THE INVESTIGATION AND MANAGEMENT OF THE SMALL-FOR-GESTATIONAL-AGE FETUS

1. Aim
The aim of this guideline is to make recommendations regarding the diagnosis and management of small-for-gestational-age (SGA) fetuses. It does not address multiple pregnancies or pregnancies with fetal abnormalities.

2. Introduction and background
SGA refers to a fetus that has failed to achieve a specific biometric or estimated weight threshold by a specific gestational age. Various thresholds (2.5th, 3rd, 5th, 10th, 15th and 25th centiles and 1.0, 1.5 or 2.0 standard deviations below the population average) are used for various fetal measures. The commonly used threshold is the tenth centile for abdominal circumference and estimated birthweight.1

SGA fetuses are a heterogeneous group comprising fetuses that have failed to achieve their growth potential (fetal growth restriction, FGR) and fetuses that are constitutionally small. Approximately 50–70% of fetuses with a birthweight below tenth centile for gestational age are constitutionally small,2,3 and the lower the centile for defining SGA, the higher the likelihood of FGR. On the other hand, a fetus with growth restriction may not be SGA.4

SGA fetuses are at greater risk of stillbirth,1–4 birth hypoxia,4 neonatal complications4 impaired neurodevelopment,5,10 and possibly type 2 (non-insulin-dependent) diabetes and hypertension in adult life.11–13 The reason that studies on SGA fetuses have shown poor perinatal outcome is likely to be the high incidence of true FGR in this group.14,15 However, the vast majority of term SGA infants have no appreciable morbidity or mortality.16

3. Identification and assessment of evidence
The Cochrane Library, Medline, Embase and the NHS Economic Evaluation Database were searched for diagnostic studies, randomised trials, systematic reviews and meta-analyses and economic studies relating to ‘fetal-growth-retardation’, ‘infant-small-for-gestational-age’ and other relevant Medical Subject Heading terms and text-words. The date of the last search was November 2000.

The levels of evidence and the grades of recommendations used in this guideline for effectiveness originate from the US Agency for Healthcare Research and Quality (Appendix I). However, this system of grading is not suitable for diagnostic accuracy studies as a randomised controlled trial may not be an appropriate study design for assessing accuracy and therefore a system of grading devised by the NHS Centre for Reviews and Dissemination is used for diagnostic studies (Appendix II). Where possible, recommendations are based on, and explicitly linked to the evidence that supports
them. When effectiveness grading (Appendix I) is used, the grade is marked with a subscript ‘E’ and when diagnostic accuracy grading (Appendix II) is used, the grade is marked with a subscript ‘D’. Areas lacking evidence are highlighted and annotated as ‘Good practice points’.

4. Diagnosis

Methods employed to detect SGA fetuses include abdominal palpation, measurement of symphysis fundal height, ultrasound biometry, ultrasound estimated fetal weight and ultrasound Doppler flow velocimetry. Four important issues need to be considered with the use of these tests:

- most measurements require an accurate estimation of gestation as a prerequisite
- most tests attempt to diagnose SGA fetuses rather than growth-restricted fetuses
- most studies use a one-off measurement (size) to predict SGA while there is evidence that it is the trend (growth) that is of more value in predicting poor fetal outcome
- in most situations no allowance is made for important prognostic factors for SGA, such as maternal height, weight, ethnicity, parity and fetal gender.

It should also be noted that, although an individual test alone may not be predictive of SGA or FGR, a composite of abnormal results such as an ultrasonically small fetus with reduced liquor or abnormal uterine artery Doppler may indicate pathology.

A distinction needs to be made between biometric tests (tests to measure size) and biophysical tests (tests to assess fetal wellbeing). Biometric tests are designed to predict size and, if performed longitudinally, growth, but not wellbeing. Biophysical tests, on the other hand, are not designed to predict size but fetal wellbeing. The presence of fetal wellbeing implies the absence of fetal acidaemia. This distinction implies that the diagnosis of SGA would rely on biometric tests while abnormal biophysical tests are more indicative of FGR than SGA.

4.1 Abdominal palpation

Abdominal palpation has limited diagnostic accuracy to predict an SGA fetus.

Physical examination of the abdomen by inspection and palpation detects as few as 30% SGA fetuses. If SGA is suspected, it is necessary to supplement abdominal palpation with ultrasound biometric tests.

4.2 Fundal height

Symphysis fundal height (SFH) measurement has limited diagnostic accuracy to predict an SGA neonate.

Although early studies reported sensitivities of 56–86% and specificities of 80–93% for fundal height to predict SGA neonates, a large study of 2941 women found the sensitivity and specificity to be 27% and 88%, respectively. Serial measurements may improve sensitivity and specificity.

The impact on perinatal outcomes of measuring fundal height is uncertain. A systematic review found only one controlled trial with 1639 patients and showed that SFH measurement did not improve any of the perinatal outcomes measured.

Low sensitivity, high false positive rates, significant intra- and inter-observer variation make this test alone unsuitable for diagnosis. Therefore, if SGA is suspected, it is necessary to supplement fundal height measurement with ultrasound biometric tests.
Start the measurement by first identifying the variable point, the fundus, and then measuring to the fixed point, the symphysis pubis, with the cm values hidden from the examiner.\textsuperscript{35}

**B\textsubscript{D}** Use of a customised fundal height chart improves accuracy to predict an SGA fetus.

A customised SFH chart is adjusted for physiological variables such as maternal height, weight, parity and ethnic group. Use of such charts was found to result in improvement in sensitivity (29% and 48% using non-customised and customised charts, respectively), resulting in increased antenatal detection of SGA babies with a reduction in unnecessary hospital investigations for fetal growth.\textsuperscript{36} Calculation of customised centiles (both for fundal height and ultrasound growth) requires computer software\textsuperscript{37} that can be downloaded from the Internet (\texttt{www.gestation.net}), free of charge, for personal or institutional use. These charts can then be printed and incorporated into patient-held records at the time of booking.

4.3 Ultrasound biometry

**B\textsubscript{D}** Use abdominal circumference and estimated fetal weight to diagnose SGA.

Abdominal circumference (AC) and estimated fetal weight (EFW) are the most accurate diagnostic measurements to predict SGA.\textsuperscript{1} In high-risk women, AC at less than the tenth centile has sensitivities of 72.9–94.5% and specificities of 50.6–83.8% in the prediction of fetuses with birthweight at less than the tenth centile. The respective figures for EFW are sensitivities of 33.3–89.2% and specificities of 53.7–90.9%.

Several studies have compared various formulas for estimating birthweight.\textsuperscript{38–40} Most of these suffer from methodological and analytical faults.\textsuperscript{41} A methodologically sound prospective study by Chien \textit{et al.},\textsuperscript{41} compared four formulas (Shepard \textit{et al.},\textsuperscript{42} Aoki,\textsuperscript{43} Campbell and Wilkin\textsuperscript{44} and Hadlock \textit{et al.}\textsuperscript{45}) and found the Shepard and Aoki formulas to have the best interclass correlation coefficient, with EFW showing the smallest mean difference from actual birthweight (Table 1). Thus, the Shepard and Aoki formulas are recommended for estimating fetal weight. However, these formulas were validated over birthweights of 2080–4430g and therefore their use outside this range may be inappropriate. The Hadlock formula may be more appropriate when the fetus is expected to be very small.\textsuperscript{46}

<table>
<thead>
<tr>
<th>TABLE 1. \textbf{FORMULAE FOR ESTIMATION OF FETAL WEIGHT}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shepard’s formula:</strong>\textsuperscript{42} \texttt{Log10EFW} = 1.2508 + (0.166 \times BPD) + (0.046 \times AC) – (0.002646 \times AC \times BPD)</td>
</tr>
<tr>
<td><strong>Aoki’s formula:</strong>\textsuperscript{43} \texttt{EFW} = (1.25647 \times BPD^3) + (3.50665 \times FAA \times FL) + 6.3</td>
</tr>
<tr>
<td><strong>Hadlock’s formula:</strong>\textsuperscript{45} \texttt{Log10EFW} = 1.3596 – 0.00386(AC \times FL) + 0.0064(HC) + 0.00061(BPD \times AC) + 0.0425 (AC) + 0.174 (FL).</td>
</tr>
</tbody>
</table>

\texttt{EFW} = estimated fetal weight (g)

\texttt{BPD} = biparietal diameter (cm)

\texttt{FAA} = fetal abdominal area (cm\textsuperscript{2})

\texttt{FL} = femur length (cm)

\texttt{AC} = abdominal circumference (cm)
Use below tenth centile threshold for both EFW and AC.

A systematic review by Chang et al. found that a threshold of the tenth centile had better sensitivities and specificities than other commonly used centiles. Another study where receiver operator curves were used to determine the cut-off to obtain the best sensitivities and specificities found the customised eighth centile for EFW to be the optimal threshold to predict operative delivery for fetal distress and admission to the neonatal care unit.

Use customised ultrasound charts.

Customised birthweight or ultrasound EFW charts that are adjusted for important independent physiological variables, such as maternal weight, maternal height, ethnic group and parity, have better sensitivities for identifying SGA fetuses and identifying morphometric evidence of FGR, have lower false-positive rates and are predictive of poor perinatal events.

Use growth velocity in addition to size.

Serial measurements of AC and EFW (growth velocities) are superior to single estimates of AC or EFW in the prediction of FGR (abnormal neonatal ponderal index and skinfold thickness) and predicting poor perinatal outcome. However, use of fetal growth alone to diagnose growth restriction (especially when the interval between the scan is less than two weeks) can lead to high numbers of false positives.

The reference charts of fetal biometry based on cross-sectional data are commonly used for assessing growth velocity. However, it is charts based on longitudinal growth studies that reflect growth correctly. Such charts are available for British population from a well-designed study and should be used for measuring growth velocity. Use of standard deviation scores, as opposed to simply “eyeballing” the growth pattern, is likely to result in a more reliable and accurate assessment of growth.

Ratio measures, such as head to abdominal circumference (HC/AC) and femoral length to abdominal circumference (FL/AC) ratios are poorer than AC or EFW alone in predicting SGA or neonatal ponderal index.

A systematic review in the Cochrane Database of Systematic Reviews has shown that routine ultrasound after 24 weeks in low-risk pregnancy does not improve perinatal outcome.

4.4 Biophysical tests to diagnose SGA/FGR

All biophysical tests, including amniotic fluid volume (AFV), Doppler, cardiotocography and biophysical scoring, are poor at diagnosing a small or growth-restricted fetus. The diagnostic accuracies of AFV and uterine artery Doppler are given below as examples of limited accuracy of biophysical tests in diagnosing SGA/FGR.

AFV has minimal value in diagnosing FGR.

Despite the positive association between AFV and neonatal morphometry, the likelihood ratios remain low. For amniotic fluid index (AFI), a positive test result has a likelihood ratio (LR) of 2.4 for predicting skinfold thickness below the tenth
centile and an LR of 1.2 for predicting neonatal ponderal index below 25th centile. The respective negative LRs are 0.6 and 0.8. Serial measurements of AFI have similarly disappointing results (see Table II for interpretation of LRs).

### TABLE 2. INTERPRETATION OF LIKELIHOOD RATIOS

With a positive test result, a LR greater than one increases the probability that SGA or FGR will be present. The greater the LR, the larger the increase in probability of SGA/FGR and the more clinically useful the test result. With a negative test result, a LR of less than one decreases the probability that SGA/FGR is present. The smaller the LR, the larger the decrease and the more clinically useful the test result.

<table>
<thead>
<tr>
<th>Likelihood ratio</th>
<th>Changes in probability of the condition</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 10 or &lt; 0.1</td>
<td>Large</td>
<td>Conclusive</td>
</tr>
<tr>
<td>5–10 or 0.1–0.2</td>
<td>Moderate</td>
<td>Moderately useful</td>
</tr>
<tr>
<td>2–5 or 0.2–0.5</td>
<td>Small</td>
<td>Sometimes useful</td>
</tr>
<tr>
<td>&lt; 2 or &gt; 0.5</td>
<td>Tiny</td>
<td>Rarely useful</td>
</tr>
<tr>
<td>1.0</td>
<td>No change at all</td>
<td>Not useful</td>
</tr>
</tbody>
</table>

**AD** Uterine artery Doppler has limited use in predicting FGR.

A systematic review with meta-analysis published in 2000 found that uterine artery Doppler had limited accuracy in predicting FGR and perinatal death. In the low-risk population the pooled LR to predict FGR was 3.6 for a positive test and 0.8 for a negative test. Even in the high-risk population the pooled LRs were 2.7 and 0.7 for positive and negative tests, respectively.

Although various Doppler studies of fetal circulation such as aortic to middle cerebral artery pulsatility index ratio are used to predict FGR fetuses, their use needs to be evaluated further in primary and secondary studies.

### 5. Management

#### 5.1 Assessment for chromosomal defects

When a small fetus is diagnosed, assess for risk of chromosomal defects.

Up to 19% of fetuses with an AC and EFW less than the fifth centile may have chromosomal defects. The risk is higher when growth restriction is associated with structural abnormalities, a normal liquor volume or a normal uterine or umbilical artery Doppler. Therefore, all growth-restricted fetuses need an ultrasound anatomical survey as a minimum. It may also be appropriate to offer karyotyping.

#### 5.2 Surveillance of the small fetus

**AE** Use umbilical artery Doppler as the primary surveillance tool.

A systematic review with meta-analysis has provided compelling evidence that the use of umbilical artery Doppler to monitor high-risk fetuses reduces perinatal morbidity and mortality. In addition, there was a significant reduction in the number of antenatal admissions and inductions of labour.
associated with Doppler use. A study comparing fetal heart-rate monitoring, biophysical profile and umbilical artery Doppler found that only umbilical artery Doppler had value in predicting poor perinatal outcomes in SGA fetuses.\textsuperscript{62} Use of Doppler does not lead to increased interventions as the rates of positive test are low (2.7\% of all umbilical artery tests in high-risk women).\textsuperscript{63} There is evidence that use of Doppler ultrasound to manage SGA fetuses reduces the use of resources compared with cardiotocography.\textsuperscript{64}

Screening a low-risk or unselected population by umbilical artery Doppler, however, does not reduce perinatal mortality or morbidity.\textsuperscript{65,66} Thus umbilical artery Doppler is not recommended for screening this population.

A variety of descriptor indices of umbilical arterial Doppler waveform, such as resistance index, systolic/diastolic ratio, pulsatility index and diastolic average ratio, is used for predicting perinatal outcome. An analysis using receiver operator curves in a well-conducted study found resistance index had the best discriminatory ability to predict abnormal outcomes such as SGA, poor Apgar scores, abnormal cardiotocograph, umbilical cord pH and admission to neonatal unit.\textsuperscript{67}

When an anomaly scan and umbilical artery Doppler are normal, the small fetus is likely to be a ‘normal small fetus’.\textsuperscript{62,68} Evidence suggests that outpatient management of such fetuses is safe.\textsuperscript{69} In addition, a randomised controlled trial\textsuperscript{70} of two regimens of fetal surveillance for SGA fetuses with normal umbilical artery Doppler found that twice-weekly compared with fortnightly monitoring resulted in earlier deliveries and more inductions of labour with no difference in neonatal morbidity. This suggests frequency of monitoring in SGA fetuses with normal Doppler need not generally be more than once every fortnight. The management of SGA fetuses with abnormal Doppler is discussed in Section 5.3.

**B** Measure liquor volume using either AFI or pocket depth as both tests have similar diagnostic accuracy.

Abnormal liquor volume has been variously defined as single cord-free 1-cm, 2-cm, 1-x-1-cm, 2-x-1-cm and 2-x-2-cm pockets or an AFI below the fifth centile for the gestation or ≤ 5 cm.\textsuperscript{71} Both AFI and single-pocket measurements poorly correlate with actual amniotic fluid volume: An AFI of less than 5 cm and single pocket of less than 2 cm have a positive LR of 2.5 each and a negative LR of 0.94 and 0.97, respectively, in the prediction of amniotic volume at less than the fifth centile for gestation.\textsuperscript{72} However, this study used amniocentesis with dye-dilution and spectrophotometry as the gold standard test and there is evidence that this gold standard itself is inaccurate.\textsuperscript{73}

Despite this, a systematic review with meta-analysis\textsuperscript{74} of 18 studies with over 10 000 patients found an antepartum AFI of ≤ 5.0 cm was associated with an increased risk of an Apgar score of less than seven at five minutes (RR = 5.2; 95\% CI:2.4–11.3). A poor correlation between AFI and neonatal acidosis was noted in the only study that examined this outcome. Other large studies\textsuperscript{75,76} have shown that a reduction in liquor volume is associated with increased perinatal mortality compared with controls with normal liquor volume.

**A** Use biophysical profile and cardiotocography infrequently.

The biophysical profile has not been shown to improve perinatal outcome but sufficient data do not exist to rule out its value: a systematic review\textsuperscript{77} found only four poor-quality studies with fewer than 3000 patients. Authors of the systematic review...
concede that to make a meaningful conclusion about the impact of biophysical profile on perinatal mortality, in excess of 10,000 women would need to be studied. However, there is evidence from uncontrolled observational studies that biophysical profile in high-risk women has good negative predictive value, i.e. fetal death is rare in women with a normal biophysical profile.78

Given the absence of benefit from randomised trials and that biophysical profile is a time-consuming test, it cannot be recommended for routine monitoring in low-risk/unselected pregnancies or for primary surveillance in SGA fetuses. However, when primary surveillance with umbilical artery Doppler is found to be abnormal, biophysical profile is likely to be useful given its good negative predictive value in high-risk populations.79 This recommendation is further supported by evidence that, in high-risk women, the biophysical profile was rarely abnormal when Doppler findings were normal.63

Use of cardiotocography (CTG) antepartum to assess fetal condition is not associated with better perinatal outcome; in fact, a systematic review of randomised trials showed that there was a trend towards increased mortality in the group receiving CTG compared with those who did not.79 Computer systems for interpretation of CTG have better accuracy than clinical experts in predicting umbilical acidosis and depressed Apgar scores.80,81 However, further evaluation of this technology is required before clinical recommendations could be made regarding its widespread use.82

5.3 Delivery

There is wide variation in practice in the timing of delivery of growth restricted fetuses.83 The Growth Restriction Intervention Trial (GRIT) is attempting to answer this issue and publication is expected in 2003. But thus far there is no evidence from GRIT that early delivery to pre-empt severe hypoxia and acidosis reduces any adverse outcome. The following recommendations may need to be revised when more definitive evidence from GRIT is published.

**CE** When end diastolic flow is present (PED), delay delivery until at least 37 weeks, provided other surveillance findings are normal.

Absent or reversed end diastolic flow is associated with increased perinatal mortality and morbidity.84-86 The odds ratio for perinatal mortality in pregnancies complicated by absent end diastolic flow (AED) and reversed end diastolic flow (RED) were 4.0 and 10.6, respectively, compared with when end diastolic flow was present.84 The incidences of respiratory distress syndrome and necrotising enterocolitis were not increased with absent or reversed end diastolic volume but there was an increase in cerebral haemorrhage, anaemia and hypoglycaemia.84

Evidence level IIa

**CE** When end diastolic flow is absent or reversed, admission, close surveillance and administration of steroids are required. If other surveillance results (biophysical profile, venous Doppler) are abnormal, delivery is indicated. If gestation is over 34 weeks, even if other results are normal, delivery may be considered.

The interval between first occurrence of AED and an abnormal CTG/biophysical profile has ranged from 1–26 days.87,88 Gestational age, the presence of hypertension and venous Doppler abnormalities (notably pulsations in the umbilical vein) are the key prognostic factors affecting this interval.87 The optimal surveillance strategy in fetuses with AED/RED is unclear. Options include daily CTG/BPP
and/or venous Doppler with delivery when the CTG becomes pathological (decelerations with reduced variability), the biophysical profile becomes abnormal (≤ 4), there is reversal of Doppler velocities in ductus venosus during atrial contraction or there are umbilical vein pulsations. Under these circumstances, delivery is likely to be by caesarean section.

CE Use gestation- and birthweight-specific charts to determine the likelihood of survival if early delivery is required. A study by Draper et al. produced gestation- and birthweight-specific survival rates from 24 weeks of gestation, taking into account factors such as fetal sex, ethnicity and singleton or twin. Use of these tables, rather than ones based on gestation or estimated birthweight alone, is likely to lead to more accurate estimation of survival. The EPICURE study produced data on developmental disability for extremely preterm infants. Although the data were not specific to growth-restricted infants, they have provided useful information on morbidity when extremely premature delivery is considered.

AE Administer steroids if gestation is below 36 weeks. Antenatal steroids significantly reduce the incidence of respiratory distress syndrome.

CE Deliver in a unit where optimal neonatal expertise and facilities are available.

CE A skilled resuscitator who is trained and competent in resuscitation of the newborn should be present at delivery. Where possible, a neonatologist should be present if gestation is extremely preterm or growth restriction is severe.

CE Intrapartum monitoring with continuous CTG is recommended. Although a systematic review of randomised trials found that intrapartum electronic monitoring did not reduce perinatal mortality, there are substantial observational data to suggest that intrapartum CTG in high-risk populations is likely to be of benefit in reducing perinatal death. Current data are not sufficient to justify a policy of elective caesarean section of all small for gestational age babies.

5.4 Evidence summary for other interventions that have been tried

- Most prenatal interventions do not show any significant effects on perinatal outcome.
- Smoking cessation programmes, particularly behavioural strategies, can be effective for a small minority of smokers in increasing birthweight but there are no data to suggest that this intervention improves perinatal outcome.
- Although a meta-analysis of 13 trials evaluated the use of aspirin in the prevention of growth restriction and found that it reduced the incidence of FGR, only a few studies have used aspirin in the treatment of FGR. These trials are small and have shown conflicting results. Further trials are needed to assess the value of aspirin in the treatment of FGR.
- There is not enough evidence to assess the value of oxygen therapy, nutrient therapy, hospitalisation and bedrest, betamimetics, calcium channel blockers, hormonal therapy and plasma volume expansion in treating growth restriction.
5.5 Plan of management

Local protocols for management of SGA fetus may be developed. An example of a locally developed management plan, in the form of a flow chart, can be viewed on [www.ncl.ac.uk/nfmmg/guidelines/sga%20guide.htm](http://www.ncl.ac.uk/nfmmg/guidelines/sga%20guide.htm).

6. Auditable standards (items iii–vi previously suggested by the RCOG)

i The outcome for all fetuses classified as severe SGA or FGR should be audited.

ii All women with evidence of FGR should be offered surveillance with umbilical artery Doppler and biometry as a minimum.

iii Corticosteroids should be offered to all women who may need delivery between 24 and 36 weeks.112

iv All women who smoke should be given full information regarding the significance of smoking in pregnancy and the postnatal period.112

v Women, and partners, who wish to stop smoking should be given support – e.g. stop-smoking packages, referral to a self-help group.112

vi Neonatal resuscitation: local guidelines should list the equipment to be available for every delivery. Spot checks should be carried out regularly. All items listed should be instantly available. All personnel attending deliveries should receive training in neonatal resuscitation that includes practising practical skills on mannequins and which is repeated yearly.112

References


Levels of evidence and grades of recommendations used for effectiveness studies

Clinical guidelines are: ‘systematically developed statements which assist clinicians and patients in making decisions about appropriate treatment for specific conditions’. Each guideline is systematically developed using a standardised methodology. Exact details of this process can be found in Clinical Governance Advice No 1: Guidance for the Development of RCOG Green-top Guidelines (available on the RCOG website www.rcog.org.uk/medical/greentopguide.html). These recommendations are not intended to dictate an exclusive course of management or treatment. They must be evaluated with reference to individual patient needs, resources and limitations unique to the institution and variations in local populations. It is hoped that this process of local ownership will help to incorporate these guidelines into routine practice. Attention is drawn to areas of clinical uncertainty where further research may be indicated.

The evidence used in this guideline was graded using the scheme below and the recommendations formulated in a similar fashion with a standardised grading scheme.

**Classification of evidence levels for effectiveness studies**

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>Evidence obtained from meta-analysis of randomised controlled trials.</td>
</tr>
<tr>
<td>Ib</td>
<td>Evidence obtained from at least one randomised controlled trial.</td>
</tr>
<tr>
<td>IIa</td>
<td>Evidence obtained from at least one well-designed controlled study without randomisation.</td>
</tr>
<tr>
<td>IIb</td>
<td>Evidence obtained from at least one other type of well-designed quasi-experimental study.</td>
</tr>
<tr>
<td>III</td>
<td>Evidence obtained from well-designed non-experimental descriptive studies, such as comparative studies, correlation studies and case studies.</td>
</tr>
<tr>
<td>IV</td>
<td>Evidence obtained from expert committee reports or opinions and/or clinical experience of respected authorities.</td>
</tr>
</tbody>
</table>

**Grades of recommendations effectiveness studies**

- **A** Requires at least one randomised controlled trial as part of a body of literature of overall good quality and consistency addressing the specific recommendation. (Evidence levels Ia, Ib)

- **B** Requires the availability of well-controlled clinical studies but no randomised clinical trials on the topic of recommendations. (Evidence levels IIa, IIb, III)

- **C** Requires evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities. Indicates an absence of directly applicable clinical studies of good quality. (Evidence level IV)

**Good practice point**

- Recommended best practice based on the clinical experience of the guideline development group.
APPENDIX II

Classification of evidence levels for diagnostic accuracy studies

I Evidence obtained from studies with a blind comparison of test to reference standard among an appropriate broadly defined sample of consecutive patients.

II When any one of the following is present in the study: narrow population spectrum; differential use of reference standard; reference standard not blind; case–control study design.

III When any two of the following are present in the study: narrow population spectrum; differential use of reference standard; reference standard not blind; case–control study design.

IV When any three or more of the following are present in the study: narrow population spectrum; differential use of reference standard; reference standard not blind; case–control study design.

V Expert opinion with no explicit critical appraisal, based on physiology, bench research or first principles.

Grades of recommendations for diagnostic accuracy studies

A\text{D} Requires at least one level I study.

B\text{D} Requires level II or III studies.

C\text{D} Evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities. Indicates an absence of directly applicable clinical studies of good quality. (Evidence level IV and V)

Good practice point

✓ Recommended best practice based on the clinical experience of the guideline development group.