Appendix B

Economic evaluation of alternative testing strategies in the detection of hyperbilirubinaemia

Introduction

Jaundice (a yellow colouration of the skin) is caused by hyperbilirubinaemia and is common in the newborn baby. Rarely, if bilirubin levels are sufficiently high, bilirubin can cross the blood barrier and cause a brain damaging condition called kernicterus. Kernicterus is a lifelong disabling neurological problem with manifestations of cerebral palsy and deafness with high costs of care. Hyperbilirubinaemia can also cause deafness without cerebral palsy and other adverse outcomes have been described. Levels of bilirubin can be controlled with phototherapy, but the only way to reduce very high levels in an emergency is with an exchange transfusion. This is a costly intensive care procedure which has a risk and carries a mortality {Davidson, Thilo How to make kernicterus a never event, *NeoReviews* 2000 insert ref ID]. Phototherapy is generally effective in controlling bilirubin levels, preventing them from rising to a level at which kernicterus occurs, hence some clinicians have called for kernicterus to be classified as a "never event". There is evidence that cases of kernicterus have risen recently due to inefficient detection of cases of jaundice, probably mainly as a result of earlier discharge following childbirth.

Current practice in England and Wales is varied but the GDG estimate that less than 10% of babies undergo specific testing of their bilirubin levels following visual examination. At present, babies who develop kernicterus often present late and with bilirubin levels already in the toxic range. The key to prevention of kernicterus is early detection of cases at a time when phototherapy can be effective. Any guideline recommendation which requires more widespread testing will have important resource implications for the NHS as well as requiring a change in practice in many places. Therefore the guideline recommendation regarding identification of cases by testing for hyperbilirubinaemia was highlighted by the GDG as an important priority for economic analysis. The NHS operates within resource constraints and a more intensive testing and treatment strategy can only be justified if it represents a better use of scarce resources than could be obtained in some alternative use of those resources.

Background to the economic evaluation

Kernicterus is a largely preventable disease if severe hyperbilirubinaemia is identified early and promptly treated (using phototherapy or, for more acute cases, exchange transfusion). Therefore, early identification of raised (or rapidly rising) bilirubin levels is the key to reducing severe morbidity.

There are studies which demonstrate that more intensive monitoring reduces the need for exchange transfusions. Evidence from the United States reports that during the 1970s, kernicterus was practically eradicated, which was probably due to the liberal use of phototherapy. The disease re-emerged in the 1990s, largely among babies cared for in the home environment in the neonatal period often with limited medical supervision during the first week after birth {Bhutani and Johnson, 2009 kernicterus in the 21st century REF ID]. Kernicterus has fallen again in the US since the adoption of the 1994 AAP guidelines; estimates are that the rate has fallen from 5.1 per 100,000 in 1988 to 1.5 per 100,000 {Burke, Robbins et al, trends in hospitalizations for neonatal jaundice, Paediatrics 2009, REF ID}.

In the UK, babies are discharged earlier and are monitored less often than in previous decades. Reduced contact with experienced midwives and reliance on intermittent visual examination to assess bilirubin levels may be one of the reasons for the failure to detect babies with significantly elevated serum bilirubin levels. A newborn baby might only be visited once by a midwife in the post natal period if there are no risk factors, although the norm is currently around two or three visits in the first week. Visual examination by a midwife to assess for jaundice during these post-natal visits is currently the standard of care, with a small proportion of these jaundiced children subjected to a total serum bilirubin blood test (TSB) based on clinical visual assessment of the level of bilirubin. This is known to be unreliable. There is strong evidence that visual examination alone cannot be used to assess the level of bilirubin in a baby. The inaccuracy of visual assessment for the detection of bilirubin levels, particularly in babies with dark skin tones, is likely to be a major factor responsible for the late presentation of babies with significant hyperbilirubinaemia. A more reliable strategy for the detection of babies who require treatment with phototherapy is undoubtedly required.

The cost of care of people with kernicterus throughout their lives is millions of pounds. If resources were invested in a testing strategy that was effective in reducing the number of cases of kernicterus annually by even one case per year, it would be cost saving if the total annual cost of the strategy was less than the lifetime cost of caring for one individual with the disease. Since kernicterus is a lifetime condition with extremely poor quality of life, the value that the NHS places on preventing a case of kernicterus is not only calculated as the cost saved by preventing the downstream costs but also the £20,000 per QALY over the lifetime of the condition. Clearly, if the intervention was more successful in preventing

kernicterus, then more NHS resources could be used to identify hyperbilirubinaemia and still be cost-effective.

It seems plausible that a more intensive testing strategy could be clinically effective by overcoming the limitations of visual examination alone, thereby leading to better detection and treatment. Currently, there are two methods of testing; a total serum bilirubin blood test (TSB) and a transcutaneous bilirubinometer (TCB) which is a non-invasive test on the surface of the skin. TCB is not accurate above a threshold level of 250 micromol/L of bilirubin so that TSB testing is required in babies whose TCB is above this threshold level. Hence a strategy involving more bilirubin measurements could be based on TSB alone or TCB with TSB for those babies whose TCB level was higher than the threshold value. Current evidence does not favour one strategy over the other for the detection of babies with bilirubin levels over 250 micromols/L. That is, even though TSB is the gold standard test, both strategies when used correctly as part of a planned monitoring protocol to test babies who are visibly jaundiced, would be equally effective at detecting hyperbilirubinaemia and preventing kernicterus. Both methods are in use in the NHS. The TSB can be analysed in hospital labs without the need for additional equipment. The TCB requires the purchase of hand held devices, sufficient for one to be available for each community midwife undertaking post natal visits on any particular day.

The economic evaluation was undertaken to determine the conditions under which increased testing would be cost-effective, and to explore which testing strategy would be cost-effective under different circumstances.

Method

In this analysis we evaluate the cost-effectiveness of moving from current practice to a more intensive test strategy in England and Wales subject to the limitations of the published evidence.

The following strategies are compared:

1. "Current practice"
    - Estimated as visual examination followed by TSB in 10% of visually jaundiced babies
2. TSB
    - A TSB on all babies with a positive visual examination
3. TCB followed by TSB if positive TSB
    - A TCB on all babies with a positive visual examination, with a TSB on those with a positive TCB

Visual examination has a high negative predictive value which means that babies who do not appear visually jaundiced are very unlikely to have clinically significant jaundice. However, visual examination has been shown to be unreliable in detecting the severity of hyperbilirubinaemia. Therefore, visual examination alone as a basis for detecting jaundice requiring phototherapy has poor sensitivity which may put jaundiced babies at a higher risk of developing kernicterus.

Detection of hyperbilirubinaemia requiring treatment or further monitoring can be better assessed using a transcutaneous bilirubin test (TCB) or a blood test to measure the total serum bilirubin levels (TSB). The TCB is done with a handheld device (e.g. Minolta JM103 or Bilicheck) which is simple to use and is placed on the baby's skin. The TSB is the gold standard test but is more invasive and distressing to the baby since it requires a blood sample. Both tests can be carried out by the midwife during the home visit.

Diagnostic tests are usually evaluated according to their sensitivity and specificity and these characteristics can be used to generate probabilities in decision analytic models. Initially, we intended to compare the alternative strategies using such an approach. However, the decision making process in this context is far more complicated than that implied by the outcomes for a "two by two table". Rather than the test result dividing the patient population neatly into positives and negatives, different test thresholds are used to stratify patients into groups requiring immediate treatment, further monitoring or transfer back to routine care. Decision making is affected implicitly in a Bayesian manner by the impact of the bilirubin level on the post-test probability of disease. The decision making is complicated further as a number of other factors, such as family history of jaundice, will also be taken into account. Furthermore, monitoring can occur at many points in time and this temporal aspect is important because thresholds for clinically significant jaundice change and the evidence base to track changes in diagnostic accuracy over the relevant time periods is lacking. Therefore, it was ultimately decided that there was not sufficient published evidence to populate such a decision model. Furthermore, it was felt that the GDG would not be able to estimate the vast array of model parameters to reflect the actual micro decision making process that occurs in actual clinical practice.

However, the GDG have set a relatively high bilirubin threshold as a basis for treatment and a relatively low bilirubin threshold for further monitoring. The rationale for this is to avoid unnecessary phototherapy (i.e. a high specificity or false positive rate in terms of treatment) whilst avoiding missed cases by continued monitoring in babies who have an intermediate bilirubin level (i.e. a high sensitivity or false negative rate in terms of monitoring). Whilst, the TCB is not thought to be reliable at high bilirubin levels (hence the need for TSB if TCB is positive) it is nevertheless thought to be accurate at the more intermediate levels.

The GDG opinion is that, using the thresholds defined in this guideline, either method of testing would be effective in detecting hyperbilirubinaemia and avoiding new cases of kernicterus. Therefore, the cost-effective strategy was estimated using a cost minimisation

approach which assumes no difference in effectiveness between testing strategies. As noted earlier, there is insufficient evidence to estimate the incremental benefit of moving from "current practice" to a more intensive testing regime, although evidence on the limitations of visual examination suggests that some benefit is likely. Therefore, a threshold analysis was undertaken to determine the number of kernicterus cases that a more intensive testing approach would have to avert in order for this to be considered cost-effective.

Model parameters and assumptions

The cost analysis was undertaken from the perspective of the NHS and personal social services which is in accordance with the NICE guidelines manual (2009). The costs were estimated using a bottom-up or "ingredient's" approach which involves detailing the physical quantity of resources used in providing treatment alongside the unit cost of those resources. From this it is possible to estimate the total cost of treatment.

It was assumed that visual examination is undertaken in the first instance in all strategies. In the "current practice" strategy it was additionally assumed that visual examination is used to determine the severity of hyperbilirubinaemia with a proportion of these having a TSB blood test as a precursor to possible phototherapy. No cost has been applied to the visual examination as it is assumed this would occur as part of the standard home visit carried out by a midwife.

The population characteristics for this analysis are shown in Table 1. Economic parameters used in the assessment of cost benefit (other than the costs of the test strategies) are shown in Table 2.

Table 1 Population characteristics

Item	Value	Source	Notes
Births	690,000	ONS (2008)	Based on 2007 births
Babies identified as jaundiced on visual exam	60%	GDG	

Babies currently tested for jaundice on basis of visual exam	10%	GDG	Testing is assumed to be with TSB
Mean number of tests per baby tested[1]	1.33	Kuzniewicz et al. (2008)	Pre-universal screening this US study estimated 0.8 tests per baby. We assumed that this was based on 60% of babies being tested, on the basis of the GDG's estimate of babies identified as visually jaundiced in the UK. A weighted average was then used to estimate the tests per baby in the 60% tested that would give an average of 0.8 tests per baby overall. In addition the GDG considered this a reasonable estimate

Table 2 Cost benefit parameters

Item	Value	Source	
Kernicterus case	£5.5 million	JMW Clinical Negligence Solicitors	
Discount rate	3.5%	NICE Guidelines Manual (2009)	Both costs and QALYs are discounted
QALY gained per kernicterus case avoided	25	Calculation	This is an approximation, based on an assumption that the quality of life with kernicterus is not much better than death
Willingness to pay for a QALY	£20,000	NICE Guidelines Manual (2009)	An advisory threshold

The resource 'ingredients' and their unit costs for TSB and TCB are shown in Table 3 and Table 4 respectively. The resource items include any the additional staff time required to undertake a test as part of a routine post natal visit. It also includes equipment costs and consumables, those resources that are used up in the provision of the test that cannot be reused.

[1] A TCB positive followed by a TSB is considered as a single test for the purposes of this analysis

Table 3 TSB resources and costs

Resources	Unit cost	Source	Notes
Midwife Band 6/7	£56.00	PSSRU (2006)	It is assumed that a midwife would take 10 minutes to undertake this test
Venous blood test	£7.00		One per test
Gloves	£0.06	medisave.co.uk accessed 16 July 2009	£6.27 per 100 One pair per test

Table 3 TCB resources and costs

Resources	Unit cost	Source	Notes
Midwife Band 6/7	£56.00	PSSRU (2006)	It is assumed that a midwife would take 1 minute to undertake this test
TCB meter	£3400 £3600	Manufacturer, JM103 Manufacturer, Bilicheck	No consumables required
Calibration tips	£2.70	Manufacturer, Bilichek	
TSB	£18.39	Marginal cost of TSB (see Table 2)	It is estimated that 25% of TCB tests would be positive leading to a TSB

The purchase of medical equipment, TCB meters in this case, carries an opportunity cost which differs from operating costs such as labour and consumables in certain respects. The purchase of TCB meters involves an upfront payment (or investment) before use. However, that cost is fixed as it does not vary with the quantity of treatment provided. The equipment can often be used over a number of years before it needs to be replaced.

The equipment costs have two facets:

- Opportunity cost – the money spent on the equipment could have been invested in some other venture yielding positive benefits. This is calculated by applying an interest rate to the sum invested in the equipment.

- Depreciation cost – the equipment has a certain lifespan and depreciates over time. Eventually, the equipment has to be replaced.

In economic evaluation, the usual practice is to annuitise the initial capital outlay over the expected life of the equipment to give an 'equivalent annual cost'. Calculating the equivalent annual cost means making an allowance for the differential timing of costs, using discounting.

The formula for calculating the equivalent annual cost is given below:

$$E = (K - [S \div \{1 + r\}n]) \div A(n, r)$$

where:

E = equivalent annual cost
K = purchase price of equipment
S = resale value
r = discount (interest rate)
n = equipment lifespan
A(n, r) = annuity factor (n years at interest rate r)

To calculate the equivalent annual cost we have assumed that the meters last 5 years and have no resale value. However, the total annual equivalent cost would depend on the actual number of meters that were necessary to deliver the strategy. This is not known and service delivery is not generally part of the remit of NICE guidelines. Therefore, the results are presented as a threshold analysis, with the threshold being the number of meters at which the TSB strategy (strategy 2) would be equivalent in cost to the TCB strategy (strategy 3).

Results

The marginal cost per test (excluding equipment) was estimated to be £18.39 for TSB and £5.81 for TCB for a test requiring no consumables and £8.51 for a test requiring calibration tips. Using these figures the cost per strategy was calculated as follows:

Calculation of total costs per annum of each strategy

1.		*Current practice*
	Population	690,000
	Tested with TSB	690,000 x 0.1 = 69,000
	TSB tests	69,000 x 1.33 = 91,770
	Cost	91,770 x £18.39 = **£1.68 million**
		Strategy 2 (TSB to all visually jaundiced babies)
	Population	690,000
	Visibly jaundiced	690,000 x 0.6 = 414,000
	TSB tests	414,000 x 1.33 = 550,620
	Cost	550,620 x £18.39 = **£10.13 million**
2.		*Strategy 3 (TCB to all visually jaundiced babies followed by TSB if TCB is positive)*
	Population	690,000
	Visibly jaundiced	690,000 x 0.6 = 414,000
	TCB tests	414,000 x 1.33 = 550,620
	TSB tests	550,620 x 0.25 = 137,655
	Cost (BiliCheck®)	(550,620 x £8.51) + (137,655 x £18.39)
		= **£7.22 million plus annual equivalent equipment cost**
	Cost (Minolta®)	(550,620 x £5.81) + (137,655 x £18.39)
		= **£5.69 million plus annual equivalent equipment cost**

Figure 1: A graph to compare the incremental costs of TCB with TSB varying the number of TCB meters needed to deliver the strategy

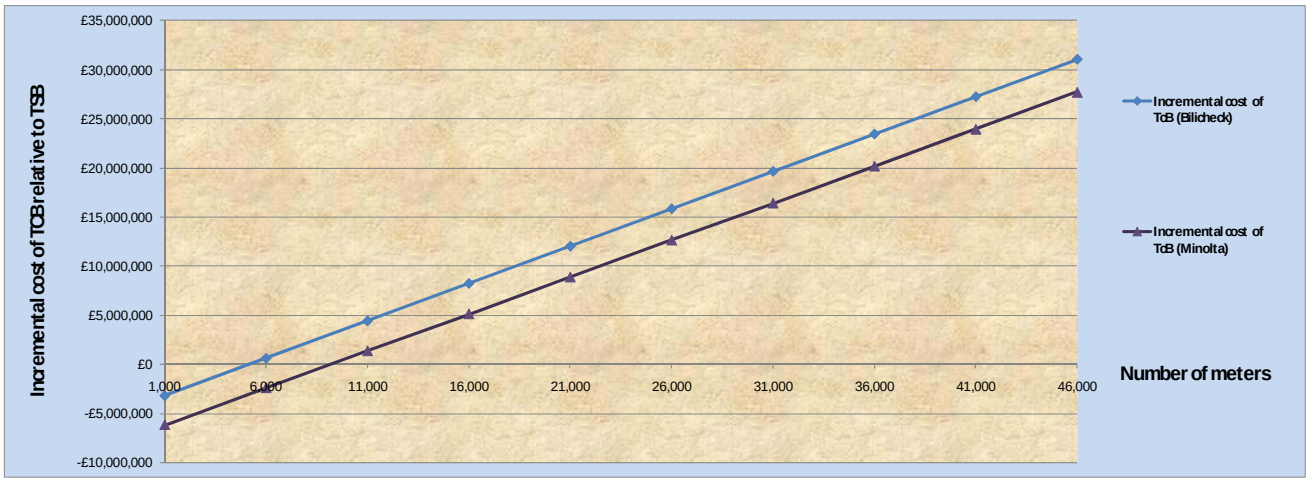
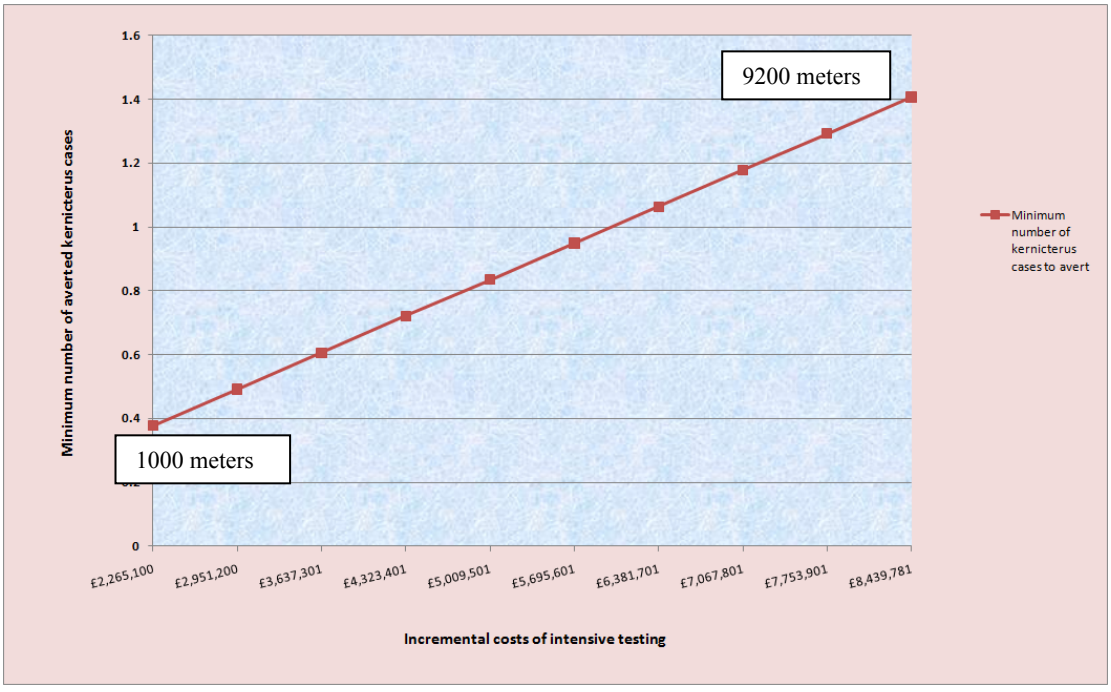


Figure 2: A graph to show the minimum number of kernicterus cases to be averted at different incremental costs of more intensive testing.

Figure 2 shows the total additional cost to the NHS of more intensive testing between a



minimum of 1000 meters and a maximum of 9,200. The cost on x- axis is the incremental cost difference between 'current practice' and more intensive testing. In this figure, the comparator with current practice is always TCB. If 9200 meters or fewer are purchased, the

cheaper option is always TCB. If more meters are required, then the TSB strategy should be preferred on cost-effectiveness grounds. The figure shows that the total cost of using 9200 meters would require an additional £8.45 million. The number of cases to be averted would have to be at least 1.4 for this to be cost-effective compared with current practice. Buying only an additional 1000 meters, the total additional cost would be £2.3 million and 0.4 cases of kernicterus would need to be averted for TCB to be more cost-effective than current practice.

Sensitivity Analysis

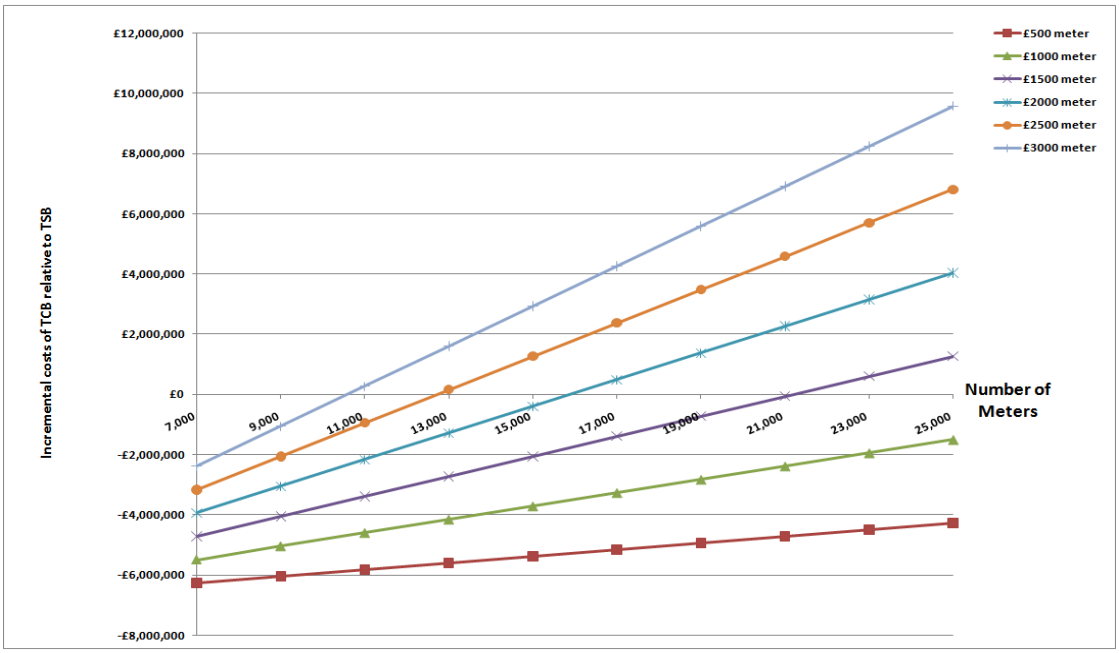
Sensitivity analysis is used in economic evaluation to assess how sensitive the results of the model are to the assumptions made about the model parameters, particularly those parameters where considerable uncertainty exists as to their actual value. One-way sensitivity analysis involves altering the value of a single parameter, holding all the others constant, to determine how sensitive the cost-effectiveness conclusion is to the assumptions made about that particular parameter.

The base-case results above were presented as threshold analyses reflecting uncertainty about the number of meters that would be needed for the TCB strategy and the number of kernicterus cases that would need to be averted in order for the additional costs of more intensive costing to be deemed an efficient use of scarce NHS resources. However, the sensitivity analyses below explore how changes in other model parameters would affect results.

i. Varying the cost of meters

In this sensitivity analysis the cost of the meters is varied between £500 and £3,000. It is assumed that the meter is a Minolta® and does not therefore require a new calibration tip for each test.

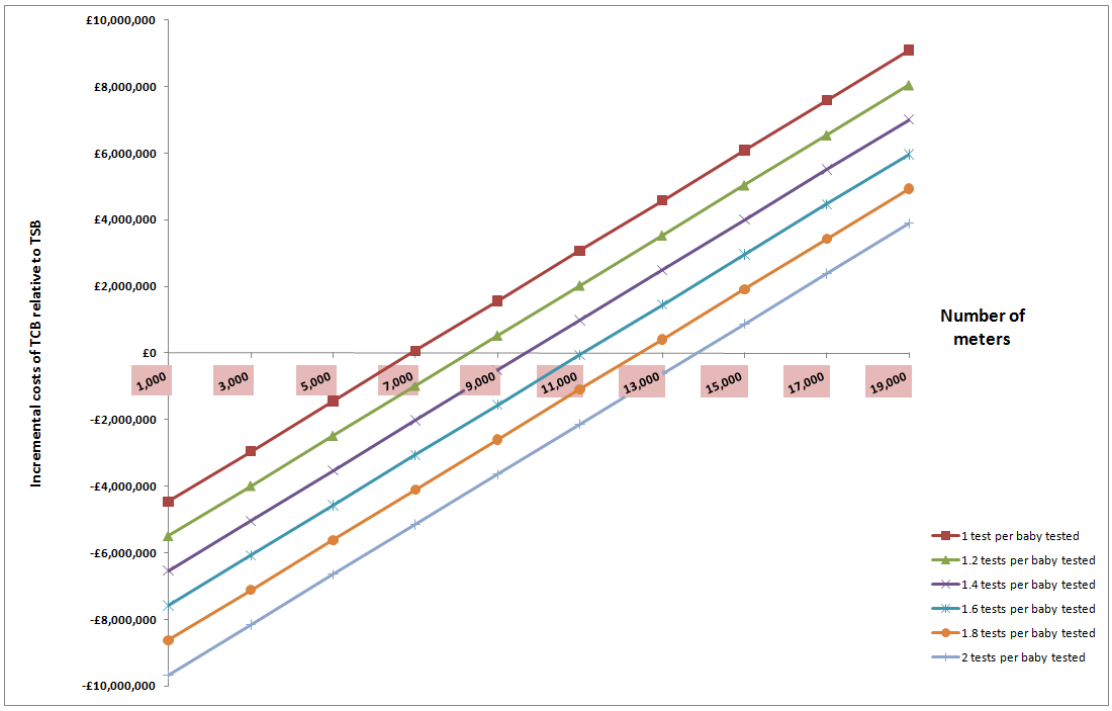
Figure 3: A graph to compare the incremental costs of TCB with TSB varying the number of TCB meters needed to deliver the strategy and the cost of the TCB meter



ii Varying the mean number of tests per baby tested

In this sensitivity analysis the mean number of tests per baby tested is varied between one and two tests.

Figure 4: A graph to compare the incremental costs of TCB with TSB varying the number of TCB meters needed to deliver the strategy and the mean number of tests per baby tested



Discussion

The analysis compared the current testing strategy with an uplift in testing using alternative strategies. In the base case analysis, the current strategy of testing only 10% of babies using TSB was £1.68 million per year. The next cheapest strategy was to use Strategy 3 (TCB to all visually jaundiced babies followed by TSB if TCB is positive) using a meter that does not require calibration tips which cost £5.73 million, or £7.22 million using a meter requiring a calibration tip. Using the TSB more intensively (on 60% babies who are visibly jaundiced) would cost £10.13 million per year. The cost difference between TSB and TCB is mainly due to the increased time estimated for a midwife to do a blood test compared with a skin test.

An important question is whether any change from current practice can be justified on cost-effectiveness grounds. In part this depends on the fixed costs, that is the number of TCB meters needed to deliver strategy 3. This determines the incremental costs of increased

testing if the TCB strategy is deemed more cost-effective than TSB, that is, the strategy with the lowest cost since this is a cost minimization analysis. In the base-case analysis, the results estimate that the maximum incremental cost of more intensive testing is around £8.45 million which is the incremental cost of an enhanced testing strategy using TSB alone relative to current practice. If the strategy using the TBC could be delivered with the purchase of only 1000 additional bilirubinometers (which would be a highly conservative estimate) then the incremental cost would be £2.2 million. Figure 2 suggests that 1.4 cases of kernicterus would have to be averted for more intensive testing to be considered cost-effective if the incremental testing costs were £8.45 million. If fewer resources were required (fewer bilirubinometers purchased) then fewer cases would need to be averted. This assumes a threshold QALY value of £20,000. At a higher threshold, say £30,000 per QALY, the number of cases of kernicterus averted in order for more intensive testing to be cost-effective would be fewer.

Figure 2 shows how this threshold of kernicterus cases that need to be averted for cost-effectiveness falls as the incremental costs of more intensive testing fall, as is the case with a smaller number of TCB meters. At this moment in time the evidence base is not sufficiently robust to assess whether more intensive testing would achieve such an incremental gain – there are approximately 6-7 new kernicterus cases per annum in England and Wales. However, given the evidence about the limitations of visual examination the GDG are opposed to relying on observations which have been demonstrated to be unreliable in the detection of severe hyperbilirubinaemia. It does seem plausible that a more intensive testing strategy using tests which are known to have greater reliability in detection of severe hyperbilirubinaemia would lead to more appropriate and timely intervention with a concomitant reduction in adverse outcomes.

The costs of the TCB testing strategy vary according to the cost of meter used. In the absence of evidence that health outcomes are different between types of meter used, the cheaper Minolta® meter should be preferred. The base-case results (see Figure 1) suggest that the Minolta® meter would be about £3 million cheaper, assuming that the meters themselves are similarly priced. Therefore, in the remainder of the discussion it will be assumed that the analysis is based on the cheaper Minolta® TCB meter.

Figures 1, 2 and 3 all show that the number of meters necessary to deliver the TCB strategy is important in determining the relative cost-effectiveness of the TCB strategy (strategy 3) to the TSB strategy (strategy 2). In the base-case analysis, TCB is cheaper than TSB providing the number of TCB meters is less than 9,200 (approximately).

If it is decided that more intensive testing is likely to be cost-effective then a secondary decision is whether initial testing should be done using TCB or TSB. Factors such as

convenience to the midwives and discomfort to the baby are not irrelevant to the decision but have not been included explicitly in this analysis because they are difficult to quantify and probably of only relatively small magnitude. This analysis suggests that the choice between TCB and TSB would depend on the number of meters that would be required. The NHS staff census as reported the NHS Information (http://www.ic.nhs.uk/statistics-and-data-collections/workforce/nhs-staff-numbers, accessed August 2009) reports the 'head count' figure for practicing midwives as 25,000 with 19,500 fulltime equivalents. The base-case analyses suggest that were all midwives required to have a TCB meter in order to implement a TCB strategy then TSB would be the cost-effective option. However, not all midwives do post-natal checks. It may be more useful to consider the number of post-natal checks undertaken per day.

If we assume that each birth typically has three post natal visits then than amounts to:

(690,000 x 3) ÷ 365 = 5,670 post natal visits per day

Community midwives would typically do 6-10 post natal visits per day which suggests that the post natal workload is managed by approximately 1,000 midwives on any given day, which might suggest that the service could actually be delivered with less than 9,200 meters.

In interpreting this analysis there are a number of caveats to be considered in addition to the most important ones already highlighted concerning the lack of evidence. The analysis assumes that 25% of infants will require a confirmatory TSB before consideration of phototherapy. If this estimate were higher, then the total cost of the TCB strategy would be higher and the cost threshold at which TSB would be the preferred option would consequently be lower. Also, the analysis assumes that a move to more intensive testing does not lead to increased phototherapy. This might seem a counter-intuitive assumption as the efficacy of more intensive testing is ultimately predicated on not missing cases that could benefit from treatment. However, the belief of the GDG is that the more likely effect of more intensive testing is improved monitoring with their recommendations on thresholds for commencing treatment perhaps leading to lower, but more targeted, intervention than currently occurs.

The analysis also assumes that the different test strategies will not differ in terms of the amount of testing undertaken and the number of follow-up home visits undertaken. Of course, it is possible that the convenience of TCB could lead to additional "downstream" costs not considered here.

An important assumption in this analysis is that phototherapy rates would not change if a more intensive testing strategy was adopted. This is based on the fact that the GDG believe that more intensive testing will correctly identify more babies requiring phototherapy, but will also identify more babies who do not require phototherapy who would have had treatment under current practice. This is a strong assumption in the model since we do not know how many more cases of hyperbilirubinaemia would be correctly identified by a change in testing strategy.

Conclusion

Based on the published limitations of visual examination, the GDG strongly believe that a more intensive testing strategy is required in order to improve outcomes in neonatal jaundice. This will require more resources, but if this reduces the incidence of kernicterus by sufficient numbers (as seems probable), it would necessarily be cost-effective to implement in the NHS. Whilst the analysis presented here is unable to demonstrate that this would be cost-effective, it does suggest that the actual number of kernicterus cases needed for more intensive testing to be cost-effective is relatively small, e.g. 0.4 cases per annum if the TCB strategy could be delivered with 1,000 meters up to 1.4 cases per annum if 9,200 meters are required. This is the cut-off above which the total cost of TSB strategy is cheaper than TCB. The reports from the US have shown a reduction of 4 cases per 100,000 births after the mid 1990s.

Determining which intensive testing strategy is cost-effective depends crucially on the number of meters which would have to be purchased in order to deliver TCB. The number of community midwives involved in home visits on any one day is far smaller than the total number working in the NHS at any one time. Therefore it seems plausible that the TCB strategy could be delivered with a number of transcutaneous meters that is sufficiently low to meet the threshold for cost-effectiveness. However, service delivery is not a remit of NICE guidelines and local commissioners may want to opt for the strategy they believe can be delivered most cost-effectively in their area.