

Screening of newborn infants for cholestatic hepatobiliary disease with tandem mass spectrometry

Imran Mushtaq, Stuart Logan, Michael Morris, Andrew W Johnson, Angie M Wade, Deirdre Kelly, Peter T Clayton

Abstract

Objective To assess the feasibility of screening for cholestatic hepatobiliary disease and extrahepatic biliary atresia by using tandem mass spectrometry to measure conjugated bile acids in dried blood spots obtained from newborn infants at 7-10 days of age for the Guthrie test.

Setting Three tertiary referral clinics and regional neonatal screening laboratories.

Design Unused blood spots from the Guthrie test were retrieved for infants presenting with cholestatic hepatobiliary disease and from the two cards stored on either side of each card from an index child. Concentrations of conjugated bile acids measured by tandem mass spectrometry in the two groups were compared.

Main outcome measures Concentrations of glycodihydroxycholanoates, glycotrihydroxycholanoates, taurodihydroxycholanoates, and taurotrihydroxycholanoates. Receiver operator curves were plotted to determine which parameter (or combination of parameters) would best predict the cases of cholestatic hepatobiliary disease and extrahepatic biliary atresia. The sensitivity and specificity at a selection of cut off values for each bile acid species and for total bile acid concentrations for the detection of the two conditions were calculated.

Results 218 children with cholestatic hepatobiliary disease were eligible for inclusion in the study. Two children without a final diagnosis and five who presented at < 14 days of age were excluded. Usable blood spots were obtained from 177 index children and 708 comparison children. Mean concentrations of all four bile acid species were significantly raised in children with cholestatic hepatobiliary disease and extrahepatic biliary atresia compared with the unaffected children ($P < 0.0001$). Of 177 children with cholestatic hepatobiliary disease, 104 (59%) had a total bile acid concentration $> 33 \mu\text{mol/l}$ (97.5th centile value for comparison group). Of the 61 with extrahepatic biliary atresia, 47 (77%) had total bile acid concentrations $> 33 \mu\text{mol/l}$.

Taurotrihydroxycholanoate and total bile acid concentrations were the best predictors of both conditions. For all cholestatic hepatobiliary disease, a cut off level of total bile acid concentration of

$30 \mu\text{mol/l}$ gave a sensitivity of 62% and a specificity of 96%, while the corresponding values for extrahepatic biliary atresia were 79% and 96%.

Conclusion Most children who present with extrahepatic biliary atresia and other forms of cholestatic hepatobiliary disease have significantly raised concentrations of conjugated bile acids as measured by tandem mass spectrometry at the time when samples are taken for the Guthrie test. Unfortunately the separation between the concentrations in these infants and those in the general population is not sufficient to make mass screening for cholestatic hepatobiliary disease a feasible option with this method alone.

Introduction

Diagnosis of cholestatic hepatobiliary disease in infancy normally starts with clinical detection of jaundice, typically associated with pale stools and dark urine. This method of ascertainment poses major problems, however, because jaundice in the newborn period is common and cholestatic hepatobiliary disease relatively rare. Between 2.4% and 15% of newborn babies remain clinically jaundiced beyond 14 days of life^{1, 2} but only 0.04% to 0.2% of newborn babies have (conjugated) hyperbilirubinaemia due to cholestatic hepatobiliary disease^{3, 4}; the remainder have transient unconjugated hyperbilirubinaemia. Thus, there is a tendency to ignore jaundice until about 6 weeks (the first health surveillance check). The consequent delay in diagnosis of cholestatic hepatobiliary disease can adversely affect outcome.^{5, 6} For infants with extrahepatic biliary atresia the resolution of jaundice and long term survival are both significantly improved when surgery is performed before 60 days.^{5, 7, 8} All infants with cholestatic hepatobiliary disease are at increased risk of bleeding due to vitamin K deficiency, and infants with undiagnosed cholestasis contribute a disproportionate number of cases of severe bleeding⁹; these haemorrhages could be prevented with vitamin K.

The Yellow Alert campaign, launched in 1993, aimed to identify those infants with cholestatic hepatobiliary disease earlier by recommending that all babies clinically jaundiced at > 14 days have their urine tested for bilirubin or their blood tested for conjugated

Biochemistry Unit,
Institute of Child
Health, University
College London,
London
WC1N 1EH

Imran Mushtaq
research fellow

Andrew W Johnson
senior lecturer in
biochemistry

Peter T Clayton
professor of paediatric
metabolic disease and
hepatology

Epidemiology and
Biostatistics Unit,
Institute of Child
Health

Stuart Logan
senior lecturer in
epidemiology

Angie M Wade
lecturer in medical
statistics

Micromass UK
Limited,
Wythenshawe,
Manchester
M23 9LZ

Michael Morris
chemist (mass
spectrometry
applications)

Liver Unit,
Birmingham
Children's Hospital,
Ladywood
Middleway,
Ladywood,
Birmingham
B16 8ET

Deirdre Kelly
consultant paediatric
hepatologist

Correspondence to:
P T Clayton
P.Clayton@ich.ucl.
ac.uk

BMJ 1999;319:471-7

bilirubin, or both.⁵ There are two problems with this: the substantial numbers of children requiring testing and difficulties with the quality control of assays for conjugated bilirubin.¹⁰ The practice has not been widely implemented and, in 1995, the British Paediatric Surveillance Unit reported that there was still unacceptable delay in the referral of many cases of extrahepatic biliary atresia.¹¹ This led us to consider how screening for cholestatic hepatobiliary disease might be undertaken in a way that avoids the laborious aspects of Yellow Alert. One answer was to look for a marker of the disease in blood that had already been taken at 7-10 days for phenylketonuria (PKU) screening—the Guthrie test.

Technological advances (in particular tandem mass spectrometry) have increased the number of metabolites that can be detected in a blood spot.¹² We set out to determine whether cholestatic hepatobiliary disease could be ascertained by measurement of compounds that accumulate in blood in cholestasis: conjugated bilirubin and bile acids. Preliminary experiments indicated that conjugated bilirubin was photolabile and difficult to measure with tandem mass spectrometry. Thus, we investigated the feasibility of the detection of cholestatic disease, in particular extrahepatic biliary atresia, by the measurement of conjugated bile acids in the dried blood spots currently obtained from all newborn infants in the United Kingdom at 7-10 days.

Subjects and methods

Cholestatic cases

Cases were identified from King's College, Great Ormond Street, and Birmingham Children's Hospitals. Inclusion criteria were referral for conjugated hyperbilirubinaemia, age < 1 year at presentation, and birth in the United Kingdom. The case notes were reviewed to ascertain presenting history and final diagnosis. Unused blood spots originally collected for phenylketonuria screening were used for this study. The Guthrie cards were retrieved from neonatal screening centres, where they are stored at room temperature for 6 months to 15 years. Infants were excluded from the study if no diagnosis was made; cholestasis was due to prolonged parenteral nutrition; no blood spot was available; or presentation was at < 14 days.

Comparison group

To provide a comparison group matched for storage time, two anonymised screening cards stored on either side of each index card were analysed. Ethical approval required that no identifying information be given on the comparison spots.

Sample analysis

Blood spots were coded and measurements performed blind. The 6 mm discs were punched and sonicated in 250 µl methanol containing 100 pmol of four internal standards labelled with deuterium.¹³ Concentrations of glycodihydroxycholanoates, glycotrihydroxycholanoates, taurodihydroxycholanoates, and taurotrihydroxycholanoates were measured with a VG Quattro II triple quadrupole spectrometer (Micromass, Altrincham).¹³ Total bile acid concentration was calculated from the four concentrations. Settings were cone volt-

age 60 V, collision energy 40 eV, product ion mass:charge ratio (m:z) 74 for glycodihydroxycholanoates and glycotrihydroxycholanoates; 100 V, 90 eV, and m:z 80 for taurodihydroxycholanoates and taurotrihydroxycholanoates; collision gas pressure 2.5 mTorr (0.33 Pa); and spectra were acquired in multiple reaction monitoring mode to facilitate background subtraction. The detection limit was 0.5 µmol/l, the "carry over" (from a 500 µmol/l sample) negligible. Coefficients of variation within and between assays were determined on blood samples to which bile acids had been added, increasing the concentrations of the four bile acid species by 7.5 µmol/l or 75 µmol/l. Coefficients of variation within assays were 5.2, 5.4, 5.4, and 3.9% at 7.5 µmol/l enrichment and 8.2, 6.6, 8.0, and 7.1% at 75 µmol/l, respectively (n=5 assays). Coefficients of variation between assays were 14.3, 13.5, 8.4, and 11.9% at 7.5 µmol/l spiking and 8.0, 10.7, 8.1, and 8.1% at 75 µmol/l spiking, respectively (n=5). Spiked blood samples were also used to study storage of blood spots exposed to light, in darkness, or in an airtight container with desiccant. Replicate analyses (n=5) were performed after 0, 3, 6, 9, and 12 months' storage. There was no significant change in any bile acid concentration with short (3 month) or long term (>6 months) storage under any of the conditions studied.

Data analysis and statistical analysis

Data were analysed with SPSS for Windows. A weighted $1/x$ linear regression model was used to produce calibration curves that were then used to calculate concentrations. Blood bile acid concentrations in case and comparison groups showed a log normal distribution. Values are therefore expressed as the geometric mean with 95% confidence intervals. The 95% range represented the distribution of the log transformed data between the 2.5th and the 97.5th centiles. Differences in bile acid concentrations between the case and comparison groups were evaluated with unpaired *t* tests on log transformed data. Regression analysis was performed on bile acid concentrations in the comparison group to determine the extent to which the age of a blood spot influenced its bile acid content.

The ability of bile acid concentration to discriminate between case and comparison groups was investigated at various cut off concentrations for the four bile acids by using GraphROC for Windows (Maxiwatti Oy, Kerttulink, Finland). Sensitivity and specificity for each cut off concentration were calculated and receiver operating characteristic curves plotted. The ability of bile acid concentration to diagnose cholestatic hepatobiliary disease and extrahepatic biliary atresia was compared between concentrations for the four bile acids by using the methods of Hanley and McNeil to compare areas below the curves.¹⁴ Two group (cases versus comparison group) discriminant analysis (SPSS for Windows) was performed on log transformed data to investigate whether any combination of the four values discriminated better than the total bile acid value.

Ethical approval

Approval was obtained from the ethics committees of the three centres. Anonymous testing of blood spots

Table 1 Underlying cause of cholestatic hepatobiliary disease

Diagnosis	No of infants
Extrahepatic biliary atresia	61
Idiopathic neonatal hepatitis	52
Alagille syndrome	17
α_1 Antitrypsin deficiency	15
Inherited metabolic disorders*:	
Galactosaemia	5
Tyrosinaemia	1
Mitochondrial cytopathy	1
Unspecified†	3
Inspissated bile syndrome	7
Idiopathic bile duct hypoplasia	6
Cytomegalovirus hepatitis	5
Fulminant hepatitis	1
Hypopituitarism	2
Sclerosing cholangitis	1
Total	177

*Excluding α_1 antitrypsin deficiency.

†Cholestatic jaundice from 1st week of life and hepatic steatosis on liver biopsy.

for the comparison group was in keeping with established procedures.

Results

Effect of storage on bile acid content of dried blood spots

Linear regression showed the mean change (95% confidence intervals) in bile acid content per year of storage to be glycodihydroxycholanoates $-0.03 \mu\text{mol/l}$ (-0.22 to $0.16 \mu\text{mol/l}$), glycotrihydroxycholanoates $0.12 \mu\text{mol/l}$ (-0.28 to $0.52 \mu\text{mol/l}$), taurodihydroxycholanoates $-0.04 \mu\text{mol/l}$ (-0.20 to $0.13 \mu\text{mol/l}$), and taurotrihydroxycholanoates $0.22 \mu\text{mol/l}$ (-0.02 to $0.46 \mu\text{mol/l}$). Examination of the residuals showed an even distribution with no element of underfitting, suggesting that fitting a more complex model than linear regression to the data was not warranted. On this basis (as well as the previous experiments on storage of blood spots) the four conjugated bile acid species were considered to be stable under normal storage conditions and matching of comparison to index cases for statistical analyses was considered unnecessary.¹⁵

Cholestatic cases

In total, 218 infants with cholestatic hepatobiliary disease were suitable for the study. Two infants had no diagnosis and five presented at <14 days; these were excluded. Blood spots were obtained from Guthrie cards of 177 of the remaining 211 infants. The median age at presentation was 47 days (range 14-248 days); the diagnoses are listed in table 1. The blood spots were obtained when the infants were aged 1-4 days ($n=5$), 5-10 days ($n=150$), 11-18 days ($n=18$), or 28-30 days ($n=4$). The cards had been in dry, room temperature storage for 2 months to 4 years.

The blood spot bile acid concentrations in both groups are shown in table 2. Results are expressed as geometric means with 95% confidence intervals. The 95% range is given for the comparison group. Blood spot concentrations of taurotrihydroxycholanoates and total bile acids are shown as dot plots in figures 1 and 2.

Comparison group

The trihydroxycholanoates, glycotrihydroxycholanoates and taurotrihydroxycholanoates, were predominant in the normal neonates, with the mean trihydroxycholanoate:dihydroxycholanoate (TriOH/DiOH) ratio being 2.02. The glycine-conjugated bile acids—glycotrihydroxycholanoates and glycodihydroxycholanoates—showed a slight preponderance over the taurine-conjugated bile acids—taurotrihydroxycholanoate and taurodihydroxycholanoate. On the basis of the 97.5th centile value, the upper limit of “normal” for total bile acid concentration in the comparison group was $33 \mu\text{mol/l}$ (95% confidence interval 31 to $42 \mu\text{mol/l}$).

Cholestatic cases versus comparison group

Concentrations of all four bile acid types were significantly raised and the TriOH/DiOH ratio was significantly higher in the cholestatic group (table 3). No significant difference in the glycine:taurine conjugate

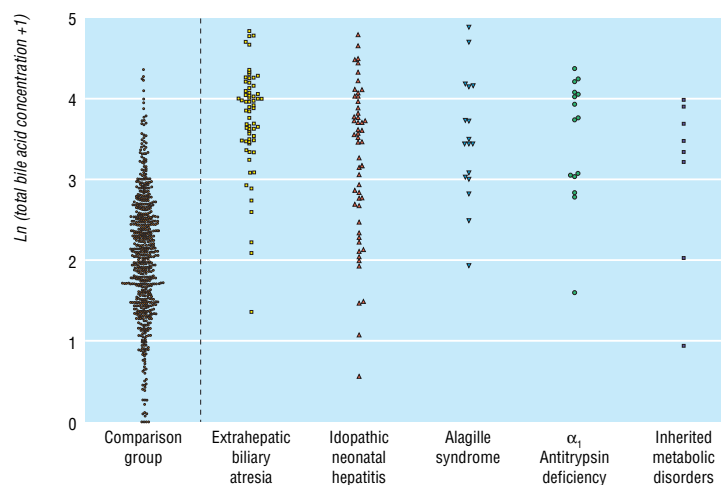


Fig 1 Total bile acid concentrations (logarithmic scale) in subgroups of infants with cholestasis and comparison group

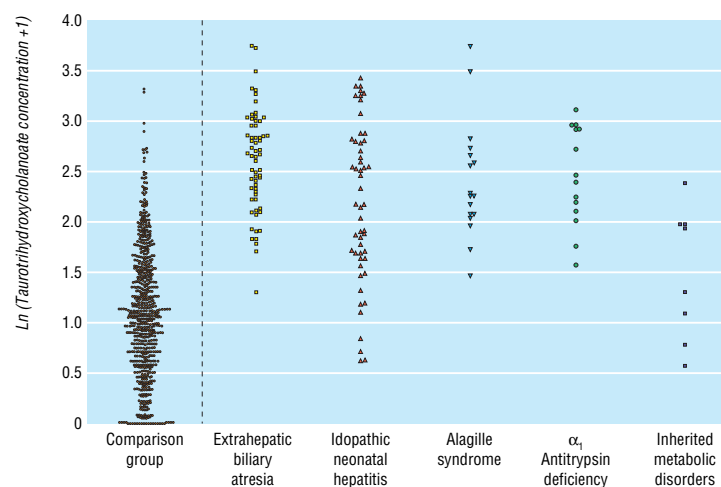


Fig 2 Taurotrihydroxycholanoate concentrations (logarithmic scale) in subgroups of infants with cholestasis and comparison group

Table 2 Blood spot bile acid concentrations (geometric mean and 95% confidence interval) in cholestatic cases and comparison group

Group	Bile acid concentration ($\mu\text{mol/l}$)					Bile acid ratio (95% confidence interval)	
	GD	GT	TD	TT	TBA	TriOH:DiOH	G:T
All cases of cholestasis (n=177)	5.2 (4.4 to 6.0)	9.7 (8.2 to 11.4)	4.1 (3.5 to 4.8)	9.2 (8.1 to 10.4)	31.0 (28.3 to 35.2)	2.62 (2.30 to 3.00)	1.30 (1.16 to 1.44)
Extrahepatic biliary atresia (n=61)	6.9 (5.6 to 8.4)	17.9 (14.7 to 21.9)	4.1 (3.4 to 4.9)	12.9 (11.2 to 14.8)	43.0 (37.5 to 0.6)	3.07 (2.71 to 3.48)	1.55 (1.36 to 1.74)
Idiopathic neonatal hepatitis (n=52)	4.3 (3.1 to 6.0)	7.0 (5.2 to 9.4)	3.9 (2.7 to 5.5)	8.0 (6.2 to 10.2)	25.3 (19.3 to 33.1)	2.60 (1.94 to 3.35)	1.15 (0.90 to 1.43)
Alagille syndrome (n=17)	4.4 (2.2 to 8.0)	9.6 (5.8 to 15.5)	4.2 (2.1 to 7.8)	10.3 (7.3 to 14.2)	33.0 (21.9 to 49.4)	3.16 (1.95 to 4.88)	1.18 (0.70 to 1.79)
α_1 Antitrypsin deficiency (n=15)	6.6 (4.1 to 10.2)	10.0 (4.8 to 19.9)	4.5 (2.7 to 7.3)	9.4 (6.1 to 14.3)	33.1 (21.2 to 51.2)	2.03 (1.25 to 3.07)	1.36 (0.89 to 1.95)
Inherited metabolic disorders (n=10)	4.9 (0.9 to 17.2)	4.8 (1.4 to 12.7)	4.4 (1.1 to 13.1)	3.5 (1.6 to 6.9)	22.5 (9.6 to 50.9)	2.00 (0.23 to 6.29)	1.50 (0.77 to 2.54)
Comparison group (n=708)	1.3 (1.2 to 1.4)	2.5 (2.3 to 2.7)	1.2 (1.1 to 1.3)	2.0 (1.8 to 2.1)	6.9 (6.5 to 7.4)	2.02 (1.90 to 2.15)	1.42 (1.33 to 1.52)
95% Range	0.0 to 5.8	0.0 to 14.9	0.0 to 5.11	0.0 to 9.5	0.0 to 33.3	0.00 to 10.57	0.00 to 6.75

GD=glycodihydroxycholanoates, GT=glycotrihydroxycholanoates, TD=taurodihydroxycholanoates, TT=taurotrihydroxycholanoates, TBA=total bile acids, TriOH=trihydroxycholanoate conjugates, DiOH=dihydroxycholanoate conjugates, G=glycine conjugates, T=taurine conjugates.

ratio was noted between the comparison group and the cholestatic group or subgroups. Of the infants in the cholestatic group, 104 (59%) had a total bile acid concentration $>33 \mu\text{mol/l}$.

Idiopathic neonatal hepatitis subgroup—Although concentrations of all four bile acid species were significantly raised in this group (n=52), the increases were less distinct than in other cholestatic disorders (table 3). Twenty seven cases (52%) had total bile $>33 \mu\text{mol/l}$.

Extrahepatic biliary atresia subgroup—This subgroup had the highest concentrations of glycodihydroxycholanoates, glycotrihydroxycholanoates, and taurotrihydroxycholanoates in the cholestatic group (n=61). The trihydroxycholanoates—glycodihydroxycholanoates and taurotrihydroxycholanoates—were particularly raised, with a mean TriOH:DiOH ratio of 3.07 compared with 2.02 in the comparison group and 2.54 in the idiopathic neonatal hepatitis subgroup ($P<0.03$). Forty seven (77%) out of the 61 cases in the extrahepatic biliary atresia subgroup had total bile acid values $>33 \mu\text{mol/l}$, seven cases had total bile acid levels of 20–33 $\mu\text{mol/l}$ and seven had values $<20 \mu\text{mol/l}$. Figures 1 and 2 and tables 2 and 3 suggest that measurements of taurotrihydroxycholanoates may be the most useful for distinguishing extrahepatic biliary atresia from other cholestatic disorders—for example, inherited metabolic disorders.

Other subgroups—Bile acid concentrations were significantly raised in babies with Alagille syndrome and α_1 antitrypsin deficiency when compared with the comparison group. Babies with metabolic disorders had only moderately raised bile acids.

Effect of age at time of Guthrie test

The data were analysed to examine the influence of age at sampling on bile acid concentrations in neonates destined to present with cholestatic hepatobiliary disease. The neonates with cholestasis were divided into those sampled before 7 days (n=76), from 7 to 10

days (n=79), and after 10 day (n=22). By using unpaired *t* tests on log transformed data differences between the three sampling periods were investigated, first for all causes of cholestatic hepatobiliary disease and then for the major subgroups: extrahepatic biliary atresia, idiopathic neonatal hepatitis, and α_1 antitrypsin deficiency. Concentrations of glycodihydroxycholanoates, taurodihydroxycholanoates, and taurotrihydroxycholanoates and total bile acids were significantly higher in those blood samples from the cholestatic group that were obtained after 10 days. Examination of the individual subgroups made it clear that the effect of sample time applied only to idiopathic neonatal hepatitis.

Performance of the tandem mass spectrometry analyses as a screening test

Receiver operating characteristic curves—Figure 3 shows receiver operating characteristic curves plotted for cholestatic hepatobiliary disease and extrahepatic biliary atresia. There is a significantly larger area under the curve for taurotrihydroxycholanoates and total bile acid concentration ($P<0.01$) compared with taurodihydroxycholanoates, glycotrihydroxycholanoates, and glycodihydroxycholanoates concentration, with no significant difference between taurotrihydroxycholanoates and total bile acids ($P=0.09$). This indicates that taurotrihydroxycholanoates and total bile acids are the best predictors of cholestatic hepatobiliary disease and extrahepatic biliary atresia. Discriminant analysis did not yield any significantly better combination for prediction of cholestatic hepatobiliary disease and extrahepatic biliary atresia.

Sensitivity and specificity of a total bile acid assay: positive and negative predictive values—Sensitivity and specificity values calculated for various total bile acid cut off levels are shown in tables 4 and 5. For cholestatic hepatobiliary disease (table 4) a cut off of 25 $\mu\text{mol/l}$ produced a sensitivity of 65.9%, a specificity of 94.0%, a positive likelihood ratio of 11.0, and a negative likelihood ratio of 0.36. With a cut off of 35 $\mu\text{mol/l}$ these figures were 53.3%, 97.8%, 24.2, and 0.48, respectively. For extrahepatic biliary atresia (table 5) a cut off of 25 $\mu\text{mol/l}$ produced figures of 85.3%, 94.0%, 14.2, and 0.16, and a cut off of 35 $\mu\text{mol/l}$ 70.5%, 97.8%, 32.0, and 0.30, respectively. Tables 4 and 5 also indicate the positive and negative predictive values of a screening test based on total bile acid measurement for different prevalences of cholestatic hepatobiliary disease and extrahepatic biliary atresia. (Estimates of prevalence vary from 1 in 500 to 1 in 2500 for cholestatic hepatobiliary disease and from 1 in 10 000 to 1 in

Table 3 Fold changes in bile acid concentration in cholestasis

Group	GD	GT	TD	TT	TBA	TriOH:DiOH ratio	G:T ratio
Cases of cholestasis (n=177)	4.0*	3.9*	3.4*	4.6*	4.5*	1.3†	0.9
Extrahepatic biliary atresia (n=61)	5.3*	7.2*	3.4*	6.5*	6.2*	1.5*	1.1
Idiopathic neonatal hepatitis (n=52)	3.3*	2.8*	3.3*	4.0*	3.7*	1.3	0.8
Alagille syndrome (n=17)	3.5‡	3.9*	3.5‡	5.2*	4.8*	1.6	0.8
α_1 Antitrypsin deficiency (n=15)	5.1*	4.0‡	3.8†	4.7*	4.8*	1.0	1.0
Inherited metabolic disorders (n=10)	3.8	1.9	3.7	1.8	3.3‡	1.0	1.1

* $P<0.0001$, † $P<0.001$, and ‡ $P<0.05$ compared with comparison group (unpaired *t* test). GD=glycodihydroxycholanoates, GT=glycotrihydroxycholanoates, TD=taurodihydroxycholanoates, TT=taurotrihydroxycholanoates, TBA=total bile acids, TriOH=trihydroxycholanoate conjugates, DiOH=dihydroxycholanoate conjugates, G=glycine conjugates, T=taurine conjugates.

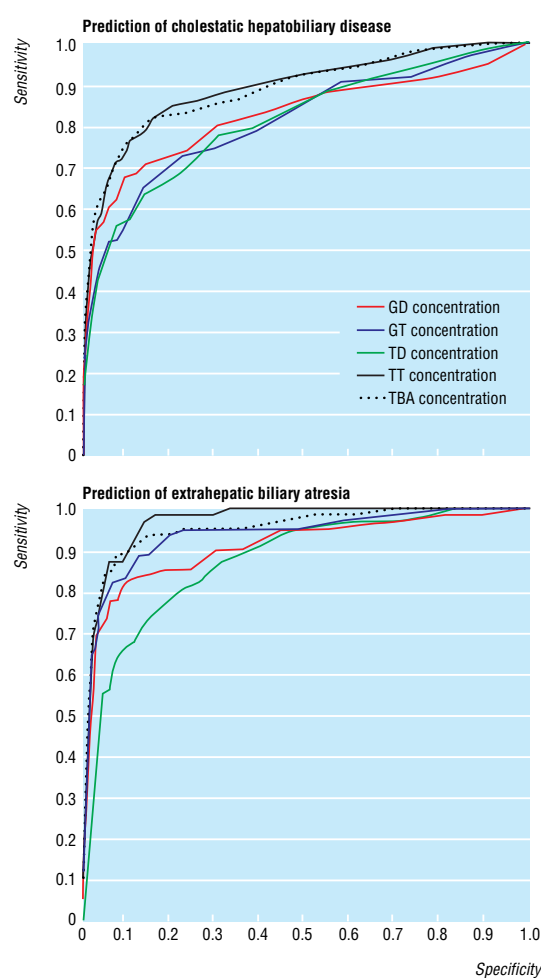


Fig 3 Receiver operating characteristic curves for blood spot bile acid concentrations for cholestatic hepatobiliary disease and extrahepatic biliary atresia. GD=glycodihydroxycholanoates; GT=glycotrihydroxycholanoates; TD=taurodihydroxycholanoates; TT=taurotrihydroxycholanoates; TBA=total bile acids

20 000 for extrahepatic biliary atresia.) If the prevalence of cholestatic hepatobiliary disease is as high as 1 in 500, the positive predictive value of a total bile acid value $> 35 \mu\text{mol/l}$ is 4.6%; if the prevalence is only 1 in 2500, the value is 0.96%.

The data were reanalysed to determine whether sensitivity, specificity, and positive and negative predictive values were improved if children who presented after 90 days were excluded (on the basis that children presenting between 3 and 10 months may not have cholestasis at 7-10 days). Concentrations of taurotrihy-

droxycholanoates and total bile acid were evaluated for the 147 children with cholestatic hepatobiliary disease (including 57 with extrahepatic biliary atresia) who presented between 14 and 90 days. The sensitivity, specificity, and negative and positive predictive values were no better than those obtained when all subjects (presenting from 14 to 284 days) were included.

Discussion

Infants with cholestatic hepatobiliary disease have raised concentrations of blood bile acids at presentation.^{13 16 17} We have shown that the mean concentrations of bile acids measured by tandem mass spectrometry on blood spot samples taken at 7-10 days are already greatly increased in neonates with cholestatic hepatobiliary disease, particularly those with extrahepatic biliary atresia. Can this be used for screening?

In 1995 Matsui et al reported a pilot study of screening for extrahepatic biliary atresia. They used an enzymatic method to measure total bile acids in blood spots taken at 0-10 days.^{18 19} They identified only seven of the 11 children with extrahepatic biliary atresia in the screened cohort, four of these beyond the optimum time for referral. Tandem mass spectrometry has provided technology capable of measuring individual bile acid species in a blood spot and of measuring total bile acids with greater precision. If a good test for cholestasis based on tandem mass spectrometry could be performed on blood spots obtained in the United Kingdom at 7-10 days, screening for cholestatic hepatobiliary disease might be practicable.

Problems with a single test

Unfortunately, although significant differences in mean bile acid concentrations were found in this study, a small but important overlap exists between the population distributions of unaffected neonates and those with cholestatic hepatobiliary disease. The separation is greater between neonates with extrahepatic biliary atresia and normal neonates, but even here the overlap is too great to make screening by this method alone a feasible option. A decision rule producing a specificity of 97.8% (that is, a 2.2% recall rate for diagnostic testing) results in a sensitivity of only 53% for cholestatic hepatobiliary disease and 70% for extrahepatic biliary atresia. Thus, for every 100 000 infants screened there will be 2200 false positive results. If the prevalence of cholestatic hepatobiliary disease is 1 in 500 and that of extrahepatic biliary atresia 1 in 10 000, the 100 000 babies screened will include 200 with

Table 4 Sensitivity and specificity of measurements of total bile acid concentrations in prediction of cholestatic hepatobiliary disease. Positive predictive value (PPV) and negative predictive value (NPV) of test, assuming prevalence of cholestasis of 1 in 500 or 1 in 2500

Cut off ($\mu\text{mol/l}$)	% Sensitivity (95% CI)	% Specificity (95% CI)	Prevalence of disease			
			1 in 500		1 in 2500	
			PPV %	NPV %	PPV %	NPV %
25.0	65.9 (59.6 to 71.2)	94.0 (92.0 to 94.7)	2.14	99.93	0.43	99.99
30.0	61.5 (55.4 to 67.9)	96.3 (94.9 to 97.2)	3.22	99.92	0.66	99.98
35.0	53.3 (47.1 to 59.8)	97.8 (97.2 to 99.1)	4.63	99.90	0.96	99.98
48.5	35.2 (29.3 to 41.1)	99.0 (98.1 to 99.5)	6.83	99.87	1.44	99.97
69.0	12.1 (8.4 to 16.9)	99.5 (98.7 to 99.8)	4.22	99.82	0.87	99.96
79.0	7.1 (4.3 to 11.1)	100.0 (99.6 to 100.0)	100	99.81	100	99.96

Table 5 Sensitivity and specificity of total bile acid concentration measurements in prediction of extrahepatic biliary atresia. Positive predictive value (PPV) and negative predictive value (NPV) of test, assuming prevalence of extrahepatic biliary atresia of 1 in 10 000 or 1 in 20 000

Cut off ($\mu\text{mol/l}$)	% Sensitivity (95% CI)	% Specificity (95% CI)	Prevalence of extrahepatic biliary atresia			
			1 in 10 000		1 in 20 000	
			PPV %	NPV %	PPV %	NPV %
25.0	85.3 (75.5 to 92.0)	94.0 (92.3 to 95.3)	0.14	100.00	0.07	100.00
30.0	78.7 (68.1 to 86.9)	96.3 (94.9 to 97.4)	0.21	100.00	0.11	100.00
35.0	70.5 (59.4 to 79.9)	97.8 (96.7 to 98.6)	0.32	100.00	0.16	100.00
48.5	50.8 (39.6 to 62.0)	99.0 (98.2 to 99.5)	0.53	100.00	0.26	100.00
67.0	21.3 (13.1 to 31.9)	99.5 (98.7 to 99.8)	0.39	99.99	0.19	100.00
79.0	9.8 (4.4 to 18.7)	100.0 (99.6 to 100.0)	100.00	99.99	100.00	100.00

cholestatic hepatobiliary disease, of whom 106 are detected by the test, and 10 babies with extrahepatic biliary atresia of whom seven are detected. If the prevalence of cholestatic hepatobiliary disease is 1 in 2500 and that of extrahepatic biliary atresia 1 in 20 000, these figures fall to 21 out of 40 babies with cholestatic hepatobiliary disease detected and three out of five babies with extrahepatic biliary atresia detected. It is unlikely that recall rates of 2% or higher would gain acceptance and setting the specificity higher leads to a rapid decline in sensitivity.

The low specificity of the test is due to high blood concentrations of bile acids in a small proportion of the comparison infants. These may have been premature or small for gestational age infants. Although gestational age has little effect on blood bile acid concentrations in the first week of life, concentrations may be significantly higher in premature infants at 14 days.²⁰⁻²¹ Babies who are small for gestational age can at 8 days have bile acid concentrations as high as our cholestatic group.²¹ Low birthweight infants have particularly high blood bile acid concentrations when they are ill—for example, with asphyxia, respiratory distress, or sepsis. Exclusion of low birthweight infants may produce a screening test with improved specificity.

Possible improvements

One reason for the low sensitivity of the test for detection of cholestatic hepatobiliary disease was that the

subgroup with idiopathic neonatal hepatitis included a high proportion of infants whose bile acid concentrations fell within the normal range. The data relating blood bile acid concentrations to time of sampling suggested that, in infants with idiopathic neonatal hepatitis, bile acid concentrations are rising in the first 2 weeks and sensitivity would be improved if all blood spots were obtained at 10-14 days. Another determinant of blood bile acid concentrations that we could not correct for is the rise occurring 30 minutes to 2 hours after a feed. If all blood spot samples could be taken more than 2 hours 30 minutes after a feed the variation in values in all subject groups would be reduced, and this would probably reduce the overlap and improve sensitivity and specificity.

Summary

In summary, this study has shown that tandem mass spectrometry analysis of blood spot bile acids cannot be used as a screening test on its own with a simple cut off value (or set of values) applicable to all neonates. We have indicated some ways in which sensitivity and specificity might be improved. Another way of improving performance of the screening test would be to assay a second compound in the spot. Conjugated bilirubin can be detected by tandem mass spectrometry in blood spots but is susceptible to photodegradation, suggesting that this is unlikely to be a useful option unless blood spots are protected from the light. Cholestasis leads to a rise in plasma phospholipid concentrations, and phospholipid profiles can be generated from a blood spot by tandem mass spectrometry (unpublished observations). It is unclear, however, whether such profiles will provide the required differentiation between cholestatic hepatobiliary disease and unaffected children. Finally, consideration could be given to combining bile acid analysis with clinical detection of jaundice. Until this has been evaluated, clinicians seeing children with jaundice in the third week of life should ensure that blood or urine, or both, are tested for conjugated bilirubin and that children with positive results are promptly referred to a specialist centre.

We thank Professor G Mieli-Vergani for her collaboration and the staff of the screening centres who provided us with samples.

Contributors: IM (guarantor) reviewed the literature, performed the validation experiments, sought ethical approval and funding, collected the Guthrie spots, performed the mass spectrometry analyses, reviewed the case notes, analysed the data, formulated the conclusions, and wrote the first draft of the paper. PTC (guarantor) had the original idea for the study, supervised the overall conduct of the project and interpretation of the results and revised the manuscript. SL (guarantor) provided advice on the selection of cases and the comparison

Key messages

- The prognosis of cholestatic hepatobiliary disease in infancy, in particular biliary atresia, is improved by early detection
- Infants destined to present with cholestatic jaundice in the first few months of life have raised concentrations of bile acids in the blood spots obtained at 7-10 days for current neonatal screening programmes
- Tandem mass spectrometry can be used to detect this marker of neonatal cholestasis
- 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
- Tandem mass spectrometry is a powerful tool for neonatal screening but every potential application must be carefully assessed

group and on analysis and interpretation of the data. AMW supervised the statistical analyses. AWJ helped to set up the mass spectrometry method for bile acid analysis. MM advised on optimum configurations for mass spectrometry analyses with the Quattro II mass spectrometer. DK contributed to the planning of the study and discussion of results.

Funding: Children's Liver Disease Foundation; Mr David Drake, consultant paediatric surgeon (personal donation); EPSRC (Link scheme).

Competing interests: None declared.

- Winfield CR, MacFaul R. Clinical study of prolonged jaundice in breast and bottle fed babies. *Arch Dis Child* 1978;53:506-7.
- Kelly DA, Stanton A. Jaundice in babies: implications for community screening for biliary atresia. *BMJ* 1995;310:1172-3.
- Dick MC, Mowat AP. Hepatitis syndrome in infancy—an epidemiological survey with 10 year follow-up. *Arch Dis Child* 1985;60:512-6.
- Mathew PM, Wharton BA. Investigation and management of neonatal jaundice, a problem orientated case record. *Arch Dis Child* 1981;56:949-53.
- Mowat AP, Davidson LL, Dick MC. Earlier identification of biliary atresia and hepatobiliary disease: selective screening in the third week of life. *Arch Dis Child* 1995;72:90-2.
- Mieli-Vergani G, Howard ER, Portmann B, Mowat AP. Late referral for biliary atresia: missed opportunities for effective surgery. *Lancet* 1989;i:421-3.
- DeConti RW, Craver RD, Willis GW, Hill CB, Hayes DH, Arensman RM. Extrahepatic biliary atresia from diagnosis to liver transplantation. *Pediatr Surg Int* 1992;7:737-40.
- Ohii R, Nio M, Chiba T, Endo N, Goto M, Ibrahim M. Long-term follow-up after surgery for patients with biliary atresia. *J Pediatr Surg* 1990;25:442-5.
- McNinch AW, Tripp JH. Haemorrhagic disease of the newborn in the British Isles: two year prospective study. *BMJ* 1991;303:1105-9.
- Logan S, Stanton A. Screening for biliary atresia. *Lancet* 1993;342:256.
- MacKiernan PJ. Biliary atresia. *British Paediatric Surveillance Unit Quarterly Bulletin* 1995;4:3.
- Rashed MS, Bucknall MP, Little D, Awad A, Jacob M, Alamoudi M, et al. Screening blood spots for inborn errors of metabolism by electrospray tandem mass spectrometry with a microplate batch process and a computer algorithm for automated flagging of abnormal profiles. *Clin Chem* 1997;43:1129-41.
- Mills K, Mushtaq I, Johnson AW, Whitfield P, Clayton PT. A method for the quantitation of conjugated bile acids in dried blood spots using electrospray ionization-mass spectrometry. *Pediatr Res* 1998;43:361-8.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Diagnostic Radiol* 1982;143:29-36.
- Hennekens CH, Buring JE. *Epidemiology in medicine*. 1st ed. Boston: Little, Brown, 1987.
- Tazawa Y, Yamada M, Nakagawa M, Konno T, Tada K. Serum bile acid patterns determined by an enzymatic method and high-performance liquid chromatography in young infants with cholestasis. *J Pediatr Gastroenterol Nutr* 1984;3:394-401.
- Javitt NB, Keating JP, Grand RJ, Harris RC. Serum bile acid patterns in neonatal hepatitis and extrahepatic biliary atresia. *J Pediatr* 1977;90:736-9.
- Matsui A, Fujimoto T, Takazawa Y, Okaniwa M, Kamoshita S. Serum bile acid levels in patients with extrahepatic biliary atresia and neonatal hepatitis during the first 10 days of life. *J Pediatr* 1985;107:255-7.
- Matsui A, Sasaki N, Arakawa Y, Ishikawa T, Momoya T, Kasano T, et al. Neonatal mass screening for biliary atresia. *Screening* 1993;2:201-9.
- Barnes S, Berkowitz G, Hirschowitz BI, Wirtschafter, Cassidy G. Postnatal physiologic hypercholaemia in both premature and full-term infants. *J Clin Invest* 1981;68:775-82.
- Senger H, Boehm G, Beyreiss K, Braun W. Bile acids in the serum and duodenal content of newborn infants of different classification. *Biomed Biochim Acta* 1986;45:931-9.

(Accepted 12 May 1999)

Commentary: What is tandem mass spectrometry?

Tandem mass spectrometry is a technique that has been developed to analyse the metabolites and proteins present in samples of blood. In this paper, tandem mass spectrometry has been used to analyse the blood "spots" (heel prick samples of blood dried on filter paper) that are typically taken shortly after birth for the Guthrie test. But rather than testing the blood for the presence of just one compound (phenylalanine, in the case of phenylketonuria), tandem mass spectrometry can simultaneously examine a large number of materials from a single sample.

Mass spectrometers measure the mass of substances (molecular weight). Tandem mass spectrometry entails the use of two mass spectrometers in sequence (hence "tandem"): the first spectrometer is used to separate compounds of a single molecular mass (precursor ions) from a nebulised blood sample. The separated species are subsequently passed through a "collision cell," where the molecules are bombarded with high energy argon gas. The whole molecules are broken up by the argon, and the fragments that result (product ions) are then passed into the second spectrometer for analysis of their molecular weights. As different classes of compounds fragment in uniquely different ways, if the mass of a molecule and its common fragments are known it is possible to infer the identity of the molecule.

Tandem mass spectrometry can therefore separate a nebulised mixture into different classes of molecules and in a single, two minute experiment it is possible to generate profiles of several different classes of molecules that may be useful in determining disease states. Analysis of acylcarnitines, amino acids, and bile acids from a single (3 mm) blood spot have been well

documented (Sweetman L, Chase DH, et al. *Clinical Chemistry* 1996;42:345-6; Sweetman L, Chase DH, et al. *Clinical Chemistry* 1996;42:349-55).

It remains unclear whether tandem mass spectrometry will ultimately be used as a tool for neonatal screening (looking for evidence of phenylketonuria, fatty acid oxidation defects, and organic acidaemias, for example) or whether it will be reserved for specific investigative testing for children who are clearly unwell. Although all of these conditions are rare, with an incidence of just 1 in 3000-4000, some of them (such as medium chain acyl-CoA dehydrogenase deficiency and methylmalonic acidaemia) are eminently treatable and if left undiagnosed have potentially devastating consequences (Wilson CJ, et al. *Archives of Disease in Childhood* 1999;80:459-62; Leonard JV. *Journal of Inherited Metabolic Disease* 1995;18:430-4).

Over 50 centres worldwide are now working with tandem mass spectrometry, although only three of these are up and running a full screening service for neonates. In the United Kingdom, six centres are involved in running investigative trials.

Abi Berger *science editor, BMJ*

Correction

Comparative efficacy of three mumps vaccines during disease outbreak in eastern Switzerland: cohort study

An error occurred in this short report by Schlegel et al (7 August, pp 352-3). The first column of the first row of the table should have read: "No vaccination" [not "No (%) vaccinated" as published].

