- 8 Goodin DS, Aminoff MJ. Does the interictal EEG have a role in the diagnosis of epilepsy. *Lancet* 1984;1:837-9.
- 9 Benbadis SR, Tatum WO. Overinterpretation of EEGs and misdiagnosis of epilepsy. *J Clin Neurophys* 2003;20:42-4.
- 10 NICE. *The epilepsies: the diagnosis and management of the epilepsies in adults and children in primary and secondary care*, NICE, 2004 (www.nice.org.uk/pdf/CG020NICEguideline.pdf).
- 11 Shinnar S, Berg AT. Does antiepileptic drug therapy prevent the development of chronic epilepsy? *Epilepsia* 1996;37:701-8.
- 12 Téllez-Zenteno JF, Ronquillo LH, Wiebe S. Sudden unexpected death in epilepsy: evidence based analysis of incidence and risk factors. *Epilepsy Research* 2005;65:101-15.
- 13 Royal College of Physicians of Edinburgh. Consensus conference on better care for children and adults with epilepsy. Royal College of Physicians of Edinburgh, 2002 (www.rcpe.ac.uk/education/standards/consensus/better_care_02.php).

Screening

Newborn screening for cystic fibrosis: do we need a second IRT?

J F Price

Commentary on the paper by Massie *et al* (see page 222)

Newborn screening for cystic fibrosis (CF) has been possible since 1978, but national screening programmes have proved highly contentious and as yet have only been adopted in a minority of European countries and North American states. The main issue has been whether early diagnosis through screening results in long term clinical advantage, particularly in lung function. Recently the balance of benefit against harm has tipped in favour of screening,¹ providing that early diagnosis is followed by high quality care in specialised centres.² Screening programmes have been introduced in Northern Ireland, Wales, and Scotland. Only about 20% of babies are screened for CF in England, but a national screening programme is due to be implemented over the next three years. A good screening programme should identify the maximum number of cases that are likely to be severely affected, ensure that as many as possible of the missed cases are mildly affected, detect as few unaffected carriers as possible, allow for ethnic diversity, and generate a minimum of anxiety. Three options are available; IRT-IRT, IRT-DNA, and IRT-DNA-IRT.

IRT-IRT involves measurement of immunoreactive trypsinogen (IRT) during the first week on the Guthrie blood spot and repeating the measurement at 3-4 weeks in those with initial high levels. This protocol for screening was introduced in East Anglia in 1979. The sensitivity of a raised 3-5 day IRT is high, but the positive predictive value is low. Because blood levels of IRT decay slowly in CF infants, a second IRT at 3-4 weeks increases the specificity, but about 1 in 200 newborn infants progress to the second blood test. This generates a period of anxiety in a large number of

families who will turn out to have unaffected children.

IRT-DNA employs DNA analysis instead of a second IRT at 3-4 weeks. Infants with very high IRT in the first week undergo DNA analysis and those with at least one mutation have a sweat test. The advantage of IRT-DNA is that both tests can be done on the initial blood spot sample. New South Wales was one of the first regions to adopt an IRT-DNA screening programme, setting the cut-off for the first IRT at >99th centile and testing for the most common mutation, $\Delta F508$. In this issue, Massie *et al* describe their experience of IRT-DNA screening in New South Wales between 1991 and 2003 and make the case against the measurement of a second IRT in babies with a high first IRT but no $\Delta F508$ mutation.³ During this time the screening programme identified 209 children with CF. Eighteen cases were missed, of whom seven were identified in the newborn period in other ways (meconium ileus or siblings with CF). The positive predictive value for an elevated IRT and one $\Delta F508$ mutation was 1:8 (eight sweat tests were required to detect one case). However the yield for detecting additional cases by a second IRT in children who had a high first IRT but no $\Delta F508$ mutation was extremely low. A second IRT was needed in about 400 infants to detect one more child with CF. To detect one "unexpected case" that was missed by $\Delta F508$ screening and not diagnosed in the newborn period by other routes, a second IRT was required in nearly 700 infants.

IRT-DNA screening was commenced in Wales in 1996. At first DNA analysis was done on four common mutations ($\Delta F508$, G551D, G542X, R553X), plus one with a high local frequency in Wales (1898+1G>A), but since 1999 a panel of

31 mutations has been used. By testing for more than one mutation the false negative rate was reduced to 3% and just six sweat tests were needed to detect one case.⁴ The merits of this need to be weighed against the additional complexity and cost of the multiple DNA test, the detection of more carriers, and the identification of "mild" genotypes. Although the Welsh experience indicates that carrier identification is not perceived to be a problem,⁵ the counselling incurs additional cost and the impact of disclosure of a carrier state on the child growing up has not been assessed.

The first two stages of the IRT-DNA-IRT protocol are the same as for IRT-DNA and sweat tests are done on babies with two detected mutations. However, a second IRT is measured at 3-4 weeks in babies with one mutation and a first IRT elevated above the 99.5th centile and in babies with no mutation but a first IRT extremely elevated above the 99.9th centile. Confirmatory sweat tests are done on those with an elevated second IRT. An IRT-DNA-IRT protocol with DNA analysis for 31 mutations has been used in Scotland since 2002 and will be adopted throughout England over the next three years. A refinement built into the new protocol for England will be a two stage DNA analysis. The four most common mutations will be tested first and a further 27 mutations will be analysed only in those with one common mutation. The purpose of this is to reduce the detection of carriers with mild mutations.

Applying existing UK data to the IRT-DNA-IRT protocol, it has been estimated that for those with one mutation a second IRT will be required in about 11 infants to detect one with CF, and in those with no mutations a second IRT will be needed in about 30 infants to detect one additional case. The critical assumption for those with no mutations is that a cut-off at 99.9% rather than 99% for the first IRT will give a better yield for the second IRT. The New South Wales experience does not entirely contradict this assumption. Their data show that the positive predictive value for a second IRT is 1:397 when the cut-off for the first IRT is set at 99%, but this improves to 1:59 if the cut-off is raised to 99.9%. Eight of their 11 "unexpected undiagnosed" cases had an initial IRT >99.9%. Another factor that may enhance the value of a second IRT in a

UK screening programme is the diversity of the population. In the UK about 4% of children diagnosed with CF are non-Caucasian in origin. A DNA panel comprising the most common 31 CF mutations will detect 97% of mutations in a Caucasian population, but only just over 65% of mutations in a non-Caucasian population. The gene frequency of $\Delta F508$ in the UK Indian subcontinent CF population is less than half that in the UK Caucasian CF population.⁶ Thus a three stage IRT-DNA-IRT protocol is likely to increase the chances of detecting CF in non-Caucasian infants. The main disadvantages of IRT-DNA-IRT are its complexity and the anxiety generated for families who have to wait for the result of a second IRT. Babies with one mutation and a normal second IRT still carry a small risk of having CF. The precise "unexpected false negative rate" and the true value of a second IRT using the

IRT-DNA-IRT approach in the UK, can only be determined by careful prospective monitoring.

There is no "ideal" protocol for CF screening and variations will be influenced by local circumstances. A recent study from France suggests that it may yet be possible to screen with acceptable sensitivity and specificity without the use of genetic tests.⁷ As screening becomes more widespread, each new screening programme should be monitored carefully and reviewed. The meticulous recording of data over many years in New South Wales sets an excellent example.

Arch Dis Child 2006;**91**:209–210.
doi: 10.1136/adc.2005.085084

Correspondence to: Prof. J F Price, Paediatric Respiratory Medicine, King's College School of Medicine, King's College Hospital, Denmark Hill, London SE5 8RX, UK; john.price@kcl.ac.uk

Competing interests: none declared

REFERENCES

- 1 Sims EJ, McCormick J, Mehta G, *et al*. Newborn screening for cystic fibrosis is associated with reduced treatment intensity. *J Pediatr* 2005;**147**:306–11.
- 2 Doring G, Holby N. Consensus Study Group. Early intervention and prevention of lung disease in cystic fibrosis: a European consensus. *J Cystic Fibrosis* 2004;**3**:67–91.
- 3 Massie J, Curnow I, Tzanakos N, *et al*. Markedly elevated neonatal immunoreactive trypsinogen levels in the absence of cystic fibrosis mutations is not an indication for further testing. *Arch Dis Child* 2006;**91**:222–5.
- 4 Doull I, Hall S, Bradley D. Newborn screening for cystic fibrosis. The Welsh Experience, 1996–2004. Personal communication to the UK Newborn Screening Program Centre, 2004.
- 5 Parsons EP, Clarke AJ, Bradley DM. Implications of carrier identification in newborn screening for cystic fibrosis. *Arch Dis Child Fetal Neonatal Ed* 2003;**88**:F467–71.
- 6 McCormick J, Green M, Mehta G, *et al*. Demographics of the UK cystic fibrosis population: implications for neonatal screening. *Eur J Hum Genet* 2002;**10**:583–90.
- 7 Sarles J, Berthezene P, Le Louarn C, Somma C, *et al*. Combining immunoreactive trypsinogen and pancreatitis-associated protein assays, a method of newborn screening for cystic fibrosis that avoids DNA analysis. *J Pediatr* 2005;**147**:302–5.

Diagnosis

Can we predict vesicoureteric reflux?

J H Baumer

Commentary on the paper by Leroy *et al* (see page 241)

Many children are investigated for vesicoureteric reflux (VUR) following a urinary tract infection (UTI), but the imaging technique of a voiding cystourethrogram (VCUG) is unpleasant for the child and not without risk. A clinical scoring system capable of confidently predicting VUR would therefore be an attractive alternative approach.

The original study by Oosterbrink and colleagues¹ seemed to suggest that a score derived from the combination of clinical factors (age, sex, and positive family history), ultrasound findings and C reactive protein (CRP) result could predict the presence of VUR with high sensitivity, albeit with rather low specificity. A diagnostic test with 100% sensitivity can be useful in ruling out a condition when it is negative, even when it has low specificity (that is, cannot be relied upon when positive). This has been referred to as a "SnNout" (when a test has a high Sensitivity, a Negative result rules out the diagnosis).² However, the lower the specificity, the smaller would be the expected proportion of children without reflux

who have a negative result, and thereby avoid an unnecessary imaging investigation. A negative result, excluding VUR, was seen in 17% of children without VUR in Oosterbrink's series.

Leroy and colleagues³ are to be commended on repeating the work of Oosterbrink. When a group of patients is used to develop a diagnostic test, the performance of that test tends to be overestimated. This is particularly true when the test is composed of a number of elements, and when continuous variables (such as age, CRP) are categorised and weighted in such a way as to produce the best discrimination. A second, validation study should therefore be conducted in an independent group of patients.

Leroy *et al* found a lower specificity (3% versus 17%) for all grades of VUR using a threshold score of 0, and lower sensitivity (93% versus 100%) and specificity (13% versus 38%) for grade 3 to 5 VUR using a threshold score of 5. With a second study showing a less good performance from the decision rule, what are we to conclude? The first question is how well the second study

was conducted. A key issue is whether an independent reference or "gold standard" test was applied blind to all patients (that is, whether those interpreting the gold standard test did so unaware of the result of the clinical decision rule and vice versa). In this case the reference test was the VCUG, with the radiologist reporting the VCUG being unaware of the VUR risk score. It is not clear whether all the components of the VUR score were obtained blind to the VCUG result. The CRP result and age and gender are objective measures, and as long as the family history was obtained before the parents knew the VCUG result, these were all very unlikely to have been influenced by the latter. If the ultrasound examinations were sometimes conducted after the VCUG, however, knowledge of the latter could have influenced the interpretation of the former.

Another issue is what proportion of eligible children had all necessary information available. Ten children who did not have a VCUG performed were excluded from the study. A further six children with inadequate clinical information were excluded from analysis. All other children had both the relevant clinical information and test results available. This is a relatively small proportion (around 10%) with missing information.

The study therefore appears to have been conducted satisfactorily, apart from the one or two uncertainties about the independence of the tests. If the tests were not performed fully independently, the study could be expected to