consecutive inclusion of participants? | Yes | |
| Did the case series have complete inclusion of participants? | Yes | |
| Was there clear reporting of the demographics of the participants in the study? | No | Not reported |
| Was there clear reporting of clinical information of the participants? | No | Not reported |
| Were the outcomes or follow up results of cases clearly reported? | N/A | |
| Was there clear reporting of the presenting site(s)/clinic(s) demographic information? | Yes | |
| Was statistical analysis appropriate? | N/A | |

| | |
|---|---|
| Other comments | The data reported was for all infants with persistent conjugated jaundice referred to one of the 3 UK national paediatric liver units.<br><br>No information was provided about how infants were detected or the processes for referral to the centre. There was no demographic or clinical information about the infants.<br><br>The report provided age at referral but did not report age at surgery or time to surgery.<br><br>The report included infants with biliary atresia referred between 2012 and 2017. The applicability of the outcomes reported to current UK practice is not clear. |

# Question 3: Does screening for biliary atresia using stool colour cards improve time to surgery and clinical outcomes?

## Sub-questions – Does screening for biliary atresia using stool colour cards lead to harms?

### Systematic reviews

The systematic review by Arshad et al. 2023 [32] included in Question 1 also considered age at and outcomes of Kasai portoenterostomy following screening using stool colour cards. The included studies were described narratively with no meta-analysis. Therefore, these studies were separately considered for inclusion in this evidence summary and the systematic review was not included for this question.

### Individual studies

#### Table 33: Gu et al. 2015 [15]

| | | | | |
|---|---|---|---|---|
| Publication | Gu YH, Yokoyama K, Mizuta K, Tsuchioka T, Kudo T, Sasaki H, et al. Stool color card screening for early detection of biliary atresia and long-term native liver survival: a 19-year cohort study in Japan. Journal of Pediatrics. 2015;166(4):897-902.e1. | | | |
| Study details | Screening study | | | |
| Study objectives | To assess the performance of a stool colour card screening programme | | | |
| Study setting | One district (Tochigi Prefecture) in Japan | | | |
| Inclusions | Infants born to mothers living in the district between August 1994 and March 2011<br>Data were taken from national and regional databases | | | |
| Exclusions | Infants born to mothers who lived outside of the prefecture before giving birth | | | |
| Population | 34 infants with biliary atresia diagnosed after the implementation of the screening programme in 1994. The authors reported that 26 of these infants were identified by the screening programme<br>Infants with biliary atresia diagnosed between 1987 and 1992 in the same region before the screening programme (n not stated) | | | |
| Intervention | Stool colour cards distributed to all pregnant women via a Maternal and Child Health Handbook<br>Stool colour cards were collected at the one-month health check-up | | | |
| Comparator | No screening | | | |
| Outcomes | All infants with biliary atresia received a Kasai portoenterostomy | | | |
| | | Before screening | After screening | p value |
| Mean (SD) age at Kasai portoenterostomy (days) | | 70.3 (SD not reported) | 59.7 (± 19.4) | p=0.003 |
| Median (range) age at Kasai portoenterostomy (days) | | 65.5 (range not reported) | 58.5 (18 to 109) | Not reported |
| Proportion of infants who had Kasai portoenterostomy at ≤ 60 days old | | 34.0% (n not reported) | 19/34 (55.9%, 95%CI 39.2 to 72.6) | Not reported |
| Proportion of infants who had Kasai portoenterostomy at > 90 days old | | 13.0% (n not reported) | 2/34 (5.9%, 95%CI 2.0 to 13.8) | p<0.05 |

| | |
|---|---|
| 5-year native liver survival rate | 87.6% (standard error (SE) 0.06) |
| 10-year native liver survival rate | 76.9% (SE 0.08) |
| 15-year native liver survival rate | 48.5% (SE 0.11) |
| Misclassification | The authors stated that no infants with a false positive screening result underwent any invasive procedures |

| Quality appraisal using the JBI checklist for cohort studies | | |
|---|---|---|
| Question | Assessment | Supporting information |
| Were the 2 groups similar and recruited from the same population? | Yes | Infants born in one district in Japan |
| Were the exposures measured similarly to assign people to both exposed and unexposed groups? | Yes | Before and after screening |
| Was the exposure measured in a valid and reliable way? | Yes | |
| Were confounding factors identified? | Yes | Considered in the discussion |
| Were strategies to deal with confounding factors stated? | No | |
| Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | |
| Were the outcomes measured in a valid and reliable way? | Yes | Objective outcomes assessed |
| Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | |
| Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | Yes | |
| Were strategies to address incomplete follow up utilized? | N/A | |
| Was appropriate statistical analysis used? | Yes | |

**Other comments**

Data were reported for 34 infants diagnosed with biliary atresia after screening was implemented. However, not all the infants were detected by the screening programme. Data for the period before screening was implemented were taken from a national database. The number of infants with biliary atresia during this time period was not reported.

The authors reported historical liver survival rates from published studies for other districts in Japan (1970 to 1986), the UK (1990 to 2009), France (1986 to 2009) and the US (1972 to 1996). These data are not extracted as they may not reflect current rates in these countries and data comparing this outcome for this population compared to an area with no screening was available from Gu & Matsui 2017 [40].

Potential confounding factors were discussed, with the authors suggesting that the management of patients with biliary atresia had not changed drastically over the years. No strategies to deal with confounding factors were reported.

The authors concluded that the timing of Kasai procedure and long-term native liver survival probabilities were improved, suggesting a beneficial effect of stool colour card screening.

The report included infants with biliary atresia treated over a 19-year period up to 2011 in a county with a higher incidence of biliary atresia and a different healthcare system to the UK. The applicability of the outcomes reported to current UK practice is not clear.

This paper also reported diagnostic accuracy for screening with stool colour cards. These outcomes are reported above under Question 1.

## Table 34: Gu & Matsui 2017 [40]

| | | | | |
|---|---|---|---|---|
| Publication | Gu YH, Matsui A. Long-term native liver survival in infants with biliary atresia and use of a stool color card: Case-control study. Pediatrics International. 2017;59(11):1189-93. | | | |
| Study details | Retrospective cohort study | | | |
| Study objectives | To investigate whether long-term native liver survival in infants with biliary atresia is associated with use of a stool colour card | | | |
| Study setting | One district (Tochigi Prefecture) in Japan | | | |
| Inclusions | Infants born between 1994 and March 2011 who had undergone a Kasai portoenterostomy | | | |
| Exclusions | None stated | | | |
| Population | 34 infants with biliary atresia born after screening was implemented in 1994<br>114 infants with biliary atresia born in areas where screening had not been implemented | | | |
| Intervention | Stool colour card screening for biliary atresia<br>Stool colour cards were collected at the one-month health check-up | | | |
| Comparator | No screening | | | |
| Outcomes | Outcomes reported for all infants with biliary atresia who received a Kasai portoenterostomy | | | |
| | | Screening area | No screening | p value |
| Mean (SD) age at first open Kasai portoenterostomy (days) | | 59.7 (± 19.4) | 68.1 (± 25.6) | p=0.003 |
| Proportion of infants who had Kasai portoenterostomy at ≤ 60 days old | | 19/34 (55.9%) | 46/114 (40.4%) | p=0.109 |
| Proportion of infants who had Kasai portoenterostomy at > 90 days old | | 2/34 (5.9%) | 23/114 (20.2%) | p=0.067 |
| Native liver survival at 5 years (n not reported) | | 87.6% | 53.1% | Not reported |
| Native liver survival at 10 years (n not reported) | | 76.9% | 43.9% | Not reported |
| Native liver survival at 12.5 years (n not reported) | | 48.5% | 36.6% | Not reported |
| Screening as a risk factor for native liver survival | | screening vs no screening: adjusted hazard ratio 2.61 (95% CI 1.20 to 5.70), p=0.016 | | |
| Age at surgery as a risk factor for native liver survival | | Age at Kasai portoenterostomy at more vs less than 90 days: adjusted hazard ratio 1.37 (95%CI 0.74 to 2.52), p=0.318 | | |
| Quality appraisal using the JBI checklist for cohort studies | | | | |
| Question | | | Assessment | Supporting information |
| Were the 2 groups similar and recruited from the same population? | | | Yes | Infants born in Japan |
| Were the exposures measured similarly to assign people to both exposed and unexposed groups? | | | Yes | Infants born in areas with or without screening |
| Was the exposure measured in a valid and reliable way? | | | Yes | |

| | | |
|---|---|---|
| Were confounding factors identified? | Yes | Considered in the discussion |
| Were strategies to deal with confounding factors stated? | Yes | |
| Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | |
| Were the outcomes measured in a valid and reliable way? | Yes | Objective outcomes assessed |
| Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | |
| Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | No | Data were not available for deceased patients in the control group |
| Were strategies to address incomplete follow up utilized? | No | |
| Was appropriate statistical analysis used? | Yes | |
| **Other comments** | Data were collected through questionnaires sent out to parents/guardians of children with biliary atresia. The authors stated that they were unable to collect data on deceased patients in the control group and that this "*skewed the native liver survival of controls to longer than they actually were*".<br><br>The authors conducted a sample size calculation that suggested that 110 patients were required in each group. The study may therefore have been underpowered to detect a difference between the groups.<br><br>Potential confounding factors were discussed. Hazard ratios were adjusted for use of stool colour card, sex, age at surgery, and type of biliary atresia<br><br>The authors concluded that long-term native liver survival in infants with biliary atresia is associated with stool colour card use for early detection but is not associated with age or a threshold of age at Kasai procedure (<90 days).<br><br>The report included infants with biliary atresia treated over a 19-year period up to 2011 in a county with a higher incidence of biliary atresia and a different healthcare system to the UK. The applicability of the outcomes reported to current UK practice is not clear. | |

## Table 35: Lee et al. 2016 [18]

| | |
|---|---|
| Publication | Lee M, Chen SC, Yang HY, Huang JH, Yeung CY, Lee HC. Infant stool color card screening helps reduce the hospitalization rate and mortality of biliary atresia: A 14-year nationwide cohort study in Taiwan. Medicine. 2016;95(12):e3166. |
| Study details | Retrospective cohort study |
| Study objectives | To examine whether the implementation of the stool colour card screening programme can improve biliary atresia case outcomes |
| Study setting | Taiwan National Health Insurance Research Database |
| Inclusions | Infants born between 2005 and 2010 after a national universal stool colour card screening programme was implemented<br><br>Infants born between 1997 and 2004 before screening was implemented |
| Exclusions | None stated |
| Population | 338 infants with biliary atresia diagnosed before the implementation of the screening programme in 2004 |

| | | | |
|---|---|---|---|
| | 175 infants with biliary atresia diagnosed after screening was implemented | | |
| Intervention | Stool colour cards placed within the Child Health Handbook that is provided for every newborn<br>Stool colour cards were collected when infants were brought to hospitals or clinics for vaccinations in the first 2 months of age | | |
| Comparator | No screening | | |
| Outcomes | Outcomes are reported for infants who underwent Kasai portoenterostomy. Follow-up duration not stated. | | |
| | Before screening | After screening | p value |
| Number of patients who underwent Kasai portoenterostomy | 301/338 (89.1%) | 156/175 (89.1%) | p=0.975 |
| Mean (SD) age at Kasai portoenterostomy (days) | 59.9 (76.4) | 48.2 (24.5) | p=0.064 |
| Proportion of infants who had Kasai portoenterostomy at >60 days old | 95/301 (31.6%) | 95/301 (31.6%) | p=0.242 |
| Mean (SD) number of hospitalisations per case until 2 years of age | 6.4 (4.1) | 5.0 (3.2) | p<0.001 |
| Mortality | 144/301 (47.8%) | 33/156 (21.2%) | p<0.001 |
| Proportion undergoing liver transplant | 86/301 (28.6%) | 44/156 (28.2%) | p=0.934 |

| Quality appraisal using the JBI checklist for cohort studies | | |
|---|---|---|
| Question | Assessment | Supporting information |
| Were the 2 groups similar and recruited from the same population? | Yes | Infants born Taiwan |
| Were the exposures measured similarly to assign people to both exposed and unexposed groups? | Yes | Before and after screening |
| Was the exposure measured in a valid and reliable way? | Yes | |
| Were confounding factors identified? | Yes | Considered in the discussion |
| Were strategies to deal with confounding factors stated? | No | |
| Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | |
| Were the outcomes measured in a valid and reliable way? | Yes | Objective outcomes assessed |
| Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | For the main outcome of interest. For longer-term outcomes such as liver transplantation rate the follow-up period is not sufficient for more recently diagnosed patients |
| Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | No | Outcomes were only reported for infants who underwent surgery |

| | |
|---|---|
| Were strategies to address incomplete follow up utilized? | No |
| Was appropriate statistical analysis used? | Yes |
| Other comments | The data were taken from a national health insurance programme that covers more than 99% of the country's population.<br><br>Not all of the infants diagnosed with biliary atresia received Kasai portoenterostomy. Outcomes were only reported for infants who received surgery.<br><br>The authors stated that they were unable to detect the proportion of infants with biliary atresia who were detected by stool colour card screening because this information was not recorded in the national database used as the data source.<br><br>Potential confounding factors such as accessibility of healthcare were discussed. No strategies to deal with confounding factors were reported for the outcomes extracted for this evidence summary.<br><br>Data were collected over a 14-year period. The follow-up duration for all patients was sufficient for outcomes related to surgery but was not sufficient for longer term outcomes relating to transplantation for patients treated in the later time periods.<br><br>The authors concluded that the hospitalisation and mortality rates of biliary atresia cases in Taiwan were significantly reduced after the launch of the stool colour card screening programme.<br><br>The report included infants with biliary atresia treated over a 14-year period up to 2010 in a county with a higher incidence of biliary atresia and a different healthcare system to the UK. The applicability of the outcomes reported to current UK practice is not clear. |

### Table 36: Lin et al. 2015 [28]

| | | | | |
|---|---|---|---|---|
| Publication | Lin JS, Chen SCC, Lu CL, Lee HC, Yeung CY, Chan WT. Reduction of the ages at diagnosis and operation of biliary atresia in Taiwan: A 15-year population-based cohort study. World Journal of Gastroenterology. 2015;21(46):13080-6. | | | |
| Study details | Retrospective cohort study | | | |
| Study objectives | To describe the ages at diagnosis and operation of infants with biliary atresia | | | |
| Study setting | Taiwan National Health Insurance Research Database | | | |
| Inclusions | Infants with biliary atresia born between 1997 and 2011 | | | |
| Exclusions | None stated | | | |
| Population | Infants were divided into 3 birth cohorts:<br>1997 to 2001: 208<br>2002 to 2006: 200<br>2007 to 2011: 132 | | | |
| Intervention | A national population screening programme using stool colours cards was introduced in 2004 | | | |
| Comparator | Outcomes were compared between 3 time periods before and after screening was introduced | | | |
| Outcomes | Follow-up duration not stated | | | |
| | 1997 to 2001 | 2002 to 2006 | 2007 to 2011 | p value |
| Number of patients who underwent Kasai portoenterostomy | 188/208 (90.4%) | 174/200 (87.0%) | 122/132 (92.4%) | Not reported |

| | | | | |
|---|---|---|---|---|
| Mean (SD) age at Kasai portoenterostomy (days) | 58.2 (± 42.0) | 50.5 (± 30.8) | 46.0 (± 23.8) | p=0.006 |
| Proportion of infants that received Kasai portoenterostomy within 60 days | 144/188 (76.6%) | 154/174 (88.5%) | 99/122 (81.1%) | p=0.285 |
| Proportion of infants with biliary atresia needing liver transplantation | 68/208 (32.7%) | 82/200 (41.0%) | 39/132 (29.6%) | p=0.782 |

| | For infants aged ≤ 60 days at surgery | For infants aged > 60 days at surgery | p value |
|---|---|---|---|
| Liver transplantation rate in infants who received Kasai portoenterostomy | 89/348 (25.6%) | 44/136 (32.3%) | p=0.133 |

| Quality appraisal using the JBI checklist for cohort studies | | |
|---|---|---|
| Question | Assessment | Supporting information |
| Were the 2 groups similar and recruited from the same population? | Yes | Infants born Taiwan |
| Were the exposures measured similarly to assign people to both exposed and unexposed groups? | Yes | Infants born in different time periods |
| Was the exposure measured in a valid and reliable way? | Yes | |
| Were confounding factors identified? | Yes | Considered in the discussion |
| Were strategies to deal with confounding factors stated? | No | |
| Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | |
| Were the outcomes measured in a valid and reliable way? | Yes | Objective outcomes assessed |
| Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | For the main outcome of interest. For longer-term outcomes such as liver transplantation rate the follow-up period is not sufficient for more recently diagnosed patients |
| Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | Yes | Data taken from a national database |
| Were strategies to address incomplete follow up utilized? | N/A | |
| Was appropriate statistical analysis used? | Yes | |
| Other comments | The national health insurance programme that the data were taken from covers more than 99% of the country's population.<br><br>It is not clear if all the infants with biliary atresia diagnosed after 2004 were detected by screening.<br><br>The authors noted that the reducing trends in age at diagnosis and surgery started before the screening programme was implemented in 2004 and that the decrease was not entirely due to the programme. They sug- | |

| | |
|---|---|
| | gested that confounding factors could include medical resource availability and better accessibility to surgical interventions and may have contributed to the improvement.<br><br>Data were collected over a 15-year period. The follow-up duration for all patients was sufficient for outcomes related to surgery but was not sufficient for longer term outcomes relating to transplantation for patients treated in the later time periods.<br><br>The authors concluded that screening using stool colour cards has the potential benefit of increasing awareness of biliary atresia amongst both parents and physicians.<br><br>The report included infants with biliary atresia treated over a 15-year period up to 2011 in a county with a higher incidence of biliary atresia and a different healthcare system to the UK. The applicability of the outcomes reported to current UK practice is not clear. |

### Table 37: Woolfson et al. 2018 [22]

| | |
|---|---|
| Publication | Woolfson JP, Schreiber RA, Butler AE, MacFarlane J, Kaczorowski J, Masucci L, et al. Province-wide biliary atresia home screening program in British Columbia: Evaluation of first 2 years. Journal of Pediatric Gastroenterology & Nutrition. 2018;66(6):845-9. |
| Study details | Screening study |
| Study objectives | To assess the performance of a province-wide stool colour card screening programme |
| Study setting | 126 maternity units in British Columbia, Canada |
| Inclusions | Infants born between April 2014 and March 2016 |
| Exclusions | Babies admitted to neonatal intensive care units were excluded as infants in these units are intensively monitored so the use of parental screening with stool colour cards was felt to be unnecessary |
| Population | 87,583 infants |
| Intervention | Stool colour cards given to families at maternity discharge. Families were instructed to monitor their infant's stool colour for one month and contact the screening centre with concerns. An online tool to remind parents to check their newborn's stool daily for one month was also made available to families through sign up online or screening a QR code on the stool colour card<br><br>From June 2015, a statement was added to the stool colour card instructing physicians to order a bilirubin tests for infants jaundiced at > 2 weeks of age |
| Comparator | No comparator for the outcomes reported |
| Outcomes | |
| Misclassification | One false-positive case had a Kasai portoenterostomy at 79 days "*due to uncertain diagnosis, despite careful post-referral subspeciality evaluation".* This infant was ultimately diagnosed with Alagille syndrome |
| Other outcomes | The study also reported age at surgery for cases detected after the implementation of screening. These data have not been extracted as studies reporting this outcome for screening vs no screening were identified and prioritised<br><br>The study also reported outcomes relating diagnostic accuracy. These outcomes are discussed under Question 1 above |
| See Table 29 for the critical appraisal of Woolfson et al. 2018 | |

## Table 38: Zheng et al. 2020 [19]

| | | | | |
|---|---|---|---|---|
| Publication | Zheng J, Ye Y, Wang B, Zhang L. Biliary atresia screening in Shenzhen: implementation and achievements. Archives of Disease in Childhood. 2020;105(8):720-3. | | | |
| Study details | Cohort study | | | |
| Study objectives | To assess the implementation and achievements of stool colour card screening for biliary atresia | | | |
| Study setting | Shenzhen Children's Hospital, China | | | |
| Inclusions | Infants diagnosed with biliary atresia by cholangiography, born in Shenzhen and diagnosed between January 2013 and September 2017 | | | |
| Exclusions | None stated | | | |
| Population | 50 patients diagnosed before screening was introduced on 1 January 2015<br>68 patients diagnosed after screening was introduced | | | |
| Intervention | Stool colour card inserted into the health handbook for children which is distributed to all pregnant women (after 1 January 2015)<br>There was no system for parents/guardians to report screening results | | | |
| Comparator | No screening | | | |
| Outcomes | Outcomes reported for infants who underwent Kasai portoenterostomy. Patients were followed-up for one to 5 years after discharge from hospital | | | |
| | Before screening | After screening | p value | |
| Number of patients who underwent Kasai portoenterostomy | 34/50 (68.0% | 57/68 (83.8%) | p<0.05 | |
| Mean (SD) age at Kasai portoenterostomy (days) | 81 (± 12) | 56 (± 15) | p<0.05 | |
| Proportion of infants diagnosed within 60 days | 12/34 (35.3%) | 37/57 (64.9%) | p<0.05 | |
| Post-operative complications | 20/34 (58.8%) | 30/57 (52.6%) | p<0.05 | |
| Post-operative jaundice-free rate | 16/34 (47.1%) | 31/57 (54.4%) | p<0.05 | |
| Mortality | 7/34 (20.6%) | 6/57 (10.5%) | p<0.05 | |
| 2-year liver survival rate | 15/34 (44.4%) | 30/57 (52.6%) | p<0.05 | |
| Liver transplantation rate | 13/34 (38.2%) | 23/57 (40.4%) | p>0.05 | |
| | Aged ≤ 60 days at Kasai portoenterostomy | Aged > 60 days at Kasai portoenterostomy | p value | |
| Liver transplantation rate | 15/49 (30.1%) | 21/42 (50.0%) | No statistical test reported | |
| Quality appraisal using the JBI checklist for cohort studies | | | | |
| Question | | Assessment | Supporting information | |
| Were the 2 groups similar and recruited from the same population? | | Yes | Infants born in one region in China | |
| Were the exposures measured similarly to assign people to both exposed and unexposed groups? | | Yes | Before and after screening | |
| Was the exposure measured in a valid and reliable way? | | Yes | | |

| | | |
|---|---|---|
| Were confounding factors identified? | Yes | Considered in the discussion |
| Were strategies to deal with confounding factors stated? | No | |
| Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | Yes | |
| Were the outcomes measured in a valid and reliable way? | Yes | Objective outcomes assessed |
| Was the follow up time reported and sufficient to be long enough for outcomes to occur? | Yes | For the main outcome of interest. For longer-term outcomes such as liver transplantation rate the follow-up period is not sufficient |
| Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | No | Outcomes were only reported for infants who underwent surgery |
| Were strategies to address incomplete follow up utilized? | No | |
| Was appropriate statistical analysis used? | Yes | |

| | |
|---|---|
| Other comments | The data reported was for infants born in one city in China before and after screening was introduced. The screening consisted of inserting a stool colour card into a health handbook. There was no system for parents/guardians to report screening results. The parents/guardians of infants diagnosed with biliary atresia were not asked if they had observed any abnormality based on the stool colour card. It is not clear if all the infants with biliary atresia were detected by screening.<br><br>Not all of the infants diagnosed with biliary atresia received treatment. The authors stated that this was because some parents refused treatment due to complex cultural and societal factors. Outcomes were only reported for infants who received surgery.<br><br>Potential confounding factors such as surgeon experience and access to liver transplant were discussed. No strategies to deal with confounding factors were reported. The study was conducted over a relatively short time frame, approximately 2 years before and after screening was introduced which reduces the likelihood of risk of bias due to confounding changes in clinical practice.<br><br>The follow-up duration was sufficient for outcomes related to surgery but was not sufficient for longer term outcomes relating to transplantation and mortality.<br><br>The authors concluded that the stool colour card screening programme led to earlier diagnoses, greater willingness to undergo treatment and better prognoses for infants with biliary atresia.<br><br>The report included infants with biliary atresia treated over a 5-year period up to 2017 in a county with a higher incidence of biliary atresia and a different healthcare system to the UK. The applicability of the outcomes reported to current UK practice is not clear. |

# Appendix 4 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 39.

Table 39: UK NSC reporting checklist

| Section | Item | Page no. |
|---|---|---|
| Title and summaries | | |
| Title Sheet | Identify the review as a UK NSC Evidence summary | Title page |
| Plain English summary | Plain English description of the executive summary. | 8 |
| Executive summary | Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review | 9 |
| Introduction and Approach | | |
| Background and objectives | Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews<br><br>Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search. | 13<br><br>16 |
| Methods | Briefly outline the rapid review methods used | 18 |
| Eligibility for inclusion in the review | State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided a priori | 18 |

| | | |
|---|---|---|
| Appraisal for quality/ risk of bias tool | Details of tool/ checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR. | 22 |
| Methods of analysis/synthesis | A narrative synthesis of results is presented, structured by UK NSC criterion and key question. No meta-analyses were conducted. | 22 |
| Databases/ sources searched | Give details of all databases searched (including platform/ interface and coverage dates) and date of final search. | 22 |
| | Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used. | Appendix 1 |
| | Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion. | Appendix 2 |
| Question level synthesis | | |
| Criterion | Include the UK NSC criteria under consideration | 23,34,37 |
| Question | Describe the question under investigation and briefly describe if it has been previously reviewed | 23,34,37 |
| Eligibility for inclusion in the review | Give an overview of the inclusion criteria for this question and a very brief summary of the reasons for exclusion after full paper review | 23,34,37 |
| Description of the evidence | For each question, give numbers of studies screened, assessed for eligibility, and inclusion in the review, with summary reasons for exclusion | 23,34,37 |
| Discussion of findings | For each question, provide a balanced discussion of the evidence and overview of the quality appraisal of the studies. | 24,35,38 |
| | In the appendix, include a table with each included study with the full citation, summary of the data relevant to the question (for example, study size, PICO, follow-up period, | |

| | | |
|---|---|---|
| | outcomes reported, statistical analyses etc.), results of any assessment of quality/risk of bias. | |
| Summary of findings relevant to criterion | Provide a conclusion on whether the criterion in question has been met and provide a summary justification for this conclusion. | 32,36,42 |
| Review Summary | | |
| Conclusions and implications for policy | Do findings indicate whether screening should be recommended?<br>IS further work warranted?<br>Are there gaps in the evidence highlighted by the review? | 44 |
| Limitations | Discuss limitations of the available evidence and of the review methodology if relevant. | 44 |

# References

[1] Solutions for Public Health. Newborn screening for biliary atresia: External review against programme appraisal criteria for the UK National Screening Committee. March 2017.
[2] Solutions for Public Health. Screening for biliary atresia: An evidence map to outline the volume and type of evidence related to screening for biliary atresia for the UK National Screening Committee. April 2021.
[3] Bassett MD, & Murray K F. Biliary atresia: recent progress. Journal of Clinical Gastroenterology. 2008;42(6):720–729.
[4] Livesey E, Cortina Borja M, Sharif K, Alizai N, McClean P, Kelly D, Hadzic N, Davenport M. Epidemiology of biliary atresia in England and Wales (1999-2006). Arch Dis Child Fetal Neonatal Ed. 2009;94(6):F451-5.
[5] Hartley JL, Davenport M, Kelly DA. Biliary atresia. Lancet. 2009;374(9702):1704-13.
[6] Office for National Statistics. Births in England and Wales: 2024. July 2025. Available from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2024/pdf (Accessed October 2025)
[7] Henderson P, Sutton E, Tayler R, Barclay AR, Hansen R. PWE-137 The epidemiology and outcome of biliary atresia in Scotland 2002 to 2013. Gut, Volume 64-suppl. 1: 2nd Digestive Disorders Federation Conference 22-25 June 2015 Abstracts
[8] Public Health Scotland. Births in Scotland year ending 31 March 2024. November 2024. Available from Births in Scotland - Year ending 31 March 2024 - Births in Scotland - Publications - Public Health Scotland (Accessed November 2025)
[9] NICE. Jaundice in newborn babies under 28 days. CG98. London: National Institute for Health and Clinical Excellence; 2010, last updated October 2023. Available from: Overview | Jaundice in newborn babies under 28 days | Guidance | NICE (Accessed September 2025).
[10] McKiernan PJ, Baker AJ, Kelly DA. The frequency and outcome of biliary atresia in the UK and Ireland. Lancet. 2000;355:25-29.
[11] Fawaz R, Baumann U, Ekong U, Fischler B, Hadzic N, Mack CL, et al. Guideline for the Evaluation of Cholestatic Jaundice in Infants: Joint Recommendations of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. J Pediatr Gastroenterol Nutr. 2017;64(1):154-168.
[12] Davenport M, Ong E, Sharif K, Alizai N, McClean P, Hadzic N, et al. Biliary atresia in England and Wales: results of centralization and new benchmark. J Pediatr Surg. 2011;46(9):1689-94.
[13] Wang KS. Section on Surgery, Committee on Fetus and Newborn, & Childhood Liver Disease Research Network. Newborn Screening for Biliary Atresia. Pediatrics. 2015;136(6):e1663–e1669.
[14] BMJ Best Practice. BMJ Best practice biliary atresia. BMJ Publishing Group; Last updated September 2023. Available from: Biliary atresia.pdf (Accessed September 2025)
[15] Gu YH, Yokoyama K, Mizuta K, Tsuchioka T, Kudo T, Sasaki H, et al. Stool color card screening for early detection of biliary atresia and long-term native liver survival: a 19-year cohort study in Japan. Journal of Pediatrics. 2015;166(4):897-902.e1.
[16] Gu YH, Zhao JQ, Kong YY, Yang HH, Diao M, Li L, et al. Repeatability and reliability of home-based stool color card screening for Biliary Atresia based on results in China and Japan. Tohoku Journal of Experimental Medicine. 2020;252(4):365-72.
[17] Chen SM, Chang MH, Du JC et al. Screening for biliary atresia by infant stool color card in Taiwan. Pediatrics. 2006;117(4):1147-54.

[18] Lee M, Chen SC, Yang HY, Huang JH, Yeung CY, Lee HC. Infant stool color card screening helps reduce the hospitalization rate and mortality of biliary atresia: A 14-year nationwide cohort study in Taiwan. Medicine. 2016;95(12):e3166.
[19] Zheng J, Ye Y, Wang B, Zhang L. Biliary atresia screening in Shenzhen: implementation and achievements. Archives of Disease in Childhood. 2020;105(8):720-3.
[20] Morinville V, Ahmed N, Ibberson C, Kovacs L, Kaczorowski J, Bryan S, et al. Home-based screening for biliary atresia using infant stool color cards in Canada: Quebec Feasibility Study. Journal of Pediatric Gastroenterology & Nutrition. 2016;62(4):536-41.
[21] Schreiber RA, Masucci L, Kaczorowski J, Collet JP, Lutley P, Espinosa V, et al. Home-based screening for biliary atresia using infant stool colour cards: a large-scale prospective cohort study and cost-effectiveness analysis. Journal of Medical Screening. 2014;21(3):126-32.
[22] Woolfson JP, Schreiber RA, Butler AE, MacFarlane J, Kaczorowski J, Masucci L, et al. Province-wide biliary atresia home screening program in British Columbia: Evaluation of first 2 years. Journal of Pediatric Gastroenterology & Nutrition. 2018;66(6):845-9.
[23] Ramonet MD, Gomez S, Morise S, Parga L, Caglio P, De Micheli M, et al. Early detection of neonatal cholestasis using a stool color card screening. [Spanish, English]. Archivos Argentinos de Pediatria. 2013;111(2):135-8.
[24] Madadi-Sanjani O, Blaser J, Voigt G, Kuebler JF, Petersen C. Home-based color card screening for biliary atresia: the first steps for implementation of a nationwide newborn screening in Germany. Pediatric Surgery International. 2019;35(11):1217-22.
[25] Guthery SL, Kyle Jensen M, Sean Esplin M, O'Brien E, Krong J, Srivastava R. Feasibility of biliary atresia newborn screening in an integrated health network. Journal of Pediatric Gastroenterology & Nutrition. 2024;79(5):954-61.
[26] Harpavat S, Garcia-Prats JA, Anaya C, Brandt ML, Lupo PJ, Finegold MJ, et al. Diagnostic yield of newborn screening for biliary atresia using direct or conjugated bilirubin measurements. JAMA. 2020;323(12):1141-50.
[27] Rabbani T, Shah J. Newborn screening for biliary atresia using direct bilirubin: An implementation science study. Journal of Medical Screening. 2025;32(2):61-6.
[28] Lin JS, Chen SCC, Lu CL, Lee HC, Yeung CY, Chan WT. Reduction of the ages at diagnosis and operation of biliary atresia in Taiwan: A 15-year population-based cohort study. World Journal of Gastroenterology. 2015;21(46):13080-6.
[29] Gong Z, Wu Y, Zheng L, Chen L, Lv Z. Can free carnitine or bilirubin in blood be used in neonatal screening for biliary atresia? European Journal of Pediatric Surgery. 2020;30(5):459-64.
[30] Chiu CY, Chen PH, Chan CF, Chang MH, Wu TC. Taiwan infant stool color card study G. Biliary atresia in preterm infants in Taiwan: a nationwide survey. Journal of Pediatrics. 2013;163(1):100-3.e1.
[31] Kong YY, Zhao JQ, Wang J, Qiu L, Yang HH, Diao M, et al. Modified stool color card with digital images was efficient and feasible for early detection of biliary atresia-a pilot study in Beijing, China. World Journal of Pediatrics. 2016;12(4):415-20.
[32] Arshad A, Gardiner J, Ho C, Rees P, Chadda K, Baker A, et al. Population-based screening methods in biliary atresia: a systematic review and meta-analysis. Archives of Disease in Childhood. 2023;108(6):468-73.
[33] Lee CS, Ni YH, Chen HL, Wu JF, Hsu HY, Chien YH, et al. A pilot study of biliary atresia newborn screening using dried blood spot matrix metalloproteinase-7. Journal of Pediatric Gastroenterology & Nutrition. 2023;76(4):418-23.
[34] Xiao Y, Zhou Y, Zhou K, Cai W. Targeted metabolomics reveals birth screening biomarkers for biliary atresia in dried blood spots. J Proteome Res. 2022;21(3):721-6.

[35] Gopal SH, Zebda R, Mohan A, Borovsky K, Takwoingi Y, Scandrett K, et al. Population-based screening strategies for biliary atresia in the newborn: A systematic review and meta-analysis. PLoS ONE 2024;19(8):e0307837.
[36] Bazian. Screening for biliary atresia: external review against programme appraisal criteria for the UK National Screening Committee. June 2012.
[37] Durkin N, Deheragoda M, Davenport M. Prematurity and biliary atresia: a 30-year observational study. Pediatric Surgery International. 2017;33(12):1355-61.
[38] Williams R, Alexander G, Armstrong I, Baker A, Bhala N et al. Disease burden and costs from excess alcohol consumption, obesity, and viral hepatitis: fourth report of the Lancet Standing Commission on Liver Disease in the UK. Lancet. 2018, 391(10125):1097-1107.
[39] Davenport M, Makin E, Ong EG, Sharif K, Dawrant M, Alizai N. The Outcome of a Centralization Program in Biliary Atresia: Twenty Years and Beyond. Ann Surg. 2025;281(4):608-14.
[40] Gu YH, Matsui A. Long-term native liver survival in infants with biliary atresia and use of a stool color card: Case-control study. Pediatrics International. 2017;59(11):1189-93.