Supplementary Appendix - Table 1: PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4 - 6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5 - 6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5 - 6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5 - 6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5 - 6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5 - 6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5 - 6

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5 - 6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5 - 6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5 - 6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5 - 6
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5 - 6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5 - 6
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	7
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	7 - 10
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	7 - 10
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	7
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	7
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11 - 15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data), role of funders for the systematic review.	5

Supplementary Appendix Table 2: MOOSE Checklist

Reporting Criteria	Reported (Yes/No)	Reported on Page
Reporting of Background		
Problem definition	Yes	4
Hypothesis statement	Yes	4 - 5
Description of Study Outcome(s)	Yes	4 - 5
Type of exposure or intervention used	No	N/A
Type of study design used	Yes	5
Study population	Yes	5
Reporting of Search Strategy		
Qualifications of searchers (eg, librarians and investigators)	Yes	1
Search strategy, including time period included in the synthesis and keywords	Yes	5 - 6
Effort to include all available studies, including contact with authors	Yes	5 -6
Databases and registries searched	Yes	5
Search software used, name and version, including special features used (eg, explosion)	Yes	5 - 6
Use of hand searching (eg, reference lists of obtained articles)	Yes	5
List of citations located and those excluded, including justification	Yes	7
Method for addressing articles published in languages other than English	Yes	5
Method of handling abstracts and	Yes	5

unpublished studies		
Description of any contact with authors	Yes	5
Reporting of Methods		
Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	Yes	5
Rationale for the selection and coding of data (eg, sound clinical principles or convenience)	Yes	5
Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)	Yes	5
Assessment of confounding (eg, comparability of cases and controls in studies where appropriate	Yes	5
Assessment of study quality, including blinding of quality assessors. stratification or regression on possible predictors of study results Y	Yes	5
Assessment of heterogeneity	No	Cannot be conducted within out study. Bias assessments conducted
Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient	Yes	5

detail to be replicated		
Provision of appropriate tables and graphics	Yes	See Tables and Figures
Reporting of Results		
Table giving descriptive information for each study included	Yes	See Tables and Figures
Results of sensitivity testing (eg, subgroup analysis)	Yes	See Tables and Figures
Indication of statistical uncertainty of findings	Yes	See Tables and Figures
Reporting of Discussion		
Quantitative assessment of bias (eg, publication bias)	Yes	7
Justification for exclusion (eg, exclusion of non–English-language citations)	Yes	5
Assessment of quality of included studies	Yes	7
Reporting of Conclusions		
Consideration of alternative explanations for observed results	Yes	11 - 15
Generalization of the conclusions (i.e., appropriate for the data presented and within the domain of the literature review)	Yes	11 - 15
Guidelines for future research	Yes	15
Disclosure of funding source	Yes	5

Supplementary Appendix Table 3: Databases searched for systematic review of population-based screening for Biliary Atresia

Database	Date range searched	Date searched	Number of results
Cochrane Central Register of Controlled Trials	1946 - current	10.09.2022	306
EBSCO–CINAHL (Cumulative Index to Nursing and Allied Health Literature)	1971 - current	10.09.2022	206
Google Scholar	-	10.09.2022	1670
Ovid–EMBASE	1974 – September 10th 2022	10.09.2022	1836
Ovid–HMIC (Health Management Information Consortium)	1979 to September 2022	10.09.2022	0
Ovid–MEDLINE	1946 - current	10.09.2022	953
Ovid–MEDLINE E-pub ahead of print	September 10th 2022	10.09.2022	0
Ovid–MEDLINE In-Process and Other Non-Indexed Citations	1946 - current	10.09.2022	56
PubMed	1963 - current	10.09.2022	807
Scopus	-	10.09.2022	1463

Web of Knowledge (science citation index expanded and conference proceedings citation index science)	1969 - current	10.09.2022	2080

Supplementary Appendix Table 4: Medline Ovid Search Strategy to identify studies on population-based screening for Biliary Atresia

1. exp Infant/ or exp Infant, Newborn/
2. exp Child/
3. neonate.mp
4. baby.mp
5. newborn.mp
6. neonates.mp
7. neonatal.mp
8. Screen.mp.
9. Screening.mp
10. exp Mass Screening/
11. exp Neonatal Screening/
12. exp Jaundice/ or exp Jaundice, Obstructive/ or exp Jaundice, Neonatal/
13. exp Cholestasis, Extrahepatic/ or exp Cholestasis/ or exp cholestasis, Intrahepatic/
14. exp Liver/
15. exp Bilirubin/
16. biliary.mp or exp Biliary Tract Surgical Procedures/ or Biliary Tract Diseases/ or exp Biliary Atresia/ or exp Biliary Tract/ or exp Liver Cirrhosis, Biliary/
17. cohort*.tw.
18. exp Epidemiologic Methods/
19. exp Case-Control Studies/
20. (case$ and control$).tw.
21. exp Cohort Studies/
22. exp Retrospective Studies/
23. exp Cross-Sectional Studies/
24. Animals/
25. animal stud*.mp.
26. exp "Review"/
27. exp Case Reports/

28. 1 or 2 or 3 or 4 or 5 or 6 or 7
29. 8 or 9 or 10 or 11
30. 12 or 13 or 14 or 15 or 16
31. 17 or 18 or 19 or 20 or 21 or 22 or 23
32. 24 or 25 or 26 or 27
33. 28 and 29 and 30 and 31
34. 33 not 32
35. Limit 34 to (English language and yr= '1975 – current')

Supplementary Appendix Table 5: Characteristics of included studies exploring population-based screening for Biliary Atresia

Study (study design)	Country	Age at Testing	Number Included	Screening Method	Outcomes	Comments
Akiyama et al. 1994 (Cross-sectional study)	Japan	Healthy Group (Mean) – 50.1 months BA group (mean) – 30 months	200 Healthy Infants, 8 BA and 8 Neonatal Hepatitis	Infrared reflectance spectrometry of Stool Samples	Sensitivity: 100%, Specificity: 95.2%	
Suzuki et al. 2011 (Prospective cohort study)	Japan	39 weeks	1148	Measurement of urinary sulfated bile acid (USBA)	Sensitivity - 100%, Specificity - 96%, NPV - 4%, PPV - 100%	Author's state high FPR may be secondary to the use of an ordinary mail collection and delivery system with no temperature regulation and the 3- to 6-day interval between sampling and receipt.
Masucci et al. 2019 (Cost-Effectiveness Analysis)	N/A	N/A	N/A	N/A	SCC cost approximately $192,000 more than no universal screening but led to eight life-years gained (incremental cost-effectiveness ratio (ICER) $24,065 per life-year gained). Screening using conjugated bilirubin testing	

					versus the colour card cost $2,369,199 more and led to five more life-years gained (ICER - $473,840 per life year gained), and so was not cost-effective.	
Gu et al. 2015 (Prospective Cohort Study)	Japan	Unclear, appears physician stools reviewed at 1 month of age	264071	SCC	Sensitivity - 76.5% (95% CI: 62.2 - 90.7), Specificity - 99.9% (95% CI: 99.9 - 100.0) NPV - 99.9% (95% CI: 99.9 - 99.9), PPV - 12.7% (8.2% - 17.3%) Age at Kasai before SCC 70.3 days. After SCC 59.7 (p = 0.03) Improved 5-, 10- and 15-year native liver survival (87.6%, 76.9% and 48.5%) compared to studies conducted in US, UK and France	
Tseng et al. 2011 (Retrospective Cohort Study)	Taiwan	Unclear, appears physician reviewed stools at 1 month of age	2,246,924 born before SCC. 1029879 born after SCC.	SCC	Median age at first presentation decreased (47 vs. 43, p = 0.028). Late referrals decreased from 9.5% to 4.9%. The median age of Kasai operation decreased (51 vs. 48. p = 0.051). The proportions of Kasai operation within 60 days decreased (68.9% vs. 73.6%, p = 0.31)	
Chen et al. 2006 (Prospective Cohort Study)	Taiwan	Stools reviewed at 1 month of age	Type 1: 29412 Type 2: 37632	SCC (Type 1: Labeled and Type 2: Unlabeled)	For the detection of BA before 60 days: • Type 1: Sensitivity 86.7%, Specificity 99.9%. NPV - 99.9%, PPV - 41.9%. • Type 2: Sensitivity: 88.8%, Specificity: 99.9%, NPV: 99.9%, PPV: 20.0	

					17 (58.6%) infants received a Kasai operation before 60 days of age.	
Woolfson et al. 2018 (Prospective Cohort Study)	Canada (British Columbia)	Stools reviewed daily up to 1 month of age	87,583	SCC	Sensitivity: 50%, Specificity: 99%, NPV: 99% and PPV: 4% Set-up and Operational Costs in 1st Year: $80,154.63. Operational costs in 2nd year: $330,033.82. Additional cost of $50,120.81 for program launch in its inaugural first year. After program start up, ISCC cost per birth, including ongoing administrative expenses was $0.86	
Harpavat et al. 2016 (Cross-sectional study)	USA	Newborns (exact age not detailed)	BA Cohort: 61 Non-BA Cohort: 9102	Conjugated Blood Measurements	Sensitivity - 100% (95% CI: 87.7 - 100), Specificity - 98.2% (95% CI: 97.9 - 98.4)	
Matsui et al. 1993 (Prospective Cohort Study)	Japan	1 month of age	104,309	Total 3x- OH bile acids were extracted from dried blood spots	Sensitivity: 63.6%, PPV: 0.62%	
Kong et al. 2016 (Prospective Cohort Study)	China	Daily check until 4 months of age	29 799	SCC	Sensitivity: 100%, Specificity: 99.9%, PPV - 8.3% (95% CI: 2.7-19.4)	
Lee et al. 2016 (Prospective Cohort	Taiwan	Review at 2 months	513 BA cases (Comparison done by BA cases before	SCC	SCC reduced the average Kasai operation age (59.9 vs. 48.2, p = 0.064).	

Study)		of age	and after screening introduced)		SCC reduced hospitalization rate in the first 2 years of life (6.4 vs. 5.0, p < 0.001). SCC also reduced the death rate within the UK (47.8 vs. 21.2, p < 0.001) and percentage of infants having neither LTX nor death (31.6% vs. 56.4%, P < 0.001). Finally, there was no significant difference in the rate of LTx (28.6% vs. 28.2, p = 0.934).	
Zhou et al. 2012 (Cross sectional study)	China	4 days after birth	292 normal infants, 17 neonatal jaundice and 8 biliary atresia	Bile acids from dried blood spots	With a cutoff of 0.63 mmol/L, produces a sensitivity: 79.1 (74.3 - 83.2), specificity: 62.5 (25 - 87.5)	
Lien et al. 2016 (Prospective Cohort Study)	Taiwan	Unclear, appears daily from birth	191 BA Infants	SCC	3-year overall survival improved after SCC implementation (64.0% vs. 89.2% P < 0.001). The 5-year survival rates with native liver in cohorts A and B were (37.5% vs. 64.3%, P = 0.01). The 5-year overall survival rates were 89.3% vs. 55.7%, (P < 0.001).	
Harpavat et al. 2020 (Prospective Cohort Study)	USA	After Birth	123,279 infants	Conjugated Blood Measurements	Sensitivity: 100.0% (95% CI, 56.1%-100.0%), Specificity: 99.9% (95% CI, 99.9%-99.9%), PPV: 5.9% (95% CI, 2.6%-12.2%), NPV: 100.0% (95% CI, 100.0%- 100.0%) Screening reduced age at presentation (56 vs. 36 days, p = 0.004) and proportion having Kasai < 30 days (12.5% vs. 57.9%, p = 0.003). Screening reduced the age the patient referred	

					to a specialist (44 vs. 25 days, p 0.003). After Kasai, infants in screening group had significantly faster time of bilirubin normalization, but no sig. difference in transplant free survival. Screening infants more likely to have a normal conjugated bilirubin by 90 days (41.7% vs. 78.9%, p = 0.03).	
Mogul et al. 2015 (Cost effectiveness Study)	N/A	N/A	N/A	SCC	With no screening, the 20-year cost was $142,479,725 with 3702 life- years, 74 deaths and 158 liver transplants. With SCC B, the cost was $133,893,563 with 3731.7 life-years, 71 deaths and 147 liver transplants. There was a >97% probability that screening with the stool color card would be cost saving and associated with an increase in life-years gained. Among all parameters, only stool color card specificity was associated with the potential for screening to no longer be cost saving.	
Gong et al. 2020 (Cross Sectional Study)	China	3 – 14 days after birth	52, 862	Free carnitine, unconjugated bilirubin (UBIL), Bilirubin monoglucuronide (BMG), and Bilirubin diglucuronide (BDG) in dry blood spots	Direct Bilirubin: • Using 30 u/mol as cut off - Sensitivity: 100%, Specificity: 52%. • Using 140 u/mol as cut off - sensitivity: 75%, Specificity: 99%. Free Carnitine: • Using 38 u/mol as cut off: sensitivity 85%, Specificity: 85%. • Using 38 u/mol as cut off - Sensitivity: 75%, Specificity: 94%.	

Zheng et al. 2020 (Retrospective cohort study)	China	Unclear – appears daily from birth	118 BA cases	SCC	SCC reduced age at Kasai (56 vs. 81, $p < 0.05$), Length of stay in hospital (44 vs. 49, $p < 0.05$). It improved 2-year native liver survival rate (44.4% vs. 52.6%, $p < 0.05$) and survival (20.6% vs. 10.5%, $p < 0.05$).	
Chiu et al. 2013 (Retrospective Cohort Study)	Taiwan	Daily from birth	197 BA Cases	SCC	Sensitivity in detecting BA using SCC before 60 days: 92.8%. 96.3% in the preterm infants	
Muraji et al. 2003 (Cross Sectional Study)	Japan	21 – 138 days	58 infants with Breast feeding Jaundice. 16 BA infants	Urinary excretion of sulfated bile acid	Sensitivity - 100%, FPR - 1.0%	
Hsaio et al. 2008 (Prospective Cohort Study)	Taiwan	Daily from Birth	422273 Infants	SCC	2004: Sensitivity - 72.5%, 2005 - 97.1% Proportion of Kasai < 60 days: 47.2% prior to SCC, vs. 60% in 2004 and 74.3% in 2005 (once SCC introduced) 1976 - 2000 ($p = 0.004$). Delayed operation rate beyond 90 days decreased over time, from 15.3% in 1976-2000 to 10.3% in 2002-2003 and 0% in 2004 and 2005	
Powell et al. 2003 (Prospective Cohort Study)	United Kingdom	Babies under 28 days	27654	Conjugated Bilirubin	Using bilirubin cut off of 18 u/mol/l: True positives: 2, False negatives: 0 False positives: 10, True negatives: 23,107	

Schreiber et al. 2014 (Prospective Cohort Study)	Canada	Daily up to 4 weeks	6187	SCC	Liver transplants decreased from 55 (no screening) to 52 (SCC). For a Canadian population, the increase in cost for passive screening, compared with no screening, is $213,584 and the gain in life years is 9.7 ($22,000 per life-year gained).	
Mushtaq et al. 1999 (Case-Control Study)	United Kingdom	Infants < 1 year	218 infants with cholestasis	Mass Spectrometry on Blood Spots	Sensitivity/Specificity/PPV/NPV: cut off of 25 umol/l produced figures of 85.3%, 94.0%, 14.2, and 0.16, and a cut off of 35 umol/l 70.5%, 97.8%, 32.0, and 0.30, respectively	Unfortunately, there is too much overlap between bile acid concentrations in infants with cholestasis and those in control infants for this to be used as a single screening test for cholestatic hepatobiliary disease in general and biliary atresia
Shen et al. 2016 (Cross Sectional Study)	China	Neonates ranging from 18 – 94 days	40 BA cases, 40 Neonates with Pneumonia	Light Spectrometry (with phone application – POOPMD)	Sensitivity - 100%, Specificity - 34/40	
Gu et al. 2017 (Case-Control Study)	Japan	Unclear – appears daily until 1 month physician review	148 BA cases	SCC	Kasai < 60 days: 55.9% vs. 40.4% (p = 0.109), Native liver survival 197.2 months before SCC vs. 81 months after SCC, p = 0.017)	

Liao et al. 2022	China	Newborns 0 – 60 days	38 BA cases	Direct Bilirubin	Using $\geq$ 1 mg dL as cut-off: Sensitivity 100% Specificity 77.26%	
Xiao et al. 2022	China	36 – 40 weeks	21 BA cases	THCA, 2-hydroxyglutaric acid, and indoleacetic acid in dried blood spots	Sensitivity of 90.48% (95% CI: 69.62% – 98.83%) and specificity of 92% (95% CI: 84.84% – 96.48%).	

SCC: Stool Colour Chart, PPV: Positive Predictive Value; NPV; Negative Predictive Value, 95% CI: 95% Confidence Intervals

Supplementary Appendix Table 6a: Quality assessment scores for cohort studies exploring population-based screening for Biliary Atresia

Study (Year)	Selection				Comparability		Exposure/Outcome			Subtotal Assessment			Conclusion
	1	2	3	4	1a	1b	1	2	3	S	C	E/O	
Cohort Studies													
Suzuki et al. (2019)	*	*	*	*	*	*	*	*	*	Good	Good	Good	Good
Gu et al. (2015)	*	*	*	*	*	*	*	*	*	Good	Good	Good	Good
Tseng et al. (2011)	*	No	*	*	No	No	*	*	*	Good	Poor	Good	Good
Chen et al. (2006)	*	*	*	*	*	*	*	*	*	Good	Good	Good	Good
Woolfson et al. (2018)	*	*	*	*	*	*	*	*	*	Good	Good	Good	Good
Matsui et al. (1993)	*	*	No	*	*	*	*	*	Follow-up rate 80%	Good	Good	Fair	Good
Kong et al. (2016)	*	*	*	*	*	*	*	*	*	Good	Good	Good	Good
Lee et al. (2016)	*	*	*	*	*	*	*	*	No statement	Good	Good	Fair	Good

Lien et al. (2011)	*	No (different provinces)	*	*	Different Years	Different Region	*	*	*	Good	Poor	Good	Good
Harpavat et al. (2020)	*	Different Source	*	*	Different Years	Different Region	*	*	*	Good	Poor	Good	Fair
Zheng et al. (2020)	*	*	No description	*	Different Years	*	*	*	*	Good	Poor	Good	Fair
Chiu et al. (2013)	*	No	*	*	N/A	N/A	*	*	No Statement	Good	-	Good	Good
Hsaio et al. (2008)	*	*	*	*	Different Years	Different Region	*	*	*	Good	Poor	Good	Good
Powell et al. (2003	*	No unexposed group	*	*	N/A	N/A	*	*	84.70%	Good	-	Good	Good
Schreiber et al. (2014)	*	No unexposed group	*	*	N/A	N/A	*	*	40% return rate	Good	-	Good	Good
Liao et al. (2022)	*	No unexposed group	*	*	N/A	N/A	*	*	*	Good	-	Good	Good

Supplementary Appendix Table 6b: Quality assessment scores for cross sectional studies exploring population-based screening for Biliary Atresia

Study (Year)	Selection				Comparability	Exposure/Outcome		Subtotal Assessment Conclusion			Conclusions
	1	2	3	4	1a	1	2	1	2	3	
Cross sectional studies											
Akiyama et al. (1994)	*	No	*	**	No	*	*	Good	Poor	Good	Good
Harpavat et al. (2016)	*	No	*	**	No	*	*	Good	Poor	Good	Good
Zhou et al. (2012)	*	No	*	**	**	**	*	Good	Good	Good	Good
Gong et al. 2020	*	No	*	**	No	**	*	Good	Poor	Good	Good
Muraji et al. (2013)	*	No	*	**	No	**	*	Good	Poor	Good	Good

Shen et al. (2016)	*	No	*	**	**	*	*	Good	Good	Good	Good
Xiao et al. (2022)	*	No	*	**	**	*	*	Good	Good	Good	Good

Supplementary Appendix Table 6c: Quality assessment scores for case-control studies exploring population-based screening for Biliary Atresia

Study (Year)	Selection				Comparability		Exposure/Outcome			Subtotal Assessment			Conclusion
	1	2	3	4	1a	1b	1	2	3	S	C	E/O	
Case-Control Studies													
Mushtaq et al. 1999	No	*	N/A	*	*	No	*	*	*	Good	Fair	Good	Good
Gu et al. 2017	No	*	N/A	*	*	No	*	*	*	Good	Fair	Good	Good

Supplementary Appendix Table 7: Summary of study data for a meta-analysis on the sensitivity and specificity of BA screening methods

Study Number	Reference	Method	Numbers of			
			True Positives	False Negatives	False Positives	True Negatives
4	Gu 2015	Stool Colour Chart	26	8	177	263859
6	Chen 2006		26	3	65	78090
10	Kong 2016		2	0	22	22775
18	Chiu 2013		181	13	n/a	n/a
20	Hsiao 2008		63	12	279	422273
21	Powell 2003	Blood Measurements of Bilirubin	2	0	10	23107
8	Harpavat 2016		35	0	166	8936
14	Harpavat 2020		7	0	122	123140
25	Liao et al. 2022		36	0	929	3157
1	Akiyama 1994	Stool Colour Saturation	8	0	10	198
24	Shen 2016		38	0	6	34
2	Suzuki 2011	Urinary Sulphated Bile Acids	1	0	6	1141
9	Matusi 1993	Bile Blood Spot Measurements	7	4	1129	103173
12	Zhou 2012		n/a	n/a	n/a	n/a
16	Gong 2020 (I)		97	0	4894	5204
16	Gong 2020 (II)		73	24	100	9908
23	Mushtaq 1999		n/a	n/a	n/a	n/a
16	Gong 2020 (III)	Carnitine Measurements	29	10	600	9408

Gong 2020 (I) results with cut-off bilirubin > 30 µmol/l.
Gong 2020 (II) results with cut-off bilirubin > 140 µmol/l.
Gong 2020 (III) results with cut-off free carnitine > 45 µmol/l.

Supplementary Appendix Table 8: Sensitivity and specificity for studies, with 95% confidence intervals.

Reference	Method	Sensitivity		Specificity	
		Estimate	Confidence Interval	Estimate	Confidence Interval
Gu 2015	Stool Colour Chart	76.47%	(58.83%,89.25%)	99.93%	(99.92%,99.94%)
Chen 2006		89.66%	(72.65%,97.81%)	99.92%	(99.89%,99.94%)
Kong 2016		100.00%	(15.81%,100.00%)	99.90%	(99.85%,99.94%)
Chiu 2013		93.30%	(88.81%,96.38%)	n/a	n/a
Hsiao 2008		84.00%	(73.72%,91.45%)	99.93%	(99.93%,99.94%)
Meta-analysis		87.90%	(80.40%, 92.80%)	99.99%	(99.99 – 99.99%)
Powell 2003	Blood Measurements of Bilirubin	100.00%	(15.81%,100.00%)	99.96%	(99.92%,99.98%)
Harpavat 2016		100.00%	(90.00%,100.00%)	98.18%	(97.88%,98.44%)
Harpavat 2020		100.00%	(59.04%,100.00%)	99.90%	(99.88%,99.92%)
Liao 2022		100.00%	(90.26%,100.00%)	77.26%	(75.95%,78.54%)
Meta-analysis		100.00%	(00.00%. 100.00%)	99.3%	(91.90% - 99.99%)
Akiyama 1994	Stool Colour Saturation	100.00%	(63.06%,100.00%)	95.19%	(91.34%,97.67%)
Shen 2016		100.00%	(90.75%,100.00%)	85.00%	(70.16%,94.29%)
Meta-analysis		100.00%	(0.00%, 100.00%)	92.4%	(83.4% - 96.7%)
Suzuki 2011	Urinary Sulphated Bile Acids	100.00%	(2.50%,100.00%)	99.48%	(98.86%,99.81%)
Matusi 1993	Bile Blood Spot Measurements	63.64%	(30.79%,89.07%)	98.92%	(98.85%,98.98%)
Zhou 2012		79.10%[1]	(74.30%,83.20%)	62.50%[1]	(25.00%,87.50%)
Gong 2020 (I)		100.00%	(96.27%,100.00%)	51.53%	(50.55%,52.51%)
Gong 2020 (II)		75.26%	(65.46%,83.46%)	99.00%	(98.79%,99.19%)
Mushtaq 1999 (I)		85.30%[1]	(75.50%,92.00%)	94.00%[1]	(92.30%,95.30%)
Mushtaq 1999 (II)		78.70%[1]	(68.10%,86.90%)	96.30%[1]	(94.90%,97.40%)
Meta-analysis		93.20%	(34.80%, 99.70%)	95.50%	(65.80% - 99.50%)
Gong 2020 (III)	Carnitine Measurements	74.36%	(57.87%,86.96%)	94.00%	(93.52%,94.46%)

[1] Sensitivity / specificity and their confidence intervals are derived from ROC curve.

Gong 2020 (I) results with cut-off bilirubin > 30 µmol/l.

Gong 2020 (II) results with cut-off bilirubin > 140 µmol/l.

Gong 2020 (III) results with cut-off free carnitine > 45 µmol/l.

Mushtaq 1999 (I) results with cut-off bilirubin > 25 µmol/l.

Mushtaq 1999 (II) results with cut-off bilirubin > 30 µmol/l.

Supplementary Appendix Table 9: Breakdown of factors included within cost-effectiveness analyses of BA screening methods

Paper	Type of BA Screening	Model Used	Cost of screening setup	Considered LTx costs	Considered Immunosuppression Cost	Considered Liver Transplant Follow-up	Considered Liver Transplant Complications
Woolfson et al.	SCC	Simple cost of set-up first and second year	Y	N	N	N	Y
Schreiber et al.	SCC	Markov Model	Y	N	N	N	Y
Masucci et al.	SCC and Conjugated Bilirubin Measurements	Markov Model	Y	Y	Y	Y	Y
Mogul et al.	SCC	Markov Model	Y	Y	Y	Y	Y

SCC: Stool Colour Chart, LTx: Liver Transplantation, Y: Yes, N: No