ORIGINAL PAPER

First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS)

N J Wald, C Rodeck, A K Hackshaw, J Walters, L Chitty, A M Mackinson

See end of article for authors' affiliations and list of investigators

Correspondence to: Prof N J Wald, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, London, EC1M 6BQ, UK; n.j.wald@qmul.ac.uk

Accepted for publication December 2003

J Med Screen 2003; **10**:56–104

CONTENTS

EXECUTIVE SUMMARY

Objectives

To identify the most effective, safe and cost-effective method of antenatal screening for Down's syndrome using nuchal translucency (NT), maternal serum and urine markers in the first and second trimesters of pregnancy, and maternal age in various combinations.

Design

A prospective study of women who booked for their antenatal care at about 8–14 weeks of gestation, with follow-up to identify pregnancies with Down's syndrome ascertained through second trimester screening or at birth.

Setting

Twenty-five maternity units (24 in the UK and one in Austria) offering second trimester Down's syndrome serum screening that agreed to collect observational data in the first trimester.

Participants

The results were based on 47,053 singleton pregnancies, including 101 pregnancies with Down's syndrome.

Measurements and tests

NT measurements were included if obtained between 9 and 13 weeks of pregnancy; serum and urine samples were also taken and stored. Another pair of serum and urine samples was collected in the second trimester and included if obtained between 14 and 20 weeks. Urine and serum samples from each affected pregnancy and five matched controls were tested for:

Serum:

- alphafetoprotein (AFP)
- total human chorionic gonadotrophin (hCG)
- unconjugated oestriol (uE_3)
- pregnancy associated plasma protein A (PAPP-A)
- free β-hCG
- dimeric inhibin-A.

Urine:

- invasive trophoblast antigen (ITA)
- β-core fragment
- total hCG
- free β-hCG.

The matching criteria were gestation (using an ultrasound crown–rump length or biparietal diameter measurement), duration of storage, and centre. Screening performance of the individual markers and combinations of markers together with maternal age was assessed using standard methods. In addition pairs of first and second trimester serum samples from 600 controls were tested to secure a larger set in which screening performance could be determined using distribution parameters based on dates (time since first day of the last menstrual period).

Main outcome measures

The following were determined for different combinations of markers:

- efficacy (by assessing screening performance, focusing on the false-positive rate (FPR) for an 85% detection rate (DR))
- safety (focusing on the number of fetal losses due to amniocentesis (or chorionic villus sampling) in 100,000 women screened)
- cost-effectiveness (focusing on the cost of screening 100,000 women and the cost per Down's syndrome pregnancy diagnosed).

Test (all include maternal age)	Measurements	FPR for 85% DR (%)	95% confidence interval (%)
Integrated test	NT and PAPP-A at 10 completed weeks AFP, uE_3, free β-hCG and inhibin-A at 14–20 completed weeks	1.2 (1.3[a])	1.0–1.4 (1.2–1.4[a])
Serum integrated test	Integrated test without NT. PAPP-A at 10 completed weeks	2.7 (4.9[a])	2.4–3.0 (4.4–5.4[a])
Combined test	NT, free β-hCG and PAPP-A at 10 completed weeks	6.1 (6.0[a])	5.6–6.5 (5.5–6.5[a])
Quadruple test	AFP, uE_3, free β-hCG, inhibin-A at 14–20 completed weeks	6.2	5.8–6.6
Triple test	AFP, uE_3, free β-hCG at 14–20 completed weeks	9.3	8.8–9.8
Double test	AFP and free β-hCG at 14–20 completed weeks	13.1	12.5–13.7
NT measurement	NT at 12–13 completed weeks	20.0	18.6–21.4

[a] NT and/or serum measurements at 12 completed weeks of pregnancy

Results

Efficacy (screening performance)

The false-positive rates for an 85% detection rate for the main screening tests are shown in the above table, in decreasing order of screening performance:

With the serum integrated test, 10 weeks is the preferred time in pregnancy for the PAPP-A measurement. For the integrated test and the combined test, the timing of the measurement of the first trimester markers is less critical.

Safety

The lower false-positive rate with the integrated test compared with other tests means that at an 85% detection rate there would be nine diagnostic procedure-related unaffected fetal losses per 100,000 women screened compared with 44 using the combined test or 45 with the quadruple test.

Cost-effectiveness

Screening using the integrated test is less costly than might be expected because the extra screening costs tend to be offset by savings in the cost of diagnosis arising from the low false-positive rate. It was estimated that to achieve an 85% detection rate the cost to the UK NHS would be £15,300 per Down's syndrome pregnancy detected. The corresponding cost using the second trimester quadruple test would be £16,800 and using the first trimester combined test it would be £19,000.

Conclusions

Implications for healthcare

The results showed that screening performance in the first trimester of pregnancy was virtually the same as that in the second trimester, and in either it was much less effective than integrating screening measurements from both trimesters into a single test. In applying these results to screening practice several conclusions can be drawn. The following tests offer the most effective and safe method of screening:

- overall: the integrated test
- if an NT measurement is not available: the serum integrated test
- for women who do not attend for antenatal care until the second trimester of pregnancy: the quadruple test
- for women who choose to have a screening test in the first trimester: the combined test.

At a constant detection rate, the cost-effectiveness of these four tests is broadly similar, any extra screening costs tending to be offset by fewer diagnostic costs. The evidence presented in this report does not support retaining the double test, the triple test, or NT measurements on their own (with or without maternal age) because each would lead to many more women having invasive diagnostic tests, without increasing the proportion of Down's syndrome pregnancies detected.

1 INTRODUCTION

We here report the results of the Serum Urine and Ultrasound Screening Study (SURUSS), a large collaborative study of antenatal screening for Down's syndrome, funded as part of the UK Health Technology Assessment (HTA) Programme, to help determine best screening practice.

Antenatal screening for Down's syndrome has developed rapidly over the last 15 years. In 1988, maternal age screening was improved by the second trimester triple test.[1] Some centres adopted the double test. (A glossary of definitions of the various screening tests and a key to abbreviations used are included at the end of this report.) The triple test was later improved by the addition of maternal serum inhibin-A to form the quadruple test. At the same time, three first trimester markers, serum pregnancy-associated plasma protein A (PAPP-A), free β-human chorionic gonadotrophin (β-hCG) and the ultrasound marker nuchal translucency (NT; see glossary) were shown to be useful in screening. A systematic review of antenatal screening for Down's syndrome published in 1997 recommended that the second trimester triple or quadruple test should be the test of choice.[2] In 1999 the integrated test was described,[3] which combined markers from the first and second trimesters to yield a screening performance better than from either trimester alone. Several urinary markers have been proposed as screening tests, notably β-core fragment[4] and invasive trophoblast antigen (ITA).[5]

The value of SURUSS is that it provides a large dataset on women seen in both the first and second trimester of pregnancy (it is the largest such dataset yet reported), without planned intervention in the first trimester. This allows a direct examination of the screening performance of all individual screening markers – NT and first and second trimester serum and urine markers. The strength of SURUSS is that it can do this in a single large unselected group of pregnant women with data collected in both trimesters, in a collaborative study from 25 centres that together reflect the provision of routine antenatal care.

2 METHODS

The study was based on women attending 25 maternity centres (24 in the UK and one in Austria). Most centres began recruiting in September 1996, following a pilot study started in January 1995 at one centre. Recruitment to the study ended in April 2000, and follow-up of pregnancy outcome was carried out to 31 May 2001.

General

The study took place at centres where second trimester serum screening for Down's syndrome was already established using the double, triple, or quadruple test. Women who booked for their antenatal care between about 8 and 14 weeks of pregnancy (based on the first day of their last menstrual period (LMP)) were invited to join the study after receiving an information leaflet and verbal explanation of the study. At their booking visit each woman had an ultrasound examination which included confirmation that the fetus or fetuses were alive, a crown–rump length (CRL) measurement (or, failing that, a biparietal diameter (BPD) measurement), and, if possible, at least three NT measurements. Hard copies of the fetal ultrasound image showing the NT measurement were sent to the coordinating centre at the Wolfson Institute of Preventive Medicine (London, UK). Two sets of serum and urine samples, from the booking visit and from the time of the second trimester screening test, were stored at –20°C or colder, and posted regularly, on dry ice, to the Wolfson Institute for storage at –40°C. Antenatal diagnostic tests for Down's syndrome and subsequent terminations of pregnancy were based on the routine second trimester double, triple, or quadruple tests; the SURUSS analyses on the serum and urine markers were not performed until the outcome of pregnancy was known. As far as possible the study was non-interventional in the first trimester. The policy was to avoid early medical intervention on the basis of the NT measurement alone. Pregnancies in which the NT measurement was ≥3 mm were flagged to ensure that the women did not miss the routine second trimester screening test. This non-interventional design aimed to compare first and second trimester screening tests without the bias caused by the diagnosis and termination of some Down's syndrome pregnancies, and the miscarriage of others between the first and second trimester. Bias that would favour first trimester screening could arise in two ways:

- some of the affected pregnancies detected by first trimester screening and terminated would have miscarried in the absence of screening
- some markers preferentially detect affected pregnancies that miscarry.

It was not possible to avoid such bias arising after about 16 weeks of pregnancy because intervention (antenatal diagnosis and selective termination of pregnancy) was offered to women following a positive second trimester screening result. The performance of screening was therefore assessed at about 17 weeks of pregnancy and not at term as in some previous observational studies in which results were based on affected livebirths.

The outcome of pregnancy was obtained in six ways:

- staff at local hospitals completing a SURUSS pregnancy outcome form at or just after delivery
- linking the SURUSS records relating to women who had a chorionic villus sampling (CVS) or amniocentesis or a karyotype at birth with information from cytogenetic laboratories
- linking the SURUSS records to cases of Down's syndrome from the National Down Syndrome Cytogenetic Register
- obtaining information from local obstetrical outcome records
- sending a form to all women with a request to return details of the outcome of their pregnancy to the Wolfson Institute
- individual searches in respect of women whose outcomes of pregnancy had not been obtained by any of the previous methods.

Details of the outcome of pregnancy included the presence or absence of Down's syndrome, other chromosomal abnormality or congenital anomaly during pregnancy or at birth, and whether the pregnancy was terminated, miscarried or resulted in a stillbirth or livebirth.

Nuchal translucency measurements

Sonographers in the study received training in NT measurement. A hard copy of every NT image (up to six per pregnancy) was independently assessed for quality and acceptability, including the position of the fetus, whether the correct translucency was measured and whether the image magnification was satisfactory. Images judged to be technically unsatisfactory were noted as such. This process of quality control was maintained throughout the study. Sonographers spent up to 20 minutes to obtain an NT image.

All 260 sonographers in the study were asked to obtain at least three NT measurements for each fetus (though some obtained less and others up to six) and return a hard copy for independent review. NT was measured using a mid-sagittal section of the fetus at optimal magnification and taken as the maximum thickness of the translucent space between the inner skin surface and the fascia covering the cervical spine. The outer electronic caliper was placed on the white/black interface, and the inner caliper was placed on the black/white interface. The measurement was made after observing fetal movement to distinguish between the fetal skin and the amniotic membrane. The name of the sonographer and the make and model of the machine used were recorded in respect of each pregnancy (there were 41 different models of ultrasound machine made by nine different manufacturers). Sonographers in a sample of seven centres were asked to record the time it took to obtain the NT measurements (this was not the time for the complete ultrasound examination which included setting up the equipment, and dating the pregnancy by measuring the CRL or BPD).

Biochemical analyses

Each pregnancy with Down's syndrome (affected pregnancy or case) was identified and each pregnancy was matched with five singleton unaffected pregnancies (controls) according to centre, maternal age (± 10 years), CRL (± 5 mm), or BPD (± 5 mm) if a CRL was not available, and duration of storage of the serum and urine samples (within 18 months). The first and second trimester serum samples were retrieved from storage, thawed and the following assays performed on each: α-fetoprotein (AFP), free β-hCG , total hCG, unconjugated oestriol (uE_3) and PAPP-A using time-resolved fluoroimmunoassay (AutoDELFIA Perkin Elmer™ (Life Sciences, USA)), and dimeric inhibin-A using the solid phase sandwich enzyme-linked immunosorbent assay (Oxford Bio-innovation and Diagnostic Systems Laboratories Incorporated, UK). The first and second trimester urine samples were also retrieved from storage and the following assays performed: ITA and β-core fragment (Quest Diagnostics, USA), and total hCG and free β-hCG (the same kits used to measure these markers in serum). Urine creatinine measurements were carried out to adjust the urine marker levels for extent of urine dilution. Samples from cases and controls were assayed in the same analytical batch without knowledge of which were from cases and which from controls.

After a preliminary analysis of the data and the observation that the urine markers did not add materially to screening performance (see below) we identified an additional 600 unaffected pregnancies at random, with both first and second trimester samples available to obtain more robust estimates of screening performance for tests using the serum markers with gestational age based on dates (time since first day of the LMP), and gestation based on an ultrasound scan examination (BPD or CRL). Together with the 490 matched controls this provided over 1000 pregnancies in which the distribution parameters of the serum markers could be compared using both methods of estimating gestational age. We also used this dataset to examine the effects of adjusting the concentration of the serum markers for maternal weight.

Statistical analyses

Data collected before 14 weeks 0 days were regarded as first trimester data. Data collected on or after 14 weeks 0 days were regarded as second trimester. (No data were collected in the third trimester.) Definitions of some of the terms used are given in the glossary.

Only 16% of pregnancies at 9 and 10 weeks booked at 9 weeks and because there were only four cases of Down's syndrome at this time, we hereafter classify them as at 10 weeks. There were few cases of Down's syndrome before 10 completed weeks and none after 20 completed weeks. Our results on screening therefore cover this 11-week period.

Markers and gestational age

In the first trimester there was a reasonable straight line fit to the relationship between gestational age and AFP, total and free β-hCG, inhibin-A, PAPP-A and NT, all expressed as $\log_{10}$ values, with the exception of $\log_{10}$ uE_3 which was quadratic. In the second trimester there was a straight line relationship between gestational age and AFP, uE_3 and PAPP-A, all expressed as $\log_{10}$ values. There was a quadratic relationship with $\log_{10}$ inhibin-A and a declining exponential relationship with total and free β-hCG. Similarly there was a log-linear or log-quadratic relationship with the urine markers. The relationship between gestational age and the serum and urine marker levels, and the fitted regression line or curve are shown in the appendix Figures 11–14, together with the observed values for affected pregnancies.

Calculating multiples of the median values

To allow for the systematic changes in marker levels with increasing gestational age all concentrations were converted into multiples of the median (MoMs) for a given gestational age among unaffected pregnancies. This "normal" median was estimated for each serum and urine marker (in the first and second trimester separately) from a weighted regression of the median marker level at each completed week of gestation in unaffected pregnancies using, in each completed week, the median gestational age in days. For NT the log-linear regression was performed separately in five groups with CRLs of 20–31, 32–42, 43–53, 54–64 and 65–75 mm which approximately correspond to 9, 10, 11, 12 and 13 completed weeks of pregnancy. Gestational age was estimated from the CRL; in women who had only a BPD available, the CRL was estimated from the strong linear relationship between CRL and BPD as (3.603 × BPD) – 15.516, based on an analysis of 12,464 women for whom both CRL and BPD values were available. The MoM value for each pregnancy was then calculated by taking the marker level for each pregnancy and dividing this by the normal median level from the regression. MoM values for the biochemical markers were adjusted for maternal weight (NT was not influenced by maternal weight) but there were insufficient data to adjust for ethnic origin (only 9% of cases were not white). The MoM values for the urine markers were corrected for the creatinine concentration (see appendix, Figures 15 and 16).

In general, the MoM values of each marker in each trimester of pregnancy fitted a log Gaussian distribution reasonably well in both affected and unaffected pregnancies as judged by observing a straight line on a probability plot. These are shown in the appendix, Figures 17–21. Urine total hCG required a square root transformation. NT in affected pregnancies was positively skewed, even after transformation into $\log_{10}$ values, but adequately fitted a log Gaussian distribution over the 10th–85th centile range. In unaffected pregnancies $\log_{10}$NT was also skewed above 2.5 MoM. On the basis of the first trimester results from SURUSS (presented later in this report), gestation-specific NT standard deviations (SDs) were used for unaffected pregnancies and gestation-specific medians in affected pregnancies for uE_3, total hCG, free β-hCG, inhibin-A, PAPP-A, and urine ITA. Because of this, screening performance will vary at each week between 10 and 13 weeks, so any single estimate of performance over the 10–13-week period would be influenced by the proportion of women screened at each week of pregnancy. Results were therefore shown for 10, 11, 12 and 13 weeks separately, and for 10–13 weeks together, based on women at each week from 10–13 weeks in SURUSS.

Distribution parameters

To estimate risk, the distributions of the markers in affected and unaffected pregnancies need to be specified, namely the mean marker levels, the SDs and the correlation coefficients between markers (all based on a $\log_{10}$ transformation except for urinary total hCG where the SD and correlation coefficient were based on a square root transformation). The mean of the transformed values (in MoMs) was estimated by the $\log_{10}$ or square root of the median to avoid the influence of outliers. The SD was estimated using a regression between the transformed marker level and corresponding centile value within the 10th–90th centile range (except for NT, for which the SD was calculated using all the data without censoring values). The correlation coefficients were estimated by dividing the covariances between each pair of markers (after excluding observations that were ± 3.5 SDs away from the mean) by the product of the corresponding SDs.

Separate estimates of the SDs for the serum markers were obtained with gestational age based on dates (time since first day of the LMP) and an ultrasound scan examination (BPD or CRL). This was done using the subset of unaffected pregnancies in which both a scan and dates estimate of gestation were available and maternal weight was recorded. In this subset the differences in variance (the square of the SD) between scan and dates (and with and without weight correction) were determined. These differences were then applied to the observed variance in the full set of affected and unaffected pregnancies, when based on a scan estimate of gestation without weight correction (such pregnancies were treated as the baseline group since weight was only recorded for about 80% of women). For example, in the subset of unaffected pregnancies the variance for uE_3 (in the second trimester) was 0.1145^2 using scan and 0.1282^2 using dates (without weight correction); the difference is 0.003325 (0.1282^2–0.1145^2). This difference was added to the variance based on the scan without weight correction, observed in the full dataset of unaffected pregnancies (0.1156^2), to give

the estimated variance for dates gestation, 0.1292^2 (0.1156^2 + 0.003325). The covariances (used to derive the correlation coefficients) were estimated in a similar manner. The method has been previously described.[6]

Risk estimation

The risk of having a pregnancy with Down's syndrome at about 17 weeks was estimated by multiplying the maternal age-specific odds of having an affected livebirth[7] (corrected to early mid-trimester by multiplying by 1/0.77 to allow for the general fetal loss of Down's syndrome pregnancies from this time in pregnancy until term) by the likelihood ratio (for a given set of marker values) obtained from the overlapping univariate or multivariate Gaussian distributions of affected and unaffected pregnancies.[8] A woman was classified as screen-positive if her risk estimate was equal to or greater than a specified cut-off level. The screening performance of each marker, considered separately and in combination with other markers, was specified in terms of detection rates (DRs) corresponding to specified false-positive rates (FPRs; see glossary). When estimated using maternal age, DRs and FPRs were estimated for a standard population of maternities in England and Wales. This was done using the distributions of risk in affected and unaffected pregnancies, found by numerically integrating the specified univariate or multivariate Gaussian distributions over each year of maternal age using the age distribution of maternities in England and Wales, from 1996–98 inclusive,[9] as previously described.[1,10] Our results on screening performance are based on estimates modelled in this way unless otherwise indicated. We provide confidence intervals (CIs) for the main estimates of performance using Monte Carlo simulation.

Data collected and data used

Table 1 shows, according to centre, the number of women recruited, the median age, the percentage aged 35 years or older and the number of Down's syndrome pregnancies. There was a total of 47,972 fetuses (47,053 singleton pregnancies, 443 twin pregnancies and 11 triplet pregnancies) among 47,507 recruited women. There were 101 Down's syndrome pregnancies identified, of which 71 were terminated, four miscarried shortly after amniocentesis, and 26 ended in a livebirth. Of these 101, three had no serum or urine samples at all (though they did have an NT measurement) and biochemical analyses were performed on the remaining 98 and their controls. This report is based on singleton pregnancies.

Five Down's syndrome pregnancies were terminated before 14 weeks; their NT measurements were 0.8, 2.0, 3.6, 3.9 and 4.9 mm (the last three measurements would usually be regarded as high). We cannot exclude the possibility that the NT measurement in some of these contributed to the decision to offer CVS. In the absence of selective abortion, about three of these pregnancies would have proceeded to the second trimester. To allow for this intervention bias we therefore censored two at random and excluded them from the analyses to avoid the over-estimation of the median marker levels in affected pregnancies. Of the 26 Down's syndrome pregnancies that went to term, 10 were among screen-negative women, two were in screen-positive

Table 1 The number of women recruited, median maternal age, percentage of women aged ≥35 years at expected date of delivery and the number of Down's syndrome pregnancies

Centre	Number of women recruited	Maternal age		Number of Down's syndrome pregnancies
		Median (years)	% ≥35	
Aberdeen Maternity Hospital	1,830	29	16	3
Alexandra Hospital, Redditch	744	28	12	2
Barnsley District Hospital	2,140	27	10	3
Billinge Hospital, Wigan	2,357	28	10	2
Birmingham Heartlands Hospital	4,154	28	14	6
Birmingham Women's Hospital	673	29	17	3
Derby City Hospital	634	29	14	0
Donauspital am SMZO, Vienna	3,530	29	13	12
Dr Gray's Hospital, Elgin	1,265	28	13	4
Kettering General Hospital	1,322	28	14	0
Leicester General Hospital	2,001	28	10	3
Liverpool Women's Hospital	1,802	28	13	2
Llandough Hospital	2,195	29	14	8
Nevill Hall Hospital, Abergavenny	227	27	10	0
Peterhead Hospital	165	26	6	0
Royal Gwent Hospital, Newport	3,363	28	11	7
Singleton Hospital, Swansea	2,339	28	14	4
Stepping Hill Hospital, Stockport	2,221	30	16	7
St John's Hospital, Chelmsford	900	30	19	3
The County Hospital, Brighton	1,324	29	13	2
The General Hospital, Jersey	2,258	30	22	7
University College Hospital, London	3,533	31	27	12
University Hospital Wales, Cardiff	2,069	29	16	3
Watford General Hospital	172	31	16	0
Whittington Hospital, London	4,289	30	24	8
All 25 centres	47,507	29	16	101

women who had declined an amniocentesis and 14 were in women who declined screening but agreed to provide a blood and urine sample. Because most women in the study had second trimester serum screening, there will have been some unrecognised affected screen-negative pregnancies that miscarried. The inability to include such pregnancies in the analysis while at the same time having to include affected screen-positive pregnancies that ended in a termination of pregnancy, but would otherwise have miscarried, would tend to overestimate the discriminatory value of the screening markers unless a correction were made. There were ten screen-negative affected pregnancies, and as an estimated 77% of affected pregnancies[11] survive from the second trimester to term, 13 would have been expected in the second trimester. To correct for this second intervention bias, three pregnancies were therefore sampled at random from the ten (and the average of five samplings used) and their marker levels were added to those already included.

The dataset used for obtaining estimates of screening performance was based on 43,712 singleton pregnancies recruited at 10–13 weeks of gestation (from the total of 47,053 recruited at all gestations). The total number of affected pregnancies in the main analyses was 102 (101 observed minus 2 terminated before 14 weeks plus 3 expected miscarriages in screen-negative women). It was recognised (and later confirmed) that the adjustments would have only a minor effect on the results but in the interests of accuracy this was done. Table 2 classifies these pregnancies according to the data available.

Completeness of ascertainment of Down's syndrome

In the absence of screening and selective abortion there would have been an estimated 81 affected livebirths (55 of the 71 that were terminated would have gone to term (77% survival rate to term)[11] plus the 26 livebirths). The expected number of affected term livebirth pregnancies, based on the maternal age distribution of the women in SURUSS, was 87, a little higher, but consistent with the estimate of 81 indicating that ascertainment of affected pregnancies in the study was adequate and probably complete. An outcome was documented for 96% of all pregnancies in SURUSS.

Supplementary fetal loss study

A supplementary nested case–control study (268 pregnancies that ended in a miscarriage and 95 that ended in a stillbirth, each matched with three livebirths) was carried out to investigate the relationship between the risk of spontaneous fetal loss and the serum markers and NT. Table 3 summarises the data used.

Table 3 Data used to estimate screening performance and to assess association between serum markers and spontaneous fetal loss

Estimation of screening performance	Number of pregnancies
NT measurements	
Down's syndrome pregnancies	85
Unaffected pregnancies	39,898
Down's syndrome case–control set	
Down's syndrome pregnancies with a first or second trimester sample available	98
Matched controls (5 for each case)	490
Additional controls	
Serum markers only	600
Fetal loss case–control study	
All fetal losses (268 miscarriages, 95 stillbirths)	363
Matched controls (3 for each case)	1,089

Note: in both studies controls were matched on centre, maternal age, CRL or BPD. In the main study they were also matched for duration of storage of serum or urine sample.

Safety and cost

To assess safety we used a review of the evidence on the risk of a fetal loss arising as a complication of the diagnostic procedure (an amniocentesis or CVS)[12] that follows a screen-positive result. The excess risk of fetal loss from amniocentesis was 0.9%. The estimate for transabdominal CVS was about the same, though there is indirect evidence to suggest that it may be higher.

The UK cost estimates that apply to a public service (NHS) screening programme were based on two sources:

- the costs of reagents and non-reagent laboratory costs and service costs were obtained from our systematic review,[2] adjusted for inflation and rounded up to the nearest 50p for items under £10
- the other costs were based on those from Gilbert and her colleagues[13] relating to:
  - the marginal costs of NT measurements (£4.50) again rounded to the nearest 50p
  - amniocentesis and CVS (including karyotype) approximated to £200 and £250 respectively with the addition of £100 for a rapid polymerase chain reaction (PCR) test (costs rounded to the nearest £50)
  - termination of pregnancy, delivery of an affected term birth and the surgical or medical evacuation of retained products of conception arising from procedure-related miscarriage, rounded to £500, £600 and £400 respectively (costs rounded to the nearest £100).

Table 2 Singleton pregnancies in the study

		Number with Down's syndrome	
	All pregnancies	Observed	Adjusted for intervention biases[a]
Total (all gestations)	47,053	101	102
Total (10–13 completed weeks[b])	43,712	94	95
NT measurement	39,983	85	86
First trimester serum sample	40,387	85	86
Second trimester serum sample	37,362	82	85
First trimester urine sample	41,546	86	87
Second trimester urine sample	35,899	65	68
NT measurement, first and second trimester serum samples	30,375	65	68
NT measurement, first and second trimester serum and urine samples	28,434	50	53

[a] The biases are described in chapter 2; [b] Based on an ultrasound scan. First trimester: 10–13 completed weeks; second trimester: 14–22 completed weeks.

Using each test we estimated the cost of screening 100,000 women (including the costs of diagnosis and termination of pregnancy) to achieve a specified DR and the corresponding cost per Down's syndrome pregnancy diagnosed. An illustration of how these costs were derived is given in the section on cost-effectiveness in chapter 3.

3 RESULTS

Nuchal translucency measurement: the role of machine and sonographer

This section is based on all 101 singleton Down's syndrome pregnancies recruited into the study at any gestational age (between 6 and 16 weeks) in which an NT measurement was attempted, without adjustment for the biases described in the section "Data collected and data used" (pages 60–1).

Figure 1 shows the number of pregnancies and the percentage in which an NT measurement could not be obtained within the 20 minutes available for the examination. Overall an NT measurement was not obtained in 9% of pregnancies. The failure rate varied with gestational age and was 33% or greater before 10 completed weeks of pregnancy, declined to 7% at 12 weeks, and then increased again to 32% or more after 14 weeks. During the course of the study the failure rate decreased significantly. Figure 2 shows the failure rate at each week of pregnancy according to study period (divided into fifths for each centre). At 10 completed weeks of pregnancy the failure rate almost halved from 14.5% during the first fifth of the study to 7.5% during the last fifth. At 12 completed weeks the rate fell from 8.1% to 4.8%. The overall failure rate between 10 and 12 weeks decreased from 9% to 5%. The results show that with experience, the ability to obtain an NT measurement is similar at 10, 11 and 12 weeks of pregnancy.

Figure 3 shows, in a similar way to Figure 1, the percentage of pregnancies in which an NT measurement was obtained, but all images for each pregnancy were judged to

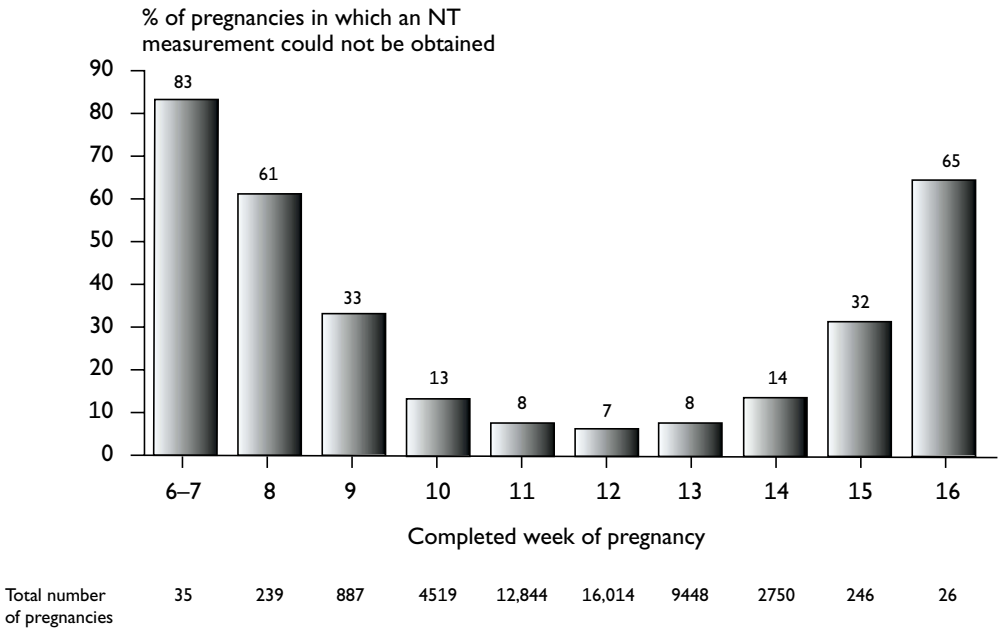

Figure 1 Percentage of pregnancies in which a NT measurement could not be obtained according to gestational age. Of the 47,053 pregnancies, 45 were excluded because they had no gestational age estimate recorded or gestation was ≥17 weeks.

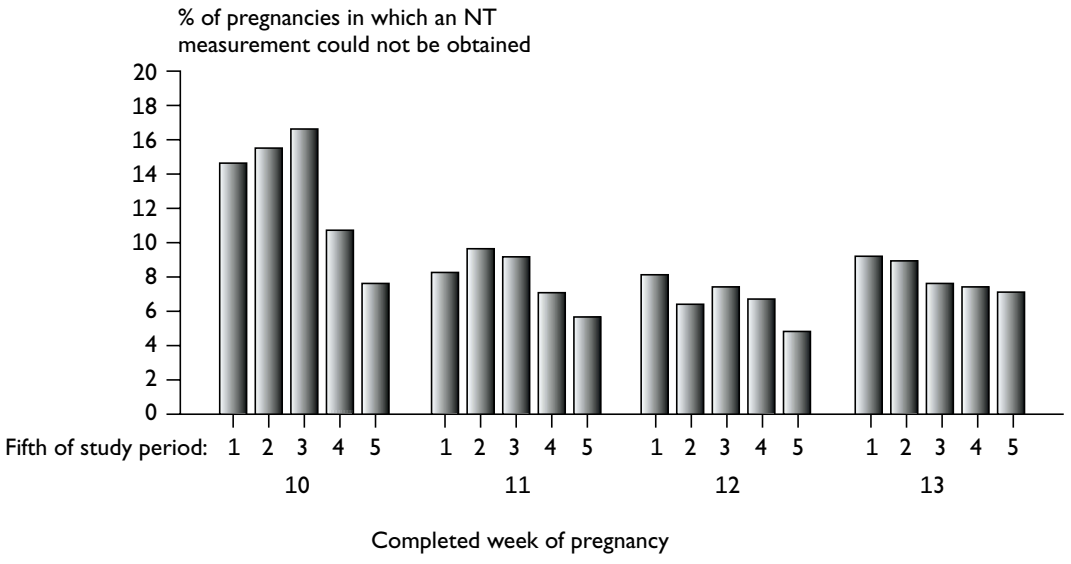

Figure 2 Percentage of pregnancies in which a NT measurement could not be obtained, according to gestational age and study period.

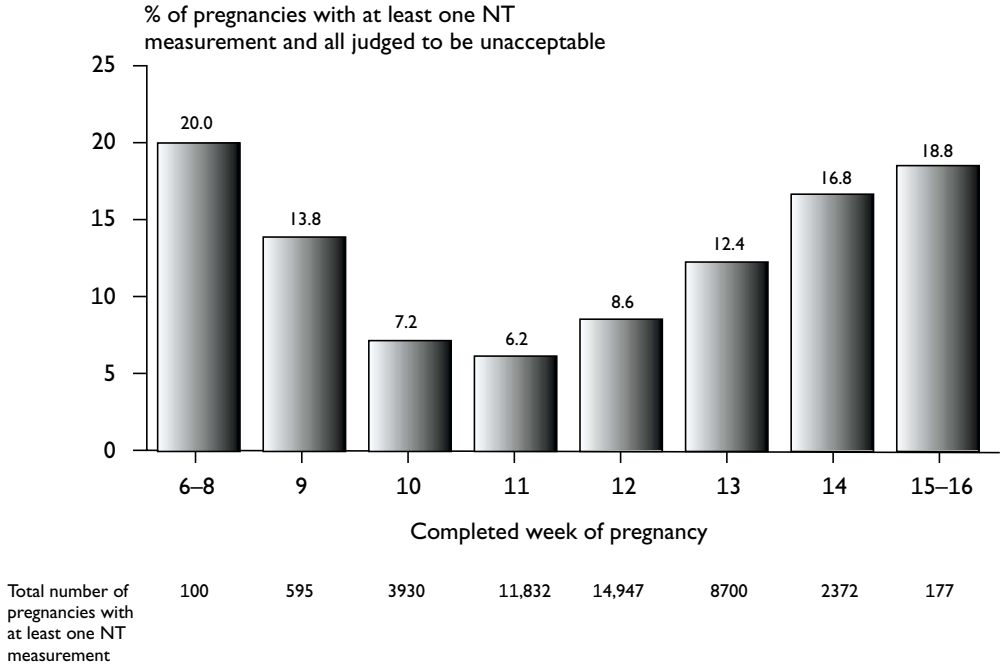

Figure 3 Percentage of pregnancies with at least one NT measurement in which all the measurements obtained were judged to be technically unsatisfactory. Of the 42,673 pregnancies with at least one NT measurement, 20 were excluded because they had no gestational age estimate recorded or gestation at ≥17 weeks.

be unsatisfactory. Overall this was 9% (3914/42,673) of pregnancies. The proportion was lowest at 11 completed weeks of pregnancy (6%). Combining the data in Figures 1 and 3, failure to obtain a satisfactory NT measurement occurred in 42% of all pregnancies at 9 completed weeks, 19% at 10, 14% at 11, 15% at 12 and 19% at 13 completed weeks.

Sonographer experience influenced the ability to obtain a satisfactory image. Among sonographers who had scanned fewer than 200 pregnancies, measurements were not obtained in 9% of all pregnancies at 10–13 completed weeks of gestation and of those in which an NT was measured, all images were judged to be technically unsatisfactory in 11% of pregnancies. Among those who had scanned 400 or more pregnancies the corresponding rates were both 7%. These results are summarised in Box 1.

The make and model of ultrasound machine used was important in influencing the ability to obtain a satisfactory image. Table 4 shows the proportion of pregnancies in which an NT measurement was not obtained and the proportion in which all images for a given pregnancy were technically unsatisfactory, classified according to the ten most commonly used ultrasound machines. With some machines, notably the Acuson 128 (Acuson, USA) and the ATL Ultramark 3000HDI (ATL, USA), failure to obtain an NT measurement was much less common than with other machines such as the Hitachi 525 (Hitachi, Japan) (≤1%, 4% and 23% respectively). The effect of the ultrasound machine was statistically independent of the experience of the sonographers and study period (divided into fifths for each centre) ($p<0.001$). Differences between ultrasound machines were apparent within centres using different machines. They could not therefore be explained by other associated differences between centres. There is likely to have been some between-sonographer confounding (better sonographers using better machines) but generally the difference between machines was too great to be explained by this. For example, two sonographers used both an ATL Ultramark 3000 HDI and a Toshiba Capasee (Toshiba, Japan). Their rates of failing to obtain an NT measurement were 0 and 3.8%, and 0 and 6.5% respectively with the two machines (both were statistically significant differences, $p<0.001$). The choice of ultrasound machine is, therefore, a factor in the performance of screening tests that include an NT measurement.

Box 1 NT measurement: technical results

Failure to obtain an NT image:	
◆ throughout study:	
– 9 completed weeks and before	33%
– at 12 completed weeks	7%
– after 14 completed weeks	32%
◆ at 10–12 completed weeks of gestation:	
– early in study	9%
– late in study	5%
Failure to obtain a "satisfactory" NT measurement:	
◆ at 10 completed weeks	19%
◆ at 11 completed weeks	14%
◆ at 12 completed weeks	15%
◆ at 13 completed weeks	19%
Average time to obtain three NT measurements:	4.7 minutes
Sonographer experience:	
◆ < 200 NT measurements at 10–13 completed weeks of gestation:	
– failure to obtain a satisfactory NT image	19%
◆ ≥ 400 NT measurements at 10–13 completed weeks of gestation:	
– failure to obtain a satisfactory NT image	14%
Ultrasound machine:	
◆ The make and model of the ultrasound machine had an important influence on the ability to obtain a satisfactory NT image.	

Figure 4 shows NT in millimetres for affected pregnancies according to CRL (gestational age is also shown to indicate the equivalence between the two). The median in unaffected pregnancies together with the regressed 5th and

Table 4 NT measurements not obtained or technically unsatisfactory according to ultrasound machine used (10 most commonly used): 10–13 weeks of gestation

Machine	Number of pregnancies	NT measurement could not be obtained in: (%)	NT images unsatisfactory in: (%)[a]	Number of machines
Acuson 128	2728	5 (0.2)	2 (0.1)	3
Aloka 2000[b]	2059	278 (13.5)	90 (5.0)	4
Aloka 650	2391	177 (7.4)	353 (15.9)	4
ATL Ultramark 3000 HDI	2193	12 (0.6)	84 (3.8)	2
ATL Ultramark 9 HDI	3507	211 (6.0)	219 (6.6)	3
Hitachi 525	2961	328 (11.1)	314 (11.9)	7
Hitachi EUB 420	1716	154 (9.0)	200 (12.8)	1
Hitachi Sumi	3406	328 (9.6)	303 (9.8)	3
Toshiba 250	6625	351 (5.3)	566 (9.0)	6
Toshiba Ecossee	2235	202 (9.0)	142 (7.0)	7

[a] As a percentage of the total number of pregnancies minus the number for which an NT measurement was not obtained; [b] Aloka, Japan.

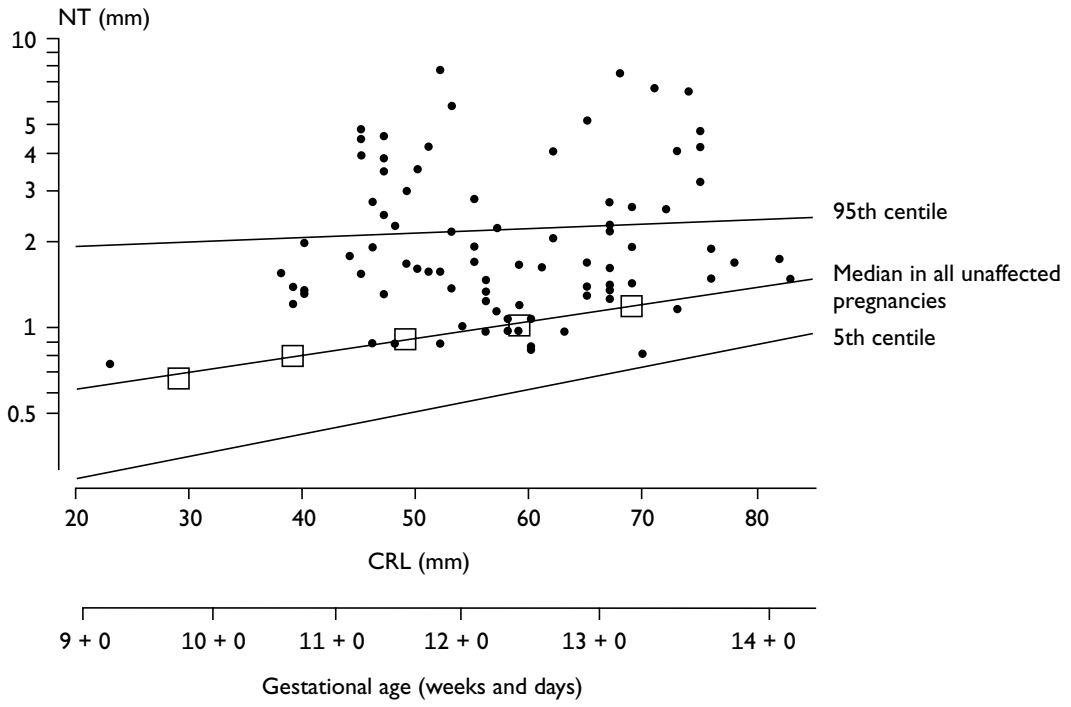

Figure 4 NT measurement (mm) in pregnancies with Down's syndrome (filled circles) according to CRL (and gestational age). The median in unaffected pregnancies (open squares) is shown together with the lines indicating half and twice the median (with corresponding centile values for unaffected pregnancies). The figure is based on all pregnancies together; in the statistical analyses centre-specific or sonographer-specific medians were used to convert NT in millimetres to MoMs.

95th centile lines is also shown. NT increased by 18% per week and the 95th and 5th centile interval decreased with gestation, indicating a progressively smaller SD. Table 5 shows the observed and expected percentages of unaffected pregnancies equal to or greater than specified NT MoM levels – the expected values based on the gestation-specific SDs assuming a Gaussian distribution. The SDs estimated using all the available data fitted the observed data better than those based on the 10th to 90th centile interval, but the resulting Gaussian distribution still tended to underestimate values above about 2.5 MoM.

Figure 5 shows the mean NT MoM value in the 75 Down's syndrome pregnancies (at 10–13 completed weeks) with at least one technically satisfactory ultrasound image and that from the ten affected pregnancies in which all images were judged unsatisfactory. The ten with technically unsatisfactory images had values that were close to the value for unaffected pregnancies (median 1.08 MoM) and they were statistically significantly lower ($p = 0.005$) than those for the 75 affected pregnancies that had at least one satisfactory image (median = 1.91 MoM).

Table 6 shows the effect on screening performance (without using maternal age) of allowing for variation between sonographers in measurement technique (by calculating sonographer-specific medians and converting NT in millimetres to MoMs) and of using only technically satisfactory images. There was about a 5 percentage point increase in detection when sonographer-specific normal NT medians were used instead of centre-specific ones (51% compared with 46%) and a further 7 percentage point increase when the quality of the NT image was always satisfactory. The improvement was not explained by some centres having a better performance than others.

Table 7 shows the effect of taking one, two, or three NT measurements and, when there are multiple measurements, the effect of using the mean, or the largest value. Taking the mean of

Table 5 Observed and expected percentage of unaffected pregnancies[a] equal to or greater than specified NT MoM levels according to gestational age

	Gestational age (completed week of pregnancy)								
	10 (n=3,379)		11 (n=9,051)		12–13 (n=17,346)		10–13 (n=29,776)		
	Observed	Expected	Observed	Expected	Observed	Expected	Observed	Expected	Expected[b]
NT MoM	–	(sd=0.1732)	–	(sd=0.1439)	–	(sd=0.1329)	–	(sd=0.1414)	(sd=0.1174)
≥1.0	54	50	50	50	52	50	52	50	50
≥1.5	12	16	9.1	11	9.3	9.2	9.5	11	6.7
≥2.0	4.7	4.1	3.1	1.8	2.0	1.2	2.6	1.7	0.52
≥2.5	2.4	1.1	1.5	0.28	0.5	0.14	1.0	0.24	0.035
≥3.0	1.8	0.29	0.9	0.05	0.3	0.02	0.6	0.04	0.002
≥3.5	1.3	0.08	0.5	0.008	0.1	0.002	0.4	0.006	<0.001
≥4.0	0.9	0.02	0.3	0.001	0.1	<0.001	0.2	0.001	<0.001

[a] Based on 29,776 unaffected pregnancies known to have reached the second trimester; sonographer-specific MoMs and satisfactory images only; sd=SD($\log_{10}$) without censoring outliers; [b] Expected value based on SD using 10th–90th centile interval.

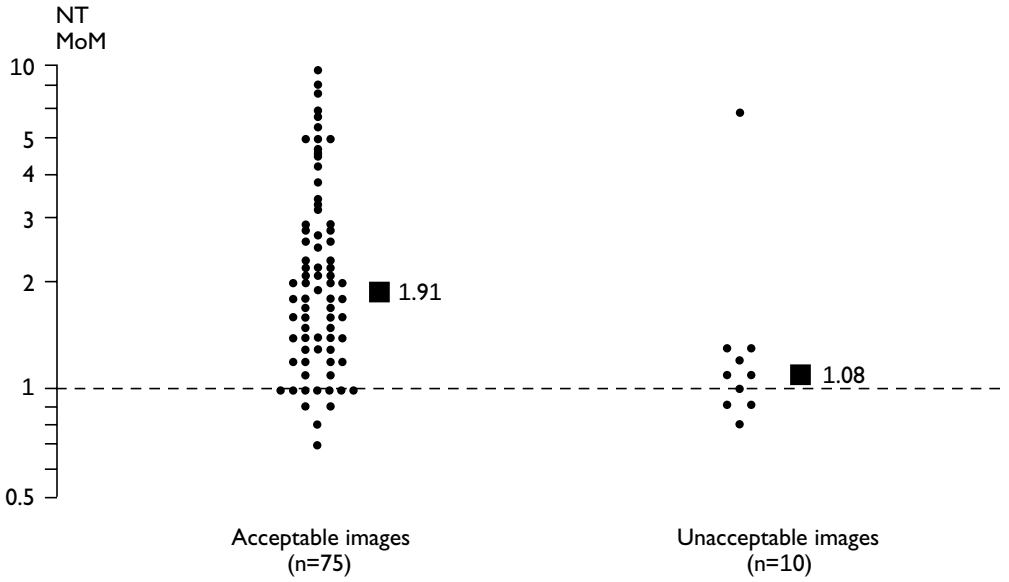

Figure 5 NT measurement (in MoM values) for the 75 Down's syndrome pregnancies with at least one technically satisfactory image and for the ten affected pregnancies where all images were technically unsatisfactory, based on the 85 observed Down's syndrome cases recruited at 10–13 completed weeks of gestation. Filled squares represent the median for each group. Centre-specific medians were used to convert NT measurements in millimetres to MoMs.

Table 6 Observed screening performance using NT measurement at 10–13 completed weeks of gestation (without use of maternal age) according to the method of estimating NT medians in unaffected pregnancies and acceptability of NT images (based on direct observation)

	DR for 5% FPR (%)	FPR for 70% DR (%)
Centre-specific NT medians		
All images	46 (39/85)	20.5
Technically satisfactory images only	52 (39/75)	15.2
Sonographer-specific NT medians[a]		
All images	51 (36/70)	14.2
Technically satisfactory images only	58 (36/62)	9.6

[a] Data restricted to sonographers who had performed NT measurements in at least 100 pregnancies.

three measurements instead of one reduces the FPR for a DR of 70% from 13.0% to 8.8%, and using the mean is better than using the largest of multiple measurements.

In 978 pregnancies in which three NT measurements were obtained, the mean time taken to obtain these was 4.7 minutes (10th–90th centile, 1–9 minutes). The longest times recorded were 20 minutes, the time limit set for the sonographer to obtain an NT measurement; this applied to 9 out of the 978 (1%) pregnancies.

Screening performance based on single markers

Median marker levels in Down's syndrome and unaffected pregnancies at different gestational ages

Figures 6, 7 and 8 show the median marker levels (in MoMs) among affected pregnancies according to gestational age

Table 7 False-positive rate for a 70% detection rate using NT measurement at 10–13 completed weeks of gestation (without use of maternal age) according to number of measurements made and summary estimate used (based on direct observation)

	Sonographer-specific medians and technically satisfactory images only[a]	Centre-specific medians and all images
	FPR (%)	FPR (%)
Number of measurements (mean used if more than one)		
1	13.0	24.1
2	9.8	23.1
3	8.8	20.6
For two or more measurements		
Mean	9.0	20.4
Largest	9.8	21.0

[a] Data restricted to sonographers who had performed NT measurements in at least 100 pregnancies.

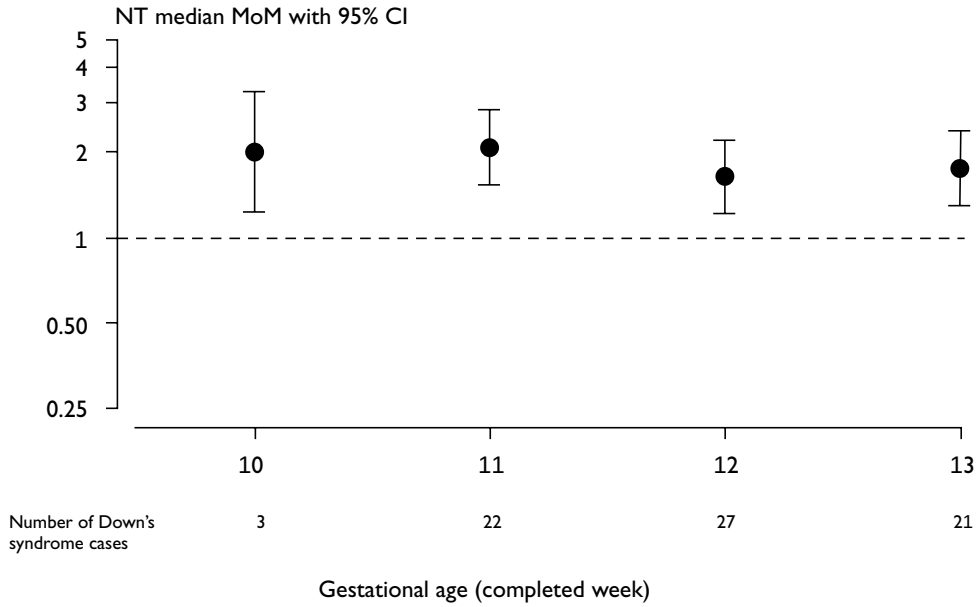

Figure 6 Median NT measurement (in MoMs) among affected pregnancies according to gestational age, together with the 95% CI. The figure is based on 73 cases at 10–13 completed weeks. There were 75 in total, of which two were censored because they were terminated before 14 completed weeks and would have been expected to miscarry. NT has been converted to MoMs using centre-specific medians and technically satisfactory measurements.

(together with the 95% CIs); Figure 6 for NT, Figure 7 for the serum markers, and Figure 8 for the urine markers. The median NT in affected pregnancies remained high from 10–13 completed weeks with no evidence of a trend (p=0.34) over this period. Serum AFP (see Figure 7) was low from 14 weeks but less so at earlier gestations. The median AFP in affected pregnancies was 0.86 (95% CI, 0.78 to 0.95) MoM in the first trimester and 0.74 (95% CI, 0.67 to 0.82) in the second trimester. There may well be a small trend in the median AFP in affected pregnancies but because the observed medians in our data were similar and Figure 7 does not show a clear trend, there was insufficient power to find such a trend. The overall median for each trimester was therefore used in the analyses. Of the other five serum markers, four showed a statistically significant trend with gestation in the first trimester; total hCG (p=0.005), free β-hCG (p=0.037), inhibin-A (p=0.008) and PAPP-A (p=0.05); hCG and inhibin-A became progressively more discriminatory and PAPP-A less discriminatory. Discrimination of PAPP-A before 10 weeks may be better than at 10 weeks but no conclusions on this could be drawn because of lack of data. The trend for uE_3 apparent in Figure 7 was not formally statistically significant (p=0.13) but was judged to be real because of the pattern of results which continued into the second trimester. Because of these results and because the SD of $\log_{10}$NT in unaffected pregnancies decreased with increasing gestational age (0.1732, 0.1439, 0.1329 at 10, 11 and 12–13 completed weeks respectively; 12–13 weeks were combined since the SDs at 12 and 13 weeks were similar), the first trimester was not considered as a single period, but gestation-specific medians were used for the six specified biochemical markers (see Table 8) and gestation-specific SDs used for NT. After 13 weeks, hCG and inhibin-A remained consistently high in affected pregnancies but PAPP-A levels were not materially different from those in unaffected pregnancies.

There was also a trend for urinary ITA medians in affected pregnancies to increase from 10–13 weeks. The pattern was less clear for the other three urinary markers (see Figure 8) with virtually no discriminatory effect before 13 weeks.

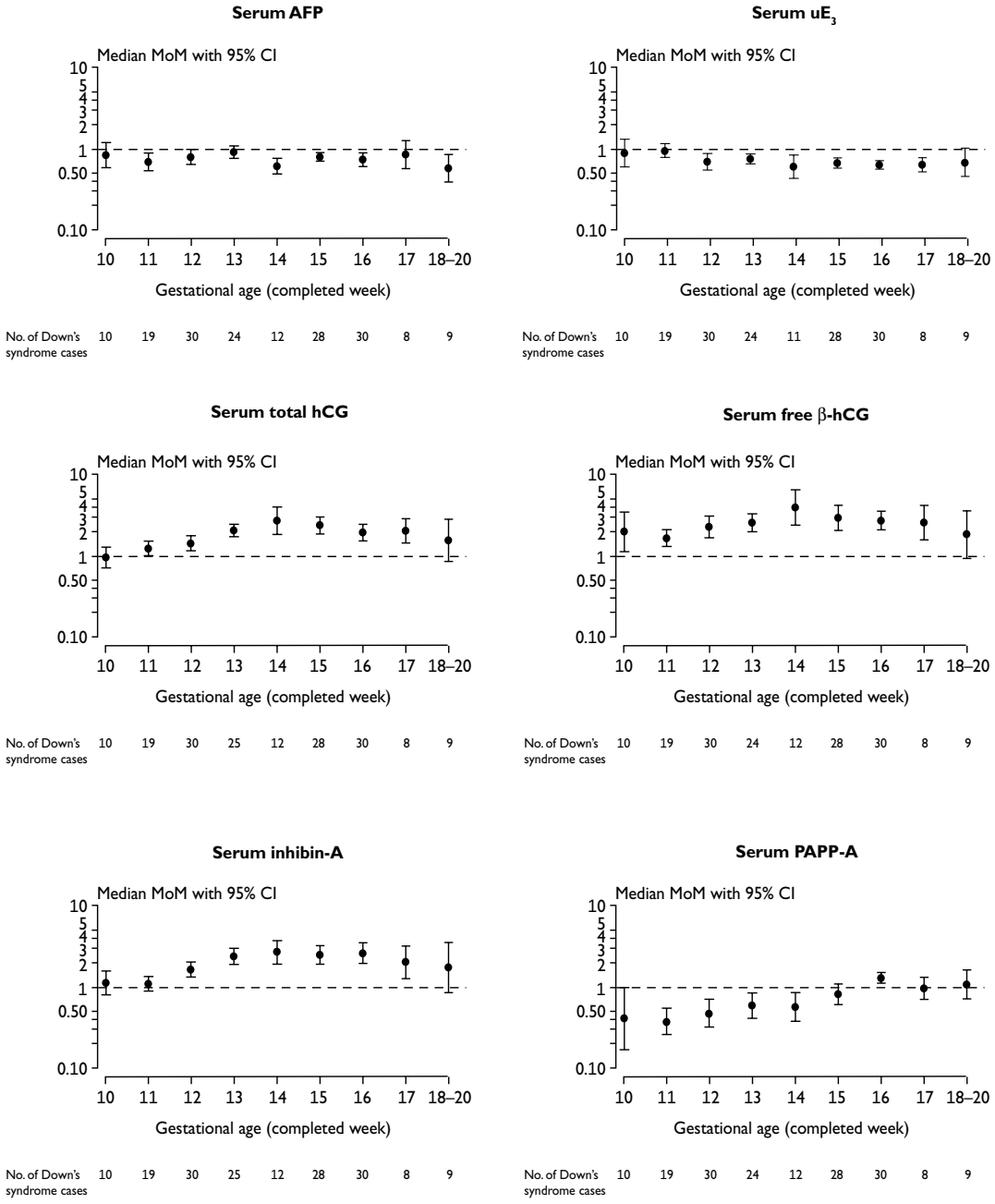

Figure 7 Median serum marker levels (in MoMs) among affected pregnancies according to gestational age, together with the 95% CI (excludes the three expected livebirths amongst screen-negatives, see section "Data collected and data used" pages 60–1, that were randomly added).

Table 8 Expected and observed median marker level (MoM) for the six markers for which the median MoM in Down's syndrome pregnancies showed an increasing or decreasing trend with gestation in the first trimester

	Median MoM in Down's syndrome pregnancies Completed week of pregnancy			
Marker	10	11	12	13
Serum				
uE_3	0.99 (0.94)	0.87 (0.99)	0.79 (0.72)	0.72 (0.78)
Total hCG	0.96 (0.96)	1.27 (1.24)	1.54 (1.45)	1.88 (2.07)
Free β-hCG	1.62 (1.94)	1.94 (1.61)	2.19 (2.22)	2.48 (2.50)
Inhibin-A	0.94 (1.14)	1.35 (1.12)	1.73 (1.62)	2.22 (2.44)
PAPP-A	0.34 (0.42)	0.42 (0.38)	0.50 (0.44)	0.58 (0.60)
Urine				
ITA	1.04 (1.29)	1.71 (1.67)	2.43 (1.97)	3.44 (3.71)

The expected medians (obtained from a regression of the observed medians with gestational age) are shown with the observed medians in brackets.

Screening performance of markers individually

Our estimates of screening performance using NT measurement are based on NT MoM values calculated using sonographer-specific medians with all images judged to be technically satisfactory. Gestation-specific SDs were used, and in estimating the risk of having an affected pregnancy an upper truncation limit of 2.5 MoM was used.

The median NT in affected pregnancies was 1.96 MoM (95% CI, 1.63 to 2.35), less than the observed value in the Fetal Medicine Foundation Study (the only other large study on the screening efficacy of NT) (2.27 MoM) but similar to the value (2.02 MoM) after it was adjusted for spontaneous fetal loss of affected pregnancies to term.[14] Our estimate of 1.96 MoM allows for spontaneous fetal loss up to the early second trimester of pregnancy (both the background excess fetal loss in Down's syndrome pregnancies and the excess fetal loss associated with high NT

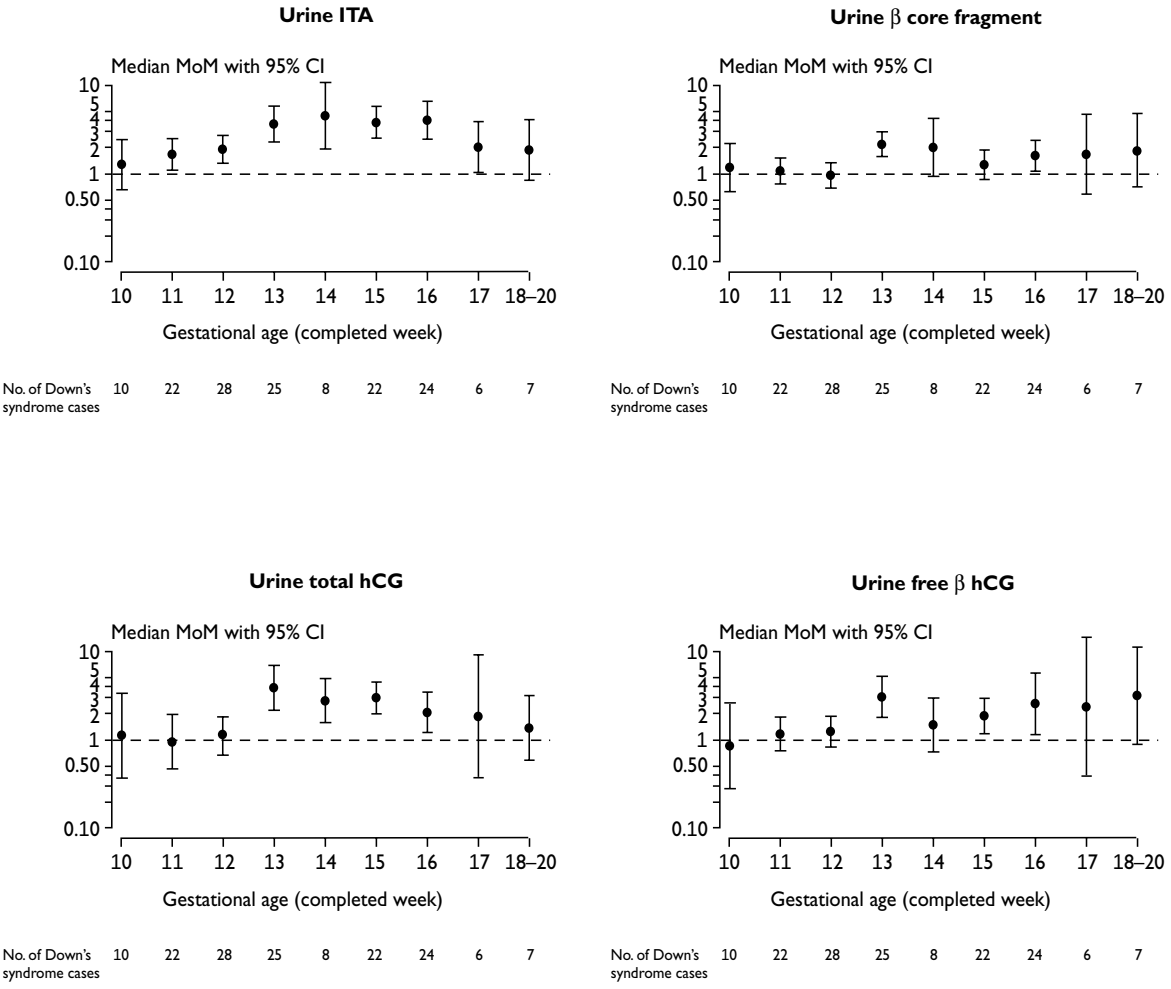

Figure 8 Median urine marker levels (in MoMs) among affected pregnancies according to gestational age, together with the 95% CI (excludes the three expected livebirths amongst screen-negatives, see section "Data collected and data used" pages 60–1, that were randomly selected).

measurements); see the section "Data collected and data used", pages 60–1 and so is accurate given that our objective in SURUSS was to estimate second trimester risks, rather than term risks. In the SURUSS population the observed DR for a 5% FPR using NT and maternal age was 63% (95% CI, 51 to 75%), close to the estimate modelled to the standard population (see chapter 2) of 67%. Similarly, the observed FPR for an 85% DR was 19.1% (95% CI, 18.7 to 19.5%), close to the modelled estimate of 21.1%. Our overall results on NT are therefore consistent with previous data and the modelled estimates are close to those observed. Our modelled estimate (67%) is also comparable to the Fetal Medicine Foundation estimate that allows for spontaneous fetal loss (73%).[14]

Table 9 shows the median levels in affected pregnancies for all the markers at 10 completed weeks and at 14–20 weeks together with the DR for a 5% FPR using each marker alone (using age alone the DR was about 34% for a 5% FPR). NT and PAPP-A were the most discriminatory markers at 10 completed weeks, followed by second trimester free β-hCG and inhibin-A at 14–20 weeks. The median inhibin-A value in affected pregnancies after 14 weeks was unexpectedly high (2.54 MoM). A study based on the Bart's screening programme (Wald N: personal observation) which, like SURUSS, was interventional in the second trimester, yielded an estimate of 2.18 MoM based on 111 affected pregnancies, and this is close to the interim unpublished estimate from a large USA study (the FASTER (first and second trimester evalu-ation of risk) project; Canick J, Women and Infants' Hospital, Rhode Island, USA: personal communication, 2002). We have no explanation for the difference other than chance. Since the use of the SURUSS estimate would overestimate screening performance of tests in which inhibin-A was a component it was felt necessary to adopt the independent estimate of 2.18 MoM. The median inhibin-A in affected pregnancies was 2.22 MoM at 13 completed weeks, higher than the estimate we used for 14 weeks and later (2.18 MoM) so we also took the median at 13 weeks to be 2.18 MoM as there is no reason to believe that levels decline at this time and rise again. Other than inhibin-A the SURUSS estimates of the other markers in affected pregnancies were similar to those in the published literature.

Appendix Table 33 gives the median $\log_{10}$MoM values for markers in affected pregnancies according to gestational age (with regression equation to estimate day specific medians) and Tables 34 and 35 give the SDs ($\log_{10}$) for the markers in affected and unaffected pregnancies together with the

truncation limits used. Table 10 shows the DR for a 5% FPR for the markers at 10, 11, 12, and 13 completed weeks of pregnancy. While some markers improved over this period (for example, NT), PAPP-A became worse.

Table 9 Median MoM values in Down's syndrome pregnancies and detection rate for a 5% false-positive rate for individual markers (without use of maternal age) according to gestation

	Completed week of pregnancy			
	10[a]		14–20	
Marker	Median (MoM)	DR (%)	Median (MoM)	DR (%)
Ultrasound				
NT	1.96	51	–	–
Serum				
AFP	0.86	8	0.74	24
uE_3	0.99	13	0.70	40
Total hCG	0.96	5	2.05	40
Free β-hCG	1.62	19	2.66	50
Inhibin-A	0.94	5	2.54 (2.18)[b]	59 (49)[b]
PAPP-A	0.34	58	1.11	5
Urine				
ITA	1.04	6	3.51	40
β-core fragment	1.10	11	1.37	14
Total hCG	1.38	17	2.18	31
Free β-hCG	1.28	5	1.99	13

[a] The medians for serum AFP and urine β-core fragment, urine total hCG and urine free β-hCG are based on 10–13 completed weeks together because they do not change materially with gestation. The SD of NT values in unaffected pregnancies is based on the one at 10 weeks; [b] Using the estimate from the Bart's screening service.

Table 10 Detection rate (%) for 5% false-positive rate (without use of maternal age) using markers for which the median in Down's syndrome pregnancies increases or decreases with gestation in the first trimester

	Completed week of pregnancy			
	10	11	12	13
Marker	DR (%)			
Ultrasound				
NT	51	59	62	62
Serum				
uE_3	13	21	29	37
Total hCG	5	15	26	41
Free β-hCG	19	28	35	44
Inhibin-A	5	16	30	48 (46)[a]
PAPP-A	58	45	35	27
Urine				
ITA	6	16	28	43

[a] Using the median of 2.18 based on the Bart's screening service.

Screening performance of markers in combination

Tables 11, 12 and 13 show the screening performance of the main markers. The screening performance of various tests was compared by estimating the FPR for a given DR and estimating the DR for a given FPR. Conventionally, the FPR is fixed, but this can conceal differences between tests when DRs are high. It is then easier to assess the relative performance of different tests by fixing the DR and comparing the FPRs.

Most of the markers were materially independent of each other. Appendix Tables 36–43 show the correlation coefficients for all combinations of the markers studied, excluding the urinary markers. Markers that were correlated included the same marker (at different times in pregnancy) or related markers (for example total hCG and free β-hCG) measured in the two trimesters, and hCG (total or free β-hCG) and inhibin-A.

The effect of using all the serum markers together is shown, so that the maximum screening performance can be seen and used to assess the value of the other combinations.

Table 11 shows the screening performance of the main first trimester markers at 10 completed weeks of pregnancy. AFP and uE_3 were excluded from most combinations because of their poor screening performance at this time in pregnancy. (Appendix Table 44 shows the estimates for specified markers at each week from 10 to 13 weeks.) Table 12 shows in the same way as Table 11 screening performance using second trimester serum markers alone and in combination (double, triple and quadruple tests, as well as the quad-ruple test with the addition of PAPP-A). Table 13 (first and second trimester markers) shows the results using the integrated test and its variants when the first markers were measured at 10 completed weeks. This table has restricted the use of markers to the first or second trimester, depending on when they are more discriminatory – so, for example, inhibin-A, which discriminates between affected and unaffected pregnancies better in the second trimester than in the first trimester, was used in the second trimester only. Appendix Table 45 shows the corresponding estimates when the first trimester markers including NT were measured at 12 completed weeks of pregnancy.

Table 14 shows the screening performance of the urinary marker ITA alone and in combination with other tests, together with the performance of those tests without ITA so that the incremental effect of adding ITA can be assessed. The performance of the other urinary markers (β-core fragment, total hCG, free β-hCG) is also shown with and without the addition of ITA. Three conclusions emerge from the table. First, the urinary markers are virtually useless in the first trimester. Secondly, of the urinary markers, ITA is the best, and little is added by using any or all of the other three markers as well. Thirdly, while ITA adds to screening performance the effect is modest.

Tables 11–14 present the screening performance of a reasonably comprehensive selection of marker combinations. The tests are ranked in groups defined by the number of markers used. Using only one or two markers has a relatively poor screening performance, and using all the markers adds little to a lesser number. In the first trimester the combined test had an estimated 83% DR for a 5% FPR. Without NT measurement, the combination of PAPP-A and free β-hCG had a 74% DR for a 5% FPR. In the second trimester the quadruple test (using free β-hCG , since this was a little better than with total hCG) was the test of choice, yielding an 83% DR for a 5% FPR.

Table 11 Screening performance using the first trimester markers at 10 completed weeks of pregnancy

Maternal age with:	DR (%) for FPR of:			FPR (%) for DR of:						DR (%) for FPR of:			FPR (%) for DR of:					
	1%	3%	5%	70%	75%	80%	85%	90%	95%	1%	3%	5%	70%	75%	80%	85%	90%	95%
	With NT									Without NT								
NT	33	50	60	9.2	12.9	18	25	36	55	–	–	–	–	–	–	–	–	–
Inhibin-A	34	51	61	9.2	12.8	18	25	36	55	12.0	24	32	35	43	52	62	73	86
Total hCG	34	51	61	9.2	12.8	18	25	36	55	11.8	24	32	35	43	52	62	74	86
AFP	35	53	61	8.5	11.8	16	23	34	52	13.0	26	33	32	40	48	57	67	80
Free β-hCG	41	57	65	7.0	9.8	13.9	20	30	48	17	30	39	24	29	36	44	55	70
PAPP-A	59	73	79	2.3	3.5	5.3	8.4	13.9	26	43	59	67	6.0	8.3	11.7	16	25	39
Total hCG, inhibin-A	35	52	62	8.7	12.1	17	24	35	54	13.2	26	34	32	40	48	59	70	84
Free β-hCG, AFP	42	58	66	6.4	9.0	12.8	19	28	45	19	32	41	22	27	34	42	53	68
Free β-hCG, inhibin-A	44	59	66	6.3	9.0	12.9	19	29	46	22	36	44	20	25	32	40	52	68
PAPP-A, total hCG	60	74	80	2.2	3.3	5.1	8.0	13.4	26	44	60	68	5.8	8.0	11.3	16	24	39
PAPP-A, AFP	60	74	80	2.2	3.3	5.0	7.9	13.3	25	45	60	68	5.7	8.0	11.2	16	24	39
PAPP-A, inhibin-A	60	74	80	2.2	3.3	5.0	7.9	13.3	25	44	60	68	5.8	8.0	11.3	16	24	39
PAPP-A, free β-hCG	66	78	83	1.5	2.3	3.7	6.1	10.8	22	52	67	74	3.8	5.5	8.0	12.1	19	33
PAPP-A, total hCG, inhibin-A	61	75	81	2.0	3.1	4.7	7.5	12.7	24	45	61	69	5.4	7.6	10.7	15	23	38
PAPP-A, free β-hCG, inhibin-A	66	78	83	1.4	2.2	3.5	5.9	10.4	21	53	68	74	3.6	5.2	7.6	11.5	18	32
PAPP-A, free β-hCG, AFP	67	78	84	1.4	2.2	3.5	5.8	10.4	21	53	67	74	3.6	5.3	7.8	11.8	19	32
PAPP-A, total hCG, inhibin-A, AFP, uE_3	65	78	83	1.5	2.3	3.6	5.9	10.2	20	50	66	73	4.0	5.7	8.2	12.0	18	31
PAPP-A, free β-hCG, inhibin-A, AFP, uE_3	70	81	86	1.1	1.7	2.7	4.6	8.3	17	56	71	78	2.8	4.0	5.9	9.0	14.2	25

The tests are ordered according to screening performance at an 85% detection rate in groups categorised by number of markers used. Screening performance for NT is based on sonographer-specific medians and technically satisfactory images. False-positive rates ≥15% are rounded to the nearest whole percentage.

Table 12 Screening performance using second trimester (14–20 completed weeks) markers

	DR (%) for FPR of:			FPR (%) for DR of:					
Maternal age with:	1%	3%	5%	70%	75%	80%	85%	90%	95%
PAPP-A	12	25	33	34	41	50	60	70	82
AFP	19	33	42	21	26	32	40	51	66
uE_3	29	44	52	13.0	17	22	29	39	56
Total hCG	29	44	53	12.5	16	21	28	38	54
Inhibin-A	38	52	59	9.8	13.5	19	26	37	56
Free β-hCG	36	52	61	8.9	11.9	16	22	32	49
AFP, total hCG	40	57	66	6.4	8.6	11.7	16	23	37
AFP, free β-hCG	46	63	71	4.8	6.5	9.1	13.1	19	32
AFP, uE_3, total hCG	51	67	74	3.7	5.2	7.4	10.9	17	28
AFP, uE_3, free β-hCG	56	70	77	2.9	4.2	6.2	9.3	14.7	26
AFP, uE_3, total hCG, inhibin-A	62	75	81	2.0	2.9	4.5	7.1	11.7	22
AFP, uE_3, free β-hCG, inhibin-A	64	77	83	1.6	2.5	3.9	6.2	10.6	20
AFP, uE_3, total hCG, inhibin-A, PAPP-A	63	77	83	1.7	2.6	3.9	6.1	10.1	18
AFP, uE_3, free β-hCG, inhibin-A, PAPP-A	66	79	84	1.4	2.2	3.4	5.4	9.0	17

The tests are ordered according to screening performance at an 85% detection rate in groups categorised by number of markers used. Screening performance for NT is based on sonographer-specific medians and technically satisfactory images. False-positive rates ≥ 15% are rounded to the nearest whole percentage.

Overall the most effective test was the integrated test with a 93% DR for a 5% FPR. At a DR of 85% the FPR was 1.2% (95% CI, 1.0 to 1.4, or 2.7%, 95% CI, 2.4 to 3.0 without NT measurement). The FPR was higher with either the first trimester combined test (6.1%, 95% CI, 5.6 to 6.5) or the second trimester quadruple test (6.2%, 95% CI, 5.8 to 6.6).

Table 15 shows the FPR for an 85% DR for selected first trimester and integrated tests according to the gestational week in which the first trimester serum sample was collected. With the integrated and the serum integrated tests, screening performance was better when the serum sample was collected at 10 completed weeks. With the standard combined test based on NT, PAPP-A and free β-hCG levels, screening performance was similar at all gestation weeks. This is because the reduction in screening performance due to the loss of discrimination of PAPP-A is compensated for by the increase in discrimination of free β-hCG.

The performance of the combined test, if performed at 12 or 13 completed weeks of pregnancy, would be enhanced by the additional measurement of inhibin-A. For example, at 12 weeks this would yield a 3.8% FPR for an 85% DR compared with a 6.0% FPR with the standard combined test (see appendix, Table 44). However, this would still be significantly less effective than an integrated test with NT and PAPP-A measured at 12 weeks (a 1.3% FPR for the same DR).

Table 16 shows the FPRs for an 85% DR together with the 95% CIs for the main screening tests in decreasing order of screening performance. Box 2 lists the screening tests that can no longer be regarded as worthwhile, in the light of the SURUSS efficacy, safety and cost results.

NT images that were judged to be satisfactory significantly improved screening performance (see page 65). Table 17 shows the effect on screening performance according to specified proportions of affected pregnancies in which the NT measurement is unsatisfactory. The loss in screening performance was smallest when an NT measurement was combined with other markers.

Box 2 Screening tests that in the light of the SURUSS results are no longer worthwhile

First trimester

- NT measurement (with or without maternal age) with no serum marker.

Second trimester

- double test
- triple test.

Influence of method of estimating gestational age

Table 18 shows the FPR for an 85% DR according to screening test, method of estimating gestational age and whether the markers were adjusted for maternal weight. For all tests the FPR was lower when gestational age was estimated using an ultrasound scan examination (by about 2 percentage points). The effect of maternal weight adjustment was smaller (reducing the FPR by about half a percentage point).

Markers and spontaneous fetal loss

Table 19 shows the association between the screening markers and spontaneous fetal loss (most of which would have been unaffected), based on the nested case–control analysis. The NT results are based on all pregnancies. All the measured markers were associated with spontaneous fetal loss. In the first trimester a large NT measurement and low levels of the biochemical markers were associated with spontaneous fetal loss. In the second trimester a high AFP and a low uE_3 were associated with fetal loss, as were both high and low levels of total hCG, free β-hCG, and inhibin A. Some of the effects were large. For example, at or above the 99th centile the odds ratio of a fetal loss for NT was 6.6, and in the second trimester 11 for AFP, 7.6 for total hCG, 5.1 for free β-hCG, and 10 for inhibin-A compared with having a measurement below the 99th centile. At or below the first

Table 13 Performance of integrated screening using first (10 completed weeks) and second trimester (14–20 completed weeks) markers

Maternal age with:		With NT									Without NT								
		DR (%) for FPR of:			FPR (%) for DR of:						DR (%) for FPR of:			FPR (%) for DR of:					
1st trimester	2nd trimester	1%	3%	5%	70%	75%	80%	85%	90%	95%	1%	3%	5%	70%	75%	80%	85%	90%	95%
	AFP, total hCG	61	76	82	2.0	2.9	4.3	6.7	11.0	21									
	AFP, free β-hCG	65	78	83	1.5	2.4	3.7	5.8	10.0	19									
	AFP, uE_3, total hCG	70	82	87	1.0	1.6	2.5	4.1	7.1	14.4									
	AFP, uE_3, free β-hCG	72	83	88	0.8	1.4	2.2	3.7	6.7	14.2									
	AFP, uE_3, total hCG, inhibin-A	76	86	90	0.6	0.9	1.6	2.8	5.2	11.7									
	AFP, uE_3 free β-hCG, inhibin-A	77	86	90	0.5	0.8	1.4	2.5	4.9	11.4									
PAPP-A	AFP, total hCG	76	86	90	0.6	0.9	1.5	2.7	5.2	11.7	65	78	83	1.6	2.4	3.7	6.0	10.3	20
PAPP-A	AFP, free β-hCG	78	87	90	0.4	0.7	1.3	2.4	4.8	11.4	68	80	85	1.2	1.9	3.0	5.1	9.1	18
PAPP-A	AFP, uE_3, total hCG	79	88	92	0.4	0.6	1.1	2.0	3.9	9.4	69	80	85	1.1	1.8	2.8	4.8	8.5	17
PAPP-A	AFP, uE_3, free β-hCG	80	89	92	0.3	0.5	0.9	1.8	3.7	9.3	71	82	86	0.9	1.4	2.4	4.2	7.7	16
PAPP-A	AFP, uE_3, total hCG, inhibin-A	83	90	93	0.2	0.4	0.7	1.3	2.8	7.4	75	85	89	0.6	1.0	1.7	3.0	5.8	13.2
PAPP-A	AFP, uE_3, free β-hCG, inhibin-A	84	91	93	0.2	0.3	0.6	1.2	2.6	7.2	77	86	90	0.5	0.8	1.5	2.7	5.3	12.5

The tests are ordered according to screening performance at an 85% dectection rate in groups categorised by number of markers used. Screening performance for NT is based on sonographer-specific medians and technically satisfactory images. False-positive rates ≥15% are rounded to the nearest whole percentage.

Table 14 Screening performance using urine markers with maternal age. False-positive rate (%) for an 85% detection rate

	FPR (%) for an 85% DR	
Markers other than ITA	With ITA[a]	Without ITA[a]
First trimester (10 completed weeks)		
None	63 (38)	–
NT	25 (12.1)	25 (20)
Combined test	5.9 (5.1)	6.1 (6.0)
Urine: β-core fragment, total hCG, free β-hCG	33 (16.5)	52
Second trimester (14–20 completed weeks)		
None	26	–
Double test	9.9	13.1
Triple test	6.7	9.3
Quadruple test	4.6	6.2
Urine: β-core fragment, total hCG, free β-hCG	9.8	29
Integrated test:[b]		
With NT	0.7 (0.8)	1.2 (1.3)
Without NT (serum integrated test)	1.8 (3.5)	2.7 (4.9)

[a] Estimates in brackets relate to 12 completed weeks of pregnancy. [b] ITA used in the second trimester.

Table 15 The effect of using first trimester gestation-specific medians on the screening performance of specified tests: false-positive rate (%) for 85% detection rate[a]

Test	Markers		Completed weeks of pregnancy				
Maternal age with:	First trimester markers	Second trimester[b]	10	11	12	13	10–13 weeks[c]
Combined test	NT, PAPP-A, free β-hCG	–	6.1	6.0	6.0	5.8	6.0
Combined test variants	NT, PAPP-A	–	8.4	10.5	13.0	15.7	12.3
	NT, free β-hCG	–	20.1	13.7	10.3	8.1	12.0
	NT, PAPP-A, free β-hCG, inhibin-A	–	5.9	5.0	3.8	2.5	4.1
Integrated test	NT, PAPP-A	Double test	2.4	2.7	3.0	3.5	3.0
	NT, PAPP-A	Triple test	1.8	1.9	2.0	2.2	2.0
	NT, PAPP-A	Quadruple test	1.2	1.2	1.3	1.5	1.3
Serum integrated test	PAPP-A	Double test	5.1	7.6	9.6	11.1	8.8
	PAPP-A	Triple test	4.2	6.0	7.5	8.5	6.9
	PAPP-A	Quadruple test	2.7	3.9	4.9	5.6	4.5

[a] Using separate gestation-specific Down's syndrome medians and NT SDs at each week from 10–13 completed weeks (for those markers listed in Table 8); [b] Free β-hCG is used; [c] Weighted average of the estimates at each week; weighted by the percentage of women who book at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%) (as observed in SURUSS). These estimates of false-positive rate will vary according to local centre, because the proportion of pregnancies screened at each week of pregnancy will vary.

Table 16 Screening performance of the main screening tests

Maternal age with:[a]	FPR (%) for an 85% DR	95% CI
Integrated test	1.2	1.0–1.4
Serum integrated test	2.7	2.4–3.0
Combined test	6.1	5.6–6.5
Quadruple test	6.2	5.8–6.6
Triple test	9.3	8.8–9.8
Double test	13.1	12.5–13.7
NT measurement:		
At 10 completed weeks	25.1	24.0–26.2
At 12–13 completed weeks	20.0	18.6–21.4

[a] Free β-hCG is used in the first and second trimesters.

Table 17 The effect of technically satisfactory and unsatisfactory NT measurements on screening performance

	Percentage of Down's syndrome pregnancies in which all the NT measurements are unsatisfactory									
	0		5		10		15		20	
Maternal age with:	DR (%) for 5% FPR	FPR (%) for 85% DR	DR (%) for 5% FPR	FPR (%) for 85% DR	DR (%) for 5% FPR	FPR (%) for 85% DR	DR (%) for 5% FPR	FPR (%) for 85% DR	DR (%) for 5% FPR	FPR (%) for 85% DR
NT measurement										
10 weeks	60	25	58	28	56	30	54	33	52	36
12–13 weeks	69	20	66	23	64	26	61	30	59	32
Combined test										
10 weeks	83	6.1	82	6.6	82	7.1	81	7.5	80	7.9
12 weeks	83	6.0	82	6.6	81	7.3	80	8.0	79	8.8
Integrated test										
10 weeks[a]	93	1.2	93	1.2	93	1.4	93	1.4	92	1.6
12 weeks[a]	93	1.3	92	1.5	92	1.7	91	1.9	91	2.1

Weeks are completed weeks of pregnancy and free β-hCG was used in the second trimester; [a] First trimester markers measured at this time in pregnancy.

Table 18 False-positive rate (%) for an 85% detection rate for different tests according to method of estimating gestational age, and maternal weight correction of serum markers

	Method of estimating gestational age			
	Dates (LMP)		Ultrasound scan (CRL or BPD)	
	Maternal weight correction?		Maternal weight correction?	
Test	No	Yes	No	Yes
Combined[a]	8.5	7.8	6.8	6.1
Triple	11.8	11.6	9.6	9.3
Quadruple	7.8	7.6	6.4	6.2
Serum integrated[a]	4.6	4.2	3.2	2.7
Integrated[a]	2.2	2.0	1.4	1.2

[a] First trimester markers measured at 10 weeks.

Table 19 The odds ratio of having a pregnancy associated with a fetal loss[a] compared with having a livebirth, according to whether the marker level is at or more extreme than specified centiles compared with being less extreme (centiles are those of the distributions for unaffected livebirths)

	Centile			
Marker	≤1st	≤5th	≥95th	≥99th
First trimester				
NT[b]	0.14	0.66	2.0	6.6
AFP	2.2	1.5	2.2	1.9
uE_3	4.8	1.6	1.4	1.4
Total hCG	4.0	2.2	1.0	1.7
Free β-hCG	3.7	1.9	1.6	1.7
Inhibin-A	4.0	2.2	1.7	2.2
PAPP-A	8.9	4.3	0.82	1.1
Second trimester				
AFP	1.3	0.77	6.3	11
uE_3	12	5.2	0.97	1.3
Total hCG	4.1	2.4	4.9	7.6
Free β-hCG	3.6	1.7	2.7	5.1
Inhibin-A	4.6	2.5	4.7	10
PAPP-A	2.7	2.8	1.2	2.2

[a] Miscarriage or stillbirth; [b] Relative risks are shown (and are based on all singleton pregnancies). The odds ratios are given to two significant figures.

centile, the odds ratio was 8.9 for first trimester PAPP-A and 12 for second trimester uE_3 compared with being above the first centile.

These data confirm previous observations that NT[15] and serum markers[16] used in Down's syndrome screening are associated with natural fetal loss, so that these markers will be more likely to identify affected pregnancies that will later miscarry than those that will not. Our study avoided bias from this source in comparing first and second trimester markers but could not exclude it after the second trimester; hence our estimates of screening performance relate to the second trimester.

Safety

Table 20 shows the number of fetal losses per 100,000 women screened in programmes using the main tests according to different DRs. For example, the integrated test offers the safest screening and antenatal diagnosis programme, with nine losses at an 85% DR compared with 94, 67, and 45 with the double, triple, and quadruple tests respectively, and 180 with NT (at 10 completed weeks) and maternal age. The double test and NT alone are the least safe methods of screening. The integrated test had about one-fifth of the fetal losses compared with the combined test (9 versus 44). If the excess risk of fetal loss were 50% higher with CVS than with mid-trimester amniocentesis, the fetal loss rates in Table 19 would have to be increased by 50% when using CVS as the diagnostic procedure. So, for example, the estimate of 44 procedure-related fetal losses per 100,000 women screened in Table 20 for the combined test at a DR of 85% would increase to 66. The table also shows the number of Down's syndrome pregnancies detected for each procedure-related fetal loss. Appendix Table 46 shows the corresponding FPRs and appendix Table 47 shows the number of fetal losses for selected first trimester and integrated tests at each week of pregnancy separately and at 10–13 completed weeks combined (Table 47). The risk cut-off levels used to achieve the screen-ing results illustrated in Table 20 are shown in the bottom section of the table. Because of the different distributions of risk estimates in Down's syndrome and unaffected pregnancies, very different risk cut-off levels are needed to achieve the same DR with the different tests. For example, to achieve an 85% DR with the quadruple test requires a risk cut-off level of 1 in 300, but 1 in 100 with the integrated test.

Table 20 Safety and false-positive rates of the main screening tests at specified detection rates

	DR (%)			
Screening test[a]	75	80	85	90
	Number of Down's syndrome pregnancies diagnosed antenatally in 100,000 women screened[b]			
	152	163	173	183
	FPR (%)			
Double	6.5	9.1	13.1	19.5
Triple	4.2	6.2	9.3	14.7
Quadruple	2.5	3.9	6.2	10.6
NT at 10 completed weeks	12.9	18.0	25.1	36.6
NT at 12–13 completed weeks	8.6	13.0	20.0	32.4
Combined[c]	2.3	3.7	6.1	10.8
Serum integrated[c]	0.8	1.5	2.7	5.3
Integrated[c]	0.3	0.6	1.2	2.6
	Number of procedure-related unaffected fetal losses in 100,000 women screened[d]			
Double	47	65	94	140
Triple	30	45	67	106
Quadruple	18	28	45	76
NT at 10 completed weeks	93	129	180	263
NT at 12–13 completed weeks	62	93	144	233
Combined[c]	17	27	44	78
Serum integrated[c]	6	11	19	38
Integrated[c]	2	4	9	19
	Number of Down's syndrome pregnancies detected for each procedure-related unaffected fetal loss[e]			
Double	3.2	2.5	1.8	1.3
Triple	5.1	3.6	2.6	1.7
Quadruple	8.5	5.8	3.8	2.4
NT at 10 completed weeks	1.6	1.3	1.0	0.7
NT at 12–13 completed weeks	2.5	1.7	1.2	0.8
Combined[c]	9.0	6.0	3.9	2.3
Serum integrated[c]	25.4	14.8	9.1	4.8
Integrated[c]	76.3	40.7	19.2	9.6
	Risk cut-off level to achieve specified DR (mid-trimester)[f] 1 in:			
Double	205	300	470	785
Triple	160	240	385	690
Quadruple	105	175	300	580
NT at 10 completed weeks	420	520	895	1245
NT at 12–13 completed weeks	370	500	930	1385
Combined[c]	115	190	335	670
Serum integrated	50	85	170	385
Integrated[c]	25	45	100	245

[a] All tests include maternal age and free β-hCG is used rather than total hCG; [b] There is a total of 226 in the second trimester, based on the expected birth prevalence of 1.74 per 1000 for England & Wales 1996–98.[7,9,11] The estimates in the table assume a 90% uptake rate of amniocentesis or CVS (higher than the 80% uptake in unaffected pregnancies because women with affected pregnancies tend to have higher risks and so are more likely to accept diagnostic testing[2]); [c] The first trimester markers PAPP-A and free β-hCG are based on the median in Down's syndrome pregnancies at 10 completed weeks and the NT SD in unaffected pregnancies is applicable to 10 completed weeks; [d] An 80% uptake rate of amniocentesis or CVS is used and a 0.9% fetal loss rate attributable to the procedure[2]; [e] Obtained by dividing the number of procedure-related unaffected fetal losses in 100,000 women screened (as above) by the number of Down's syndrome pregnancies detected (as above); [f] 1 in x where x is rounded to the nearest 5. See appendix, Tables 46 and 47 for results based on first trimester markers measured at 12 completed weeks.

Financial costs

Table 21 shows the UK unit cost estimates (in £) that apply to a public service NHS screening programme in 2001 (service A). It also shows two other cost illustrations (services B and C), in which the reagent costs are the same as A, but the diagnostic costs, termination and delivery costs are 1.5 and 2 times greater and all other costs are 4 and 6 times greater respectively. These multiples, while somewhat arbitrary, were chosen to reflect possible costs in other healthcare settings and show the impact of substantial change in the unit costs on overall cost-effectiveness. The table also shows examples of the cost of three screening tests based on the unit costs specified.

Table 22 illustrates how the cost of screening is calculated using the combined test (at an 85% DR). Table 23 shows the cost of screening, including the cost of diagnosis and termination of pregnancy, for 100,000 women according to DR and the screening test used. If these costs are divided by 100,000 one obtains the cost per woman screened. For example, for service A (UK NHS) the costs per woman screened using screening methods each set to achieve an 85% DR would be as shown in Table 24 (all methods including age).

Table 21 Unit costs for the components of an antenatal Down's syndrome screening service and examples of the cost of three tests

	Cost illustrations (£)		
Cost item	Service A (UK NHS)	Service B (4 × A[a])	Service C (6 × A[a])
Reagent cost of each serum marker (singleton assay)	1.50	1.50	1.50
NT measurement	4.50	18	27
Non-reagent laboratory costs (per sample)	3.50	14	21
Service costs (computer-assisted test interpretation and administration)	4.00	16	24
Diagnostic test			
CVS with rapid PCR	350	525	700
Amniocentesis with rapid PCR	300	450	600
Termination of pregnancy	500	750	1000
Medical evacuation of products of conception	400	600	800
Delivery	600	900	1200
Examples[b]			
Combined test (NT plus 2 serum markers)	15	51	75
Quadruple test (4 serum markers)	14	36	51
Integrated test[c] (NT plus 5 serum markers)	23	70	100

[a] Except the reagent costs which are the same as A, and the diagnostic test costs and costs of termination or delivery which are 1.5 and 2 times greater respectively; [b] Cost per test excluding diagnostic test (rounded to the nearest £); [c] Includes non-reagent laboratory costs for two samples.

Table 22 Illustration of how the costs of screening (Tables 23 and 25) are estimated using the combined test at an 85% detection rate (see Table 20). The unit costs for each item are found in Table 21

	Down's syndrome				Unaffected		
	Number of pregnancies (a)	Cost per pregnancy (b) (£)	Total cost (a x b) (£)		Number of pregnancies (c)	Cost per pregnancy (d) (£)	Total cost (c x d) (£)
Combined test:	226	15	3390	(FPR = 6.1%)	99,774	15	1,496,610
(DR = 85%) ↓					↓		
Offered CVS:	192			(80%)	6,086		
(90%) ↓					↓		
Accept CVS with rapid PCR:	173	350	60,550		4,869	350	1,704,150
(90%) ↓				(0.9%)[b]	↓		
	156	500	78,000	Fetal loss	44	400[c]	17,600
Termination of pregnancy (77%) ↓							
Down's syndrome pregnancy that would have been delivered if not terminated	120[a]	600	–72,000[a]				
Total cost (£)			69,940				3,218,360

Total cost of screening 100,000 pregnancies:
£69,940 + £3,218,360 = £3.3 million
Cost per Down's syndrome pregnancy diagnosed:
£3.3 million ÷ 173 = £19,000

[a] If the 156 had not been terminated 77%[11] would have gone to term, so the cost of delivering these pregnancies would have been avoided; [b] See Wald et al., 1997[2]; [c] Cost of removal of products of conception.

Table 23 Cost of screening 100,000 women (including diagnosis and termination of pregnancy) (£ millions) for selected screening tests according to specified detection rate using unit costs specified in Table 21

	DR (%)			
Screening test[a]	75	80	85	90
Service A (UK NHS)				
Double	2.7	3.3	4.3	5.8
Triple	2.3	2.8	3.5	4.8
Quadruple	2.0	2.3	2.9	4.0
NT at 10 completed weeks	4.5	6.0	8.0	11.3
NT at 12–13 completed weeks	3.3	4.6	6.6	10.1
Combined	2.2	2.6	3.3	4.6
Serum integrated	2.1	2.3	2.6	3.2
Integrated	2.4	2.5	2.6	3.0
Service B				
Double	6.1	7.2	8.9	11.7
Triple	5.3	6.2	7.5	9.8
Quadruple	4.7	5.3	6.2	8.1
NT at 10 completed weeks	8.2	10.0	12.6	16.8
NT at 12–13 completed weeks	6.6	8.2	10.8	15.3
Combined	6.0	6.5	7.4	9.1
Serum integrated	5.6	5.8	6.4	7.5
Integrated	7.2	7.3	7.5	8.2
Service C				
Double	8.6	10.0	12.3	15.9
Triple	7.4	8.6	10.3	13.4
Quadruple	6.6	7.4	8.7	11.2
NT at 10 completed weeks	11.4	13.9	17.3	22.9
NT at 12–13 completed weeks	9.4	11.5	14.9	20.9
Combined	8.7	9.4	10.6	12.8
Serum integrated	7.9	8.3	9.0	10.5
Integrated	10.3	10.5	10.9	11.7

[a] All tests include maternal age and, where appropriate, free β-hCG is used in the second trimester. The first trimester markers PAPP-A and free β-hCG are based on the median in Down's syndrome pregnancies at 10 completed weeks (see Table 8) and the NT SD in unaffected pregnancies is applicable to 10 completed weeks. The estimates assume that 90% of women with affected pregnancies and 80% with unaffected pregnancies accept a diagnostic test (the difference is due to affected pregnancies having on average a higher reported risk and such women are more likely to have a diagnostic test).[2] It is also assumed that 90% of pregnancies diagnosed with Down's syndrome are terminated[2]. See appendix, Table 48 for results based on first trimester markers at 12 completed weeks.

Table 24 Costs per woman screened for each screening method for service A

Screening method	Cost (£)
NT measurement	80
Double test	43
Triple test	35
Combined test	33
Quadruple test	29
Integrated test	26[a]
Serum integrated test	26[b]

[a] Precise estimate £26.50; [b] precise estimate £25.60
NB: Costs include diagnosis and termination of pregnancy where applicable. They are not the costs of each test.

Table 25 shows the cost per Down's syndrome pregnancy detected for selected screening tests according to specified DRs using the unit costs in Table 21. For service A (UK NHS) at a DR of 85%, the integrated tests are the least expensive and even with illustrations B and C they are comparable with the other options. The test with the best screening performance (the integrated test) is the one that offers the safest policy (see Table 23) and is also one of the most cost-effective (Table 25). The integrated test comprises more screening measurements and hence more costs, but these are offset by savings achieved because of the lower FPR and the reduced number of amniocenteses. As can be seen from Tables 23 and 25, the rank order of costs varies somewhat according to detection rate.

Unit costs vary from place to place and over time. To enable readers to obtain cost estimates corresponding to those in Tables 23 and 25 based on their own unit

Table 25 Cost of screening (including diagnosis and termination of pregnancy) (£1000s) per Down's syndrome pregnancy diagnosed for selected screening tests according to specified detection rates using unit costs specified in Table 21

	DR (%)			
Screening test[a]	75	80	85	90
Service A (UK NHS)				
Double	17.5	20.3	24.8	31.9
Triple	14.9	16.9	20.3	26.3
Quadruple	13.2	14.4	16.8	21.7
NT at 10 completed weeks	29.8	36.8	46.3	61.5
NT at 12–13 completed weeks	21.9	28.2	38.0	55.0
Combined	14.5	16.0	19.0	25.2
Serum integrated	13.7	13.9	14.8	17.5
Integrated	15.9	15.4	15.3	16.3
Service B				
Double	40.2	44.5	51.7	63.7
Triple	34.8	37.9	43.3	53.4
Quadruple	30.8	32.6	36.1	44.1
NT at 10 completed weeks	53.5	61.6	72.9	91.7
NT at 12–13 completed weeks	43.3	50.4	62.2	83.4
Combined	39.4	40.1	42.8	49.8
Serum integrated	36.6	35.9	36.7	40.7
Integrated	47.0	44.8	43.5	44.6
Service C				
Double	56.2	61.7	71.1	86.9
Triple	48.7	52.6	59.6	73.0
Quadruple	43.4	45.6	50.4	61.2
NT at 10 completed weeks	74.9	85.4	100.2	125.0
NT at 12–13 completed weeks	61.3	70.6	86.0	113.9
Combined	57.1	57.7	61.1	70.1
Serum integrated	51.9	51.1	52.1	57.2
Integrated	67.8	64.6	62.8	63.7

[a] All tests include maternal age and, where appropriate, free β-hCG is used in the second trimester. The first trimester markers PAPP-A and free β-hCG are based on the median in Down's syndrome pregnancies at 10 completed weeks (see Table 8) and the NT SD in unaffected pregnancies is applicable to 10 completed weeks. The estimates assume that 90% of women with affected pregnancies and 80% with unaffected pregnancies accept a diagnostic test (the difference is due to affected pregnancies having on average a higher risk and such women are more likely to have a diagnostic test).[2] It is also assumed that 90% of pregnancies diagnosed with Down's syndrome are terminated[2]. See appendix Table 49 for results based on first trimester markers at 12 completed weeks. To estimate the cost per Down's syndrome birth avoided the above costs would need to be increased by 10%.

costs we have set up an interactive website http://www.smd.qmul.ac.uk/wolfson/screencost in which local unit costs can be entered.

Appendix Tables 48 and 49 show the costs for selected first trimester and integrated tests at each week of pregnancy separately and at 10–13 completed weeks combined.

4 DISCUSSION

Overall results

The most effective screening test was the integrated test, with an estimated 85% DR for a 1.2% FPR if PAPP-A is measured at 10 completed weeks of pregnancy. The performance was much better (with far fewer false-positives for the same detection rate) than with any combination of first trimester screening markers or any combination of second trimester markers. This test also provides the safest and most cost-effective method of screening. Figure 9 compares the FPR for an 85% DR for the main tests. Figure 10 shows in a similar way to Figure 9 the odds of being affected given a positive result (OAPR). The odds were about five times greater with the integrated test than with the first trimester combined test or with the second trimester quadruple test.

While it is to be expected that combining first and second trimester markers would be better than either alone, the effect was greater than might be intuitive. Table 26 illustrates this using a discordant paired analysis. Among women with discordant integrated test and combined test

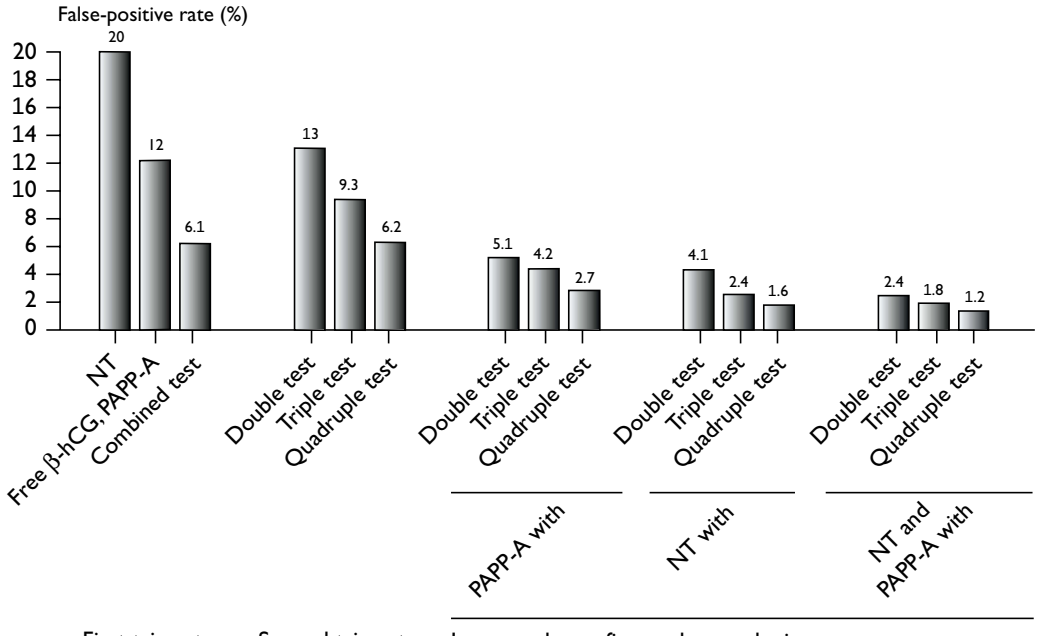

Figure 9 False-positive rate (%) for an 85% detection rate according to different tests in the first trimester (10 completed weeks if PAPP-A used, 12–13 completed weeks if PAPP-A not used), second trimester (14–20 completed weeks), or both trimesters. (Free β-hCG was the hCG measurement used and all tests included maternal age).

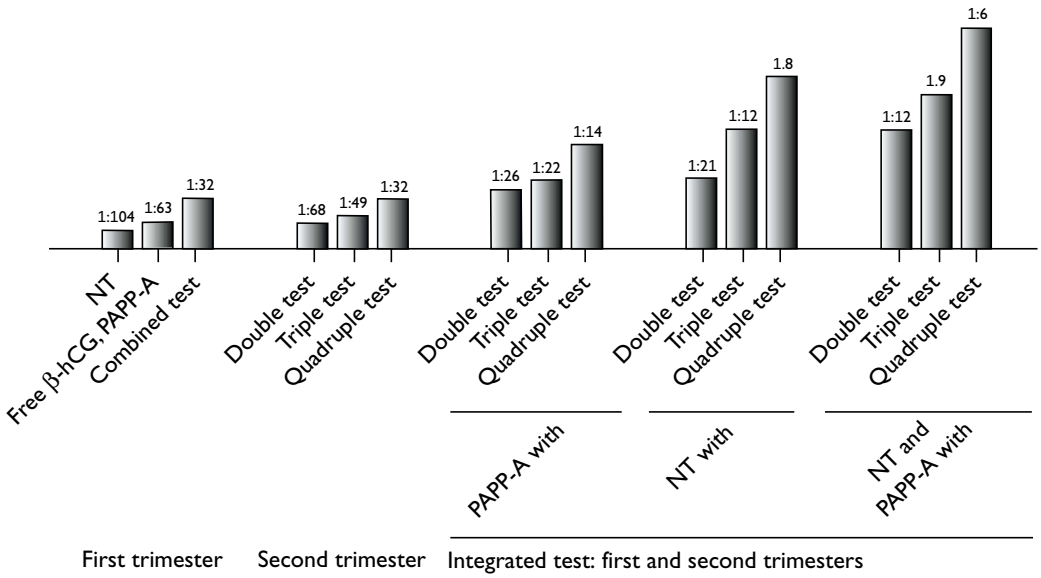

Figure 10 The odds of being affected given a positive result for the tests shown in Figure 9.

Table 26 Percentage of affected and unaffected pregnancies according to combination of test results (combined, quadruple and integrated tests) using a 1 in 250 mid-trimester risk cut-off

Screening test result				
Integrated test	Combined test	Down's syndrome %	Unaffected %	Discriminatory odds ratio[a]
+	+	81.7	1.7	
+	–	8.8	1.2	17.8
–	+	1.4	3.4	
–	–	8.1	93.7	
All		100.0	100.0	
Integrated test	Quadruple test			
+	+	81.8	1.8	
+	–	8.7	1.0	15.7
–	+	2.1	3.8	
–	–	7.4	93.4	
All		100.0	100.0	
Combined test	Quadruple test			
+	+	74.2	1.0	
+	–	8.9	4.1	1.04
–	+	9.8	4.7	
–	–	7.1	90.2	
All		100.0	100.0	

+ Screen-positive; – screen-negative. The first trimester markers NT, PAPP-A and free β-hCG are based on pregnancies at 10 completed weeks (see Table 8). The results are similar at 12 completed weeks. [a] The discriminatory odds ratio expresses the number of times the better test correctly classifies affected and unaffected pregnancies when they yield discrepant results.

results (that is, one positive and the other negative), the integrated test detected about six times more affected pregnancies than it missed (8.8% versus 1.4%) and avoided about three times as many false-positives (3.4% versus 1.2%) in this group – a big improvement. Similarly among women with discordant integrated and quadruple test results, the integrated test detected about four times more affected pregnancies (8.7% versus 2.1%) and avoided about four times as many false-positives (3.8% versus 1.0%). The discriminatory odds ratio in the table quantifies the discriminatory improvement of the better test. The table also shows a direct comparison of the combined test and the quadruple test; the two tests are very similar with each detecting about 10% of the cases missed by the other and each avoiding about half the false-positives from the other. The discriminatory odds ratio is nearly unity (1.04).

The low FPR for an 85% DR achievable with the integrated test (1.2%) can be appreciated by comparing it with the corresponding FPR for the serum integrated test (2.7%), the first trimester combined test (6.1%) and the second trimester quadruple test (6.2%).

Of the urinary markers only ITA was useful but the reduction in the FPR when added to the serum markers may not be considered worthwhile given that a separate sample has to be obtained, processed and analysed (see Table 14).

Differences in results from observational and interventional studies

In intervention studies in which the screening marker being studied is itself used in screening, women with high levels of, for example, free β-hCG are more likely to be screen-positive, and have a diagnostic test and a termination of pregnancy if the diagnosis were positive. Some of these pregnancies might have aborted spontaneously and thus would not have been included if the study had been entirely observational. There is evidence in support of this hypothesis. Table 27 classifies studies on free β-hCG in Down's syndrome pregnancies according to whether they were observational or interventional. The median free β-hCG levels in affected pregnancies in SURUSS were similar to those from other intervention studies but higher than estimates from observational studies. The median MoM values in affected pregnancies were, in general, higher in the interventional studies than the observational ones, a statistically highly significant difference. A similar effect applies to inhibin-A but the median in SURUSS was higher than estimates from other intervention studies (2.54 compared with about 2.2), hence our reason for not using the SURUSS estimate in our statistical analysis. Because of the high correlation between free β-hCG and total hCG it might be expected that SURUSS would also yield a high median total hCG MoM value in affected pregnancies, however, it did not. The median in SURUSS was 2.05 MoM and the summary estimate previously published was 2.06 (95% CI, 1.95 to 2.17).[2] If our explanation is correct, it is possible that the association between hCG and Down's syndrome pregnancies that miscarry is greater with the free β subunit than it is with total hCG.

Method of estimating screening performance

Some investigators estimate screening perform-ance for a given test by estimating the risk of having a Down's syndrome pregnancy for all women in the study and then by specifying the proportion of affected pregnancies at or above the 95th centile of risk in unaffected pregnancies, so yielding the DR for a 5% FPR. There is a problem with this approach. It is greatly dependent upon the estimate of the 95th centile which, even in a sample of several hundred unaffected pregnancies, is subject to significant random variation. This defines the cut-off against which one counts the number of

Table 27 Ranked median second trimester free β-hCG levels in Down's syndrome pregnancies from observational or interventional[a] studies

Study	Number of Down's syndrome cases	Median free β-hCG MoM	Type of study: observational (O) or interventional (I)
Macri *et al.*, 1990[17]	29	1.67	O
Cuckle *et al.*, 1996[18]	30	1.82	O
Aitken *et al.*, 1996[19]	44	1.99	O
Forest *et al.*, 1995[20]	11	2.01	O
Spencer, 1991[21]	29	2.08	O
Knight *et al.*, 1998[22]	52	2.16	O
Wald *et al.*, 1994[23]	75	2.22	O
Spencer and Macri, 1992[24]	23	2.23	O
Crossley *et al.*, 1991[25]	81	2.30	O
Norgaard-Pederson *et al.*, 1994[26]	72	2.31	O
Ryall *et al.*, 1992[27]	57	2.36	O
Macri *et al.*, 1996[28]	10	2.40	I
Wenstrom *et al.*, 1997[29]	31	2.40	O
Spencer *et al.*, 1992[30]	90	2.41	O
Stone *et al.*, 1993[31]	21	2.50	O
Spencer, 1999[32]	107	2.52	I
Bart's screening[b]	95	2.55	I
Macri *et al.*, 1992[33]	20	2.58	O
SURUSS[c]	84	2.66	I
Hsu *et al.*, 1997[34]	47	2.79	I
Chao *et al.*, 1999[35]	15	2.86	I
Birmingham screening[d]	186	2.89	I
Macri *et al.*, 1992[36]	26	3.01	O

Median for observational studies is 2.26 MoM and the median for interventional studies is 2.66 MoM – the difference is statistically significant (p = 0.004). [a] Interventional studies were those which used free β-hCG (or total hCG which is highly correlated with free β-hCG) in the second trimester in their screening programmes (so many affected pregnancies with high hCG will have been terminated including some that would subsequently have ended in spontaneous fetal loss) whereas observational studies, in this context, are those that did not; [b] Unpublished estimate. [c] Of the 85 cases with a second trimester serum sample, a free β-hCG measurement was unavailable for one. [d] Vicky Edwards, Principal Biochemist, Birmingham Women's Hospital: personal communication, 2001.

affected pregnancies exceeding the value, and this itself is subject to substantial random variation given that there are no more than about 100 affected pregnancies in SURUSS. In other words, the estimates are likely to be unstable and for this reason this method has not been used in the past, for example, in the analyses relating to the triple test.[1] The modelled estimates are more stable and are best tested against an independent dataset based on large numbers.

Validity checks

To check the validity of the SURUSS data we applied the SURUSS screening parameters and the algorithm that uses these parameters to estimate the screening performance of the second trimester screening programme at Bart's (about 35,500 pregnancies, including 78 affected pregnancies, a subset of pregnancies with a scan estimate of gestation and maternal weight correction of the marker levels). We found that the predicted estimates were close to those observed. For the quadruple test the predicted DR (at a 5% FPR) was 83% (see Table 12), and the observed rate was 78% (95% CI, 67 to 87%). A sufficiently large first trimester dataset was not available in the Bart's screening programme to test the combined test but we note that our estimate of an 83% DR for a 5% FPR (see Table 11) is similar to that reported by Spencer and his colleagues in a large series (210 pregnancies with Down's syndrome) which observed a rate of 89% without adjustment for fetal loss.[37]

Our results provide an unbiased comparison of the different tests studied because they are based on pregnancies that all reached the second trimester of pregnancy. This is an important advantage of SURUSS. The estimates of screening performance are accurate in the second trimester but DRs will be higher for a specified FPR than estimates at term because the screening markers are associated with both spontaneous fetal loss and Down's syndrome. For example, in the Oxford–Bart's study[38] (which was observational) the quadruple test (with free β-hCG) yielded a 79% DR for a 5% FPR, while the estimate from SURUSS was 83% for 5%. For the triple test the corresponding estimates were 71% and 77% respectively. Our estimates of screening performance using the integrated test (85% detection for a 1.2% FPR at 9–10 completed weeks or 1.3% at 12 completed weeks) were reasonably similar to those previously reported (85% for 0.9%).[3] The similarity in results indicates that the estimates are robust and can be relied on in informing screening policy and service provision.

We did not amalgamate the SURUSS results with those from other studies. To do so would have diminished the value of SURUSS, diluted the conclusions and led to several intractable problems in the statistical analysis of the results. For example, there are differences in the median free β-hCG and inhibin-A levels in affected pregnancies between observational (term pregnancies) and interventional studies (mid-trimester pregnancies); to combine these would yield medians that would not have been applicable to either one

or other period of pregnancy. We have, however, examined all the results from SURUSS to check that where they can be compared with other results in a fair way they are consistent, and in the one instance where they were not (median inhibin-A level in Down's syndrome pregnancies), we used an alternative appropriate estimate.

Fetal loss biases that influence estimates of Down's syndrome screening performance

Two biases tend to overestimate screening performance compared with performance determined by the proportion of term pregnancies with Down's syndrome detected. One bias is general and the other is specific to the screening markers.

General fetal loss bias

Because Down's syndrome pregnancies are more likely to miscarry than unaffected pregnancies, bias will arise in estimations of the DR when Down's syndrome pregnancies that are detected are ascertained at an earlier gestation than Down's syndrome pregnancies that are missed. This can be avoided if all Down's syndrome pregnancies are ascertained at the same time in pregnancy, for example, at term (in the absence of screening) or at the time of amniocentesis (among women who are having an amniocentesis for reasons other than a positive screening result).

Marker-related fetal loss bias

Because screening markers (for example, NT, PAPP-A and inhibin-A) are associated with miscarriage as well as with Down's syndrome pregnancies, bias will arise when the markers are used in "intervention" studies to screen for Down's syndrome with a termination of pregnancy carried out if an affected pregnancy is diagnosed in screen-positive women.

This bias could be avoided if all screening data were available from observational studies in which no intervention was carried out on the basis of the screening markers and all the pregnancies continued to term. However, such data only exist for second trimester markers, not for first trimester markers (for example, NT, PAPP-A and free β-hCG). Published studies on first trimester markers[39] have been subject to this marker-related fetal loss bias, so existing comparisons of first and second trimester screening tests will tend to exaggerate the performance of the first trimester markers relative to the second. There is unfortunately no estimate of the size of the marker-related fetal loss bias. It may lead to DRs about 5 or 10% too high for a 5% FPR if term births were used as the standard.

SURUSS, and a similar USA study (FASTER), can compare the screening performance of first and second trimester tests as well as combine them into a single integrated test, in an unbiased way because intervention was offered after the second trimester screening results were available. While data from these studies can be used to calculate second trimester (about 16–17 weeks) risk estimates, they cannot directly estimate term risks.

While the main advantage of standardising screening performance to about 16–17 weeks is that it enables all the different screening tests for Down's syndrome to be fairly and accurately compared, (which hitherto was not possible), the disadvantage is that the DR will necessarily include some pregnancies that would have miscarried after 17 weeks had a termination not been carried out. The choice between risk estimates and screening performance at term or 16–17 weeks is no longer an option unless corrections and indirect adjustments are made to try to convert one set of estimates into the others. Many centres (for example, in the USA) prefer to give a risk estimate closer to the time of the test, seeing it as the one that is more relevant, so in many places practice will not be changed.

Second trimester versus term risks

While it has been practice to report estimates of screening performance relevant to Down's syndrome pregnancies at term; for the reason given above, this is no longer possible. Screening, antenatal diagnosis, and selective abortion are now commonplace so it is not possible to directly obtain term-estimates. To do so, a study would have to be completely observational and this would be unethical. If there is intervention based on screening results the effect of bias in estimating screening performance at term cannot be avoided. Second trimester screening performance figures will, therefore, of necessity become the standard. Given this, it is probably reasonable to report, in screening programmes, second trimester risk estimates and second trimester estimates of the OAPR.

We recognise that moving risk estimates from term to early mid-trimester is, in some places, a change from previous practice. There is, however, no real choice, given that screening is standard clinical practice and it would not be ethical to perform studies on screening tests that offer no intervention. Since the comparison of different screening tests was standardised to mid-trimester (about 17 weeks), the two fetal loss biases were avoided and the risk estimates were accurate, so differences between tests performed in the first and second trimester can be fairly compared.

Nuchal translucency

A main conclusion from this study is that NT is a good screening marker that can be measured satisfactorily in routine practice, but that it has a poor performance as a screening test for Down's syndrome on its own or with maternal age alone. Its performance with maternal age is a 60% DR for a 5% FPR at 10 completed weeks or a 69% DR at 12–13 completed weeks, and this assumes that a measurement is made in every pregnancy. This performance is inferior to that of tests in which it is used in combination with two or more other markers. Conclusions on NT measurement as a screening marker for Down's syndrome that can be drawn from our results are summarised in Box 3.

Adopting sonographer-specific medians allows for systematic differences between sonographers and ensures that sonographers acquire sufficient initial experience and continuing experience to derive reliable normal median

Box 3 Conclusions on use of NT measurements

- The best time to obtain an NT measurement together with a serum PAPP-A measurement is at 10 completed weeks. The success rate in obtaining NT measurement was greatest at 12 completed weeks; with experience the rate at 10 or 11 weeks was similar to that at 12 weeks.
- NT alone is much less effective than in combination with other screening markers.
- Screening performance is better when NT measurements in millimetres are converted to MoM values using sonographer-specific medians rather than centre-specific medians.
- Taking the mean of several NT measurements obtained during one visit (for example, three) improves screening performance compared with using just one measurement.
- The make and model of the ultrasound machine has an important influence on obtaining a satisfactory NT measurement.

values for NT MoM values to be calculated. It is a simple matter for NT measurements to be linked to a particular sonographer and sonographer-specific medians calculated accordingly. It is also simple to use the mean of three measurements, a policy that we have shown improves screening performance. The independent assessment of NT images while a patient is still present in the ultrasound room is more difficult, but if introduced as routine would yield a useful improvement in screening performance. A possible alternative would be to have a system of regular local independent review of images to improve general screening performance.

Overall, an NT measurement was not obtained in 9% of all pregnancies during the first trimester (5% at 10–12 weeks in the last fifth of SURUSS recruitment). The estimates are partly explained by the fact that the sonographer did not try for more than 20 minutes and perhaps, because it was known that the measurement was for research purposes only, there was less incentive to obtain one. In practice, the percentage is likely to be less than 9%.

Table 28 shows the performance at a risk cut-off of 1 in 250 of the first trimester combined test and the integrated tests in which a varying proportion of pregnancies do not have an NT measurement. The estimates for the combined and integrated tests are similar whether the first trimester markers are measured at 10, 11 or 12 completed weeks. If a measurement is obtained but judged to be unsatisfactory it may still be used though there will be a loss of screening performance (see Table 17).

Absent nasal bone

A proposed ultrasound marker of Down's syndrome was reported after the conclusion of SURUSS – the absence of a nasal bone in the fetus in the first trimester.[40] The marker can be sought in the same ultrasound examination as that used for the NT measurement. The reported results were based on women who were screen-positive using NT measurement and maternal age. In this group the estimated DR was 73% for a 0.5% FPR. If this result is confirmed and found to be applicable to all pregnancies, the addition of the marker could improve the performance of first trimester screening using the combined test to a DR of about 95% for a FPR of 5% and improve the performance of screening using the integrated test to about 95% for a 1% FPR.

De Biasio and colleagues[41] have provided some evidence that would suggest that seeking an absent nasal bone may not be as straightforward as had been proposed. In five affected pregnancies they reported seeing a nasal bone, suggesting that DRs may not be as high as previously reported.

Gestational timing of measurements

Our results show the importance of using week-specific median PAPP-A and free β-hCG levels together with week-specific SDs for NT measurement. The timing of the first trimester measurement must represent a compromise. The value of PAPP-A diminishes with gestation while the value of free β-hCG level increases, as does the NT measurement but only to a small extent. For the combined test it matters little in which week between 10 and 13 the test is performed; there is a small advantage in carrying out the first trimester measurements for the integrated test at 10 completed weeks, and there is a substantial advantage for the serum integrated test (see Table 15). Given that sometimes an NT measurement may not be obtained, it might be sensible to aim for 10 completed weeks when arranging the first trimester measurements for the integrated test so that if a serum integrated test result is issued it will be more useful than one done later.

For these reasons in the results we have used 10 completed weeks as the timing for the first trimester measurements. We also show results relating to screening at 12 completed weeks when an NT measurement is most easily obtained. Appendix Tables 44–49 provide further details (including details on screening performance, safety and cost of screening) from when the first trimester measurements are taken at different gestational weeks.

The quadruple test, whether part of the integrated test or as a separate test, can be done from 14 completed weeks of gestation. At centres where an ultrasound anomaly scan is routine, 14–15 completed weeks is an appropriate gestation for blood collection, but for other centres this should be at about 16 completed weeks, because of AFP screening for open neural tube defects. The main conclusions on the gestational timing of measurements are given in Box 4.

Table 28 Screening performance for tests in which an NT measurement is not available for specified percentages of pregnancies. All tests are used at an early mid-trimester risk cut-off level of 1 in 250 and include maternal age

	NT measurement available on all women		NT measurement not available for									
			5%		10%		15%		20%		25%	
	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)	DR (%)	FPR (%)
NT												
At 10 completed weeks	69	8.6	68	8.8	67	8.9	66	9.1	65	9.2	64	9.4
At 12–13 completed weeks	71	6.0	70	6.3	69	6.6	67	6.9	66	7.1	65	7.4
Combined test[a]												
NT, Free β-hCG, PAPP-A	83	5.0	83	5.1	83	5.2	82	5.4	82	5.5	82	5.6
Integrated test[a,b]												
NT and PAPP-A with:												
Double test	88	3.6	88	3.7	88	3.8	88	3.8	87	3.9	87	4.0
Triple test	89	3.3	89	3.4	89	3.4	89	3.5	88	3.6	88	3.7
Quadruple test	90	2.8	90	2.9	90	2.9	90	3.0	89	3.0	90	3.1

[a] The first trimester markers PAPP-A and free β-hCG are based on the median in Down's syndrome pregnancies at 10 completed weeks (see Table 8) and the NT SD in unaffected pregnancies is applicable to 10 completed weeks. The results are similar when the first trimester measurements are taken at 12 completed weeks;
[b] Includes free β-hCG in the second trimester.

Box 4 Conclusions on gestational timing of measurement of screening markers

- All biochemical markers are best in the second trimester except PAPP-A.
- NT performance improves between 10 and 12–13 completed weeks.
- PAPP-A performance deteriorates between 10 and 12–13 completed weeks.
- Trade-off between NT and PAPP-A favours 10 completed weeks when both are to be used.
- The best screening performance is achieved by integrating results of markers measured in the first and second trimesters.

Choice of markers

The selection of the panel of screening markers to be used, when in pregnancy they should be measured, and the most effective way of combining them should be based on the efficacy and safety of the screening policy. The final decision would then be influenced by cost. We adopted the following guidelines:

- The primary aim is to maximise screening performance, that is, to maximise the DR for a given FPR, which is equivalent to minimising the FPR for a given DR. This provides the best combination of accuracy and safety, which is the chief concern of women being screened and their medical advisers.
- In general, the effect of adding further markers should be assessed by including them in order of screening performance (best ones first) unless there is a proven additional benefit from the use of the marker (for example, AFP in neural tube defect screening).
- The markers should be interpreted together as a single test, as this maximises screening performance, yielding the lowest FPR for a given DR.
- If markers from the first and second trimesters are combined (that is, in an integrated test), each marker should be used in the trimester in which it best discriminates between affected and unaffected pregnancies.

The main conclusions that follow from these four guidelines are summarised in Box 5.

Efficacy, safety, and cost-effectiveness

A summary of the efficacy, safety, and costs linked to the different screening tests of choice is given in Table 29. The screening tests are shown together with how they would be used: the integrated test offered routinely, the serum integrated test offered if an NT measurement is not available, and the quadruple test if the patient's first visit is in the second trimester.

Box 5 Main conclusions that follow from the four guidelines

- The integrated test has the best screening performance and this is most appropriately based on NT and PAPP-A measurement at 10 completed weeks and the quadruple test from 14–22 completed weeks.
- With the integrated test (and the combined test), 10 completed weeks is the preferred single time in pregnancy for the NT and PAPP-A measurement, because this maximises screening performance using these two markers together.
- If NT measurements are not available the serum integrated test has the best screening performance.
- It is best to perform the hCG measurement (free β-hCG or total hCG) in the second trimester rather than the first because it is more discriminatory in the second trimester and because it preserves the benefits of the quadruple test for women who book too late for the first trimester component of the integrated test.

The fetal loss rate among women who had an amniocentesis was 1.8% (based on 20 intrauterine deaths (≥24 weeks) and 19 miscarriages (<24 weeks) but excluding 10 stillbirths). Among women who did not have an amniocentesis the rate was 0.79%, (based on 76 intrauterine deaths and 262 miscarriages but excluding 111 stillbirths). The excess fetal loss rate was therefore 1.0% (95% CI, 0.6 to 1.4%). Some of this difference reflects the association between screening markers and miscarriages, but the upper bound of the 95% CI, 1.4%, can be taken as representing a secure upper limit to the hazard from the procedure.

The diagnostic test and termination of pregnancy

Any screening test needs to be judged in the context of the whole screening programme. In this respect the methods of antenatal diagnosis and the timing of termination of pregnancy are relevant. The antenatal diagnosis of Down's syndrome and other chromosome disorders is usually carried out by performing a karyotype on cultures of amniotic fluid or chorionic villus cells. This usually takes 7–14 days to complete. Fluorescence *in situ* hybridisation[42] or quantitative PCR (quantitative fluorescence PCR, or rapid PCR)[43,44] can provide a result within 48 hours and has been found to be reliable for the diagnosis of Down's syndrome, Edward's syndrome (trisomy 18), and Patau's syndrome (trisomy 13) as well as numerical sex chromosome disorders. The PCR test is rapid, inexpensive, and large numbers of samples can be tested with relatively few staff.

The period between undergoing an amniocentesis or CVS and receiving the result is one of great anxiety for the couple involved. In view of this, there would be advantages if the rapid PCR test were introduced as a standard diagnostic test. A conventional cytogenetic analysis could follow; the continuing requirement for a full karyotype would need to

Table 29 Screening tests of choice: efficacy, safety and cost

		Efficacy		Safety	Cost
Setting	Test	FPR (%) for an 85% DR	OAPR	Number of procedure-related fetal losses per 100,000	£ per woman screened in NHS to achieve an 85% DR[a]
Offered routinely					
In general	Integrated test	1.2	1:6	9	26
If NT test not available	Serum integrated test	2.7	1:14	19	26
If first visit is in second trimester	Quadruple test	6.2	1:32	45	29
Available but not offered routinely					
For women who request a first trimester test	Combined test	6.1	1:32	44	33

[a] Costs include diagnosis and termination of pregnancy where applicable. They are not the costs of each test.

be based on detailed comparison of the two methods in respect of the specific disorders for which a diagnosis was required.

Termination of pregnancy should, in the majority of cases, be possible at about 17 weeks, before fetal movements have been felt. The integrated test would result in fewer terminations than first trimester screening because of the spontaneous miscarriages of affected pregnancies that occur between 12 and 15 completed weeks, so fewer women would be compelled to make a decision on whether to terminate the pregnancy.

Selecting a screening policy

Table 30 shows the expected DRs and FPRs for the main screening tests at different risk cut-off levels. The table indicates the trade-off between increasing the DR and increasing the FPR for given changes in the cut-off. For example, if the integrated test were used, a risk cut-off of 1 in 200 would yield an 89% DR for a 2.4% FPR, and little would be gained in lowering the risk cut-off.

Table 31 shows the summary results for the main screening tests using an early mid-trimester risk cut-off of 1 in 250.

Meeting the criteria for a worthwhile screening test

The proposed screening methods improve on the performance of existing antenatal screening tests for Down's syndrome which themselves were judged to have met the criteria for a worthwhile screening test including those of the National Screening Committee. We have shown that the proposed methods of screening are, depending on the DR, probably no more expensive per case detected than existing methods, and may be less expensive. Demonstration projects have shown the acceptability and feasibility of the first trimester combined test,[37] the quadruple test,[49] and the integrated test (Wald N, unpublished observation from the Integrated Risk Screening (IRS) project).

Introducing the integrated test

The integrated test has been introduced at several centres in North America and Europe. It requires a new routine and cooperation between different medical departments. The main challenge relates to providing a service that can measure NT accurately and precisely. This requires adequate equipment, staff training, time for the measurement, quality assurance, and appropriate handling of the data in the calculation of risk. Experience has been gained in offering the integrated test as part of an international collaborative demonstration project, the Integrated Risk Screening Project (coordinated by Karen and Nicholas Wald, Barts and the London School of Medicine). The project involves screening women in both the public and private sectors. By August 2002 about 24,000 women had been screened. As well as demonstrating the feasibility and acceptability of screening using the integrated test (both with and without NT measurement), the project team produced educational material including information leaflets for patients and health professionals. The patient information leaflet used in

Table 30 Screening performance according to early second trimester risk cut-off for specified screening tests

	Risk cut-off																	
	1 in 50			1 in 100			1 in 150			1 in 200			1 in 250			1 in 300		
Screening test (all include maternal age)	DR (%)	FPR (%)	OAPR	DR (%)	FPR (%)	OAPR	DR (%)	FPR (%)	OAPR	DR (%)	FPR (%)	OAPR	DR (%)	FPR (%)	OAPR	DR (%)	FPR (%)	OAPR
Combined																		
10 completed wks	66	1.0	1:7	74	2.1	1:13	78	3.1	1:18	81	4.1	1:22	83	5.0	1:27	85	6.0	1:31
12 completed wks	68	0.9	1:6	75	1.9	1:11	78	2.8	1:16	81	3.8	1:21	83	4.7	1:25	84	5.5	1:29
Triple	60	1.4	1:10	70	2.9	1:18	75	4.3	1:26	79	5.7	1:32	81	6.9	1:38	83	8.0	1:43
Quadruple	67	1.3	1:8	74	2.5	1:15	79	3.6	1:20	82	4.7	1:25	84	5.7	1:30	86	6.6	1:34
Serum integrated																		
10 completed wks[a]	76	1.0	1:6	82	1.8	1:10	85	2.6	1:14	87	3.3	1:17	88	4.0	1:20	89	4.6	1:23
12 completed wks[a]	70	1.1	1:7	77	2.2	1:12	81	3.2	1:17	83	4.1	1:22	85	5.0	1:26	86	5.8	1:30
Integrated																		
10 completed wks[a]	81	0.7	1:4	86	1.3	1:7	88	1.9	1: 9	89	2.4	1:12	90	2.8	1:14	91	3.3	1:16
12 completed wks[a]	81	0.7	1:4	85	1.3	1:7	87	1.9	1:10	89	2.4	1:12	90	3.0	1:14	91	3.4	1:17

Free β-hCG is used in the triple, quadruple or integrated tests. [a] First trimester markers measured at this time in pregnancy.

Table 31 Screening tests of choice: efficacy, safety and cost at a constant 1 in 250 early mid-trimester risk cut-off level

	Efficacy			Safety	Cost[a]
Screening test	DR (%)	FPR (%)	OAPR	Number of procedure-related fetal losses per 100,000 pregnancies	£ per woman screened in NHS
Integrated[b]	90	2.8	1:14	20	30
Serum integrated[b]	88	3.4	1:20	29	29
Quadruple	84	5.7	1:30	41	28
Combined[b]	83	5.0	1:27	36	30

[a] Costs include diagnosis and termination of pregnancy where applicable. They are not the costs of each test; [b] First trimester markers measured at 10 completed weeks.

the London centre can be assessed on the Internet at http://www.smd.qmul.ac.uk/wolfson/epm/screening/scrinteg/integ.pdf

If satisfactory NT measurements are not available, a serum integrated result can be issued, possibly using the booking blood as the first trimester sample. Even if NT measurements were to be used in the future, the serum integrated test would be a useful first step. One programme, in Maine USA (Foundation for Blood Research) has now introduced it routinely.

Screening need not be based either on the serum integrated test or on the integrated test; it can be a mixture of both. Some pregnancies may have an NT measurement while others do not. NT measurements could be introduced gradually in the same way that the use of an ultrasound estimate of gestational age was introduced gradually with the triple test, used where available, but not otherwise. This could form the basis of regional screening policies first introducing the serum integrated test and moving to the full integrated test (including NT measurement when the appropriate ultrasound facilities and expertise are available).

The value of including PAPP-A as part of the integrated test diminishes with increasing gestational age. At 10 completed weeks the FPRs with and without PAPP-A at an 85% DR were 1.2% and 2.5% respectively (see Table 13), but at 12 completed weeks the estimates were 1.3% and 1.6% respectively (see appendix, Table 45). Nonetheless, it is worth preserving PAPP-A as a marker, provided that the cost of the PAPP-A measurement is not too great, since if an NT measurement cannot be obtained, PAPP-A will compensate for much of the lost screening performance. So, if a woman were first seen at 10 completed weeks, there is a clear advantage in including PAPP-A (see Table 15).

Three disadvantages might be thought to arise with the integrated test compared with first trimester screening: a possible greater risk of later terminations, the possible distress associated with a later termination, and the assumed loss of earlier reassurance. None of these, in the view of the authors, is compelling. Firstly, while there is an increased risk of a maternal death from a termination at about 17 weeks of pregnancy compared with one at 13 weeks, the difference is extremely small. According to Lawson and co-workers,[45] using USA data from 1972–87, at 13 weeks of pregnancy the maternal mortality rate associated with a termination of pregnancy was 2 per 100,000 pregnancies and at 17 weeks of pregnancy it was 7 per 100,000. In 10 million screened women, with a DR of 85% and a 90% uptake of amniocentesis among women with affected pregnancies, about 17300 women will have a termination of pregnancy on account of a diagnosis of Down's syndrome if all women with a detected affected pregnancy have a termination (see Table 20), so there would be 0.35 expected maternal deaths at 13 weeks or 1.21 at 17 weeks. The excess risk is therefore less than 1 per 10 million women screened (0.86 per 10 million), which in Britain would mean one additional maternal death every 20 years. This is likely to be an overestimate since mortality rates from termination of pregnancy have been declining over time, as would be expected from improvements in techniques for termination. The risk will therefore now be significantly lower than that reported by Lawson and co-workers[45] and needs to be judged in relation to the large number of extra invasive diagnostic procedures and consequent additional losses of unaffected fetuses that would arise from using first trimester tests instead of the integrated test (about 4000 per 10 million women screened). Secondly, while there may be less psychological distress with an earlier termination for women with unwanted pregnancies there is no reason to believe that this would also apply to women with a wanted pregnancy in which a serious fetal abnormality is diagnosed. A recent study of women with fetal abnormalities showed that the time of termination had little or no effect on the amount of distress.[46] It is the decision whether to have a termination that appears to be more stressful than the procedure itself or its timing. Also, there would be some earlier terminations of pregnancy that would otherwise have miscarried with the possibility of women feeling remorse for a termination that need not have taken place. Thirdly, the notion that women could be given early reassurance is unjustified; even a low risk estimate in a first trimester test could become a high one if the second trimester markers were used as well, and vice versa.

Table 32 shows that a policy of offering the combined test to all women, followed by the quadruple test in screen-negative women is not supported by the SURUSS results. Such a policy would lead to a DR of 93% with a FPR of 9.8%. Table 32 also shows the equivalent FPR (4.5%) at the same DR using the integrated test given to all women. With the integrated approach the FPR is more than halved with consequent significant improvement in the safety (48 fewer unaffected fetal losses) and substantial savings in costs.

Future research

There are several areas of research that would be worth pursuing. Ongoing demonstration projects of the integrated test need to be reported so that the feasibility and acceptability of this approach is documented. These are underway.

Table 32 Screening using a stepwise or integrated approach

			In 100,000 women screened:		
Policy based on:	DR[a] (%)	FPR[a] (%)	Number of false-positives	Number of unaffected fetal losses	Cost of screening (£)
Stepwise screening					
Women have first trimester combined test and screen-negatives have second trimester quadruple test (early second trimester risk cut-off 1 in 250 for both tests)	93	9.8	9728	88	6.1 million
Integrated screening					
All women have first trimester NT, PAPP-A and second trimester quadruple test (risk cut-off set to achieve same DR as stepwise policy)	93	4.5	4490	40	3.7 million
Advantage of integrated test policy		54% lower	5328 fewer false-positives	48 fewer unaffected fetal losses	2.4 million less expensive

[a] Assumes 100% uptake of CVS or amniocentesis after each test result.

The value of absent nasal bone as a marker needs to be assessed in different settings and estimates made of its screening performance alone and in combination with other markers. This would need to be linked to research into the elements needed for a satisfactory examination and the time taken to complete this. Although the value of ITA in urine did not add materially to screening performance there may be advantages in measuring this protein in maternal serum. Research in this area would be worthwhile. Research into the use of fetal cell and free fetal DNA in maternal blood as a method of screening is active.[47–49] At present the performance of this approach is poor and there are technical problems but the position may change in the future.

5 CONCLUSIONS

SURUSS has provided, for the first time, a single large dataset in which the ultrasound marker NT was measured together with a wide range of biochemical serum and urine markers in the first and second trimesters of pregnancy. The strength of the study is that it reflected screening performance within the context of the provision of routine antenatal care in 25 centres and that all the data were based on a single cohort of screened women. The data from the study permit the examination of any combination of the specified screening markers within the first trimester, the second trimester, and across both trimesters.

The study helps clarify four issues: (a) the ability to obtain adequate NT measurements in routine practice, (b) the correlation between markers measured in the first and second trimesters of pregnancy, (c) the effect of selective miscarriage early in pregnancy, and (d) the perception that because earlier published estimates of screening performance of the integrated test were derived from different tests in different women they might be incorrect.

Our results showed that overall, on the basis of efficacy, safety, and cost, the integrated test is the test of choice. Adding other markers provided little benefit. The integrated test yielded an 85% detection rate for a false-positive rate of 1.2% if a satisfactory NT measurement was obtained for all or nearly all pregnancies and PAPP-A was measured at 10 completed weeks. If an NT measurement was not available, the serum integrated test (using the same serum markers) would be the next best screening method (85% detection rate for a 2.7% false-positive rate), materially better than any first or second trimester serum screening test.

The benefit of integrating markers across the two trimesters is greater than might intuitively be expected; it decreases the false-positive rate substantially, compared with screening in either trimester alone. It therefore has a large impact in reducing the number of women requiring an invasive diagnostic procedure and hence reducing the loss of unaffected pregnancies.

For women who present for the first time in the second trimester of pregnancy, the SURUSS results suggest that the quadruple test is the test of choice, confirming the results from other studies.[2,50] For women who request a screening result and a diagnosis made before 14 completed weeks of pregnancy, the combined test was found to be the best option, though women would need to be informed that the efficacy and safety of this screening and diagnostic regimen is inferior to the use of the integrated test.

The SURUSS results show that in antenatal screening for Down's syndrome it is now possible to obtain a high level of detection (detecting 8 or 9 out of every 10 affected pregnancies) with a false-positive rate (1–2%) that is substantially lower than in the past, so achieving a significantly higher level of safety by reducing the number of women who need an invasive diagnostic test such as amniocentesis.

ACKNOWLEDGEMENTS

This report was commissioned by the NHS R&D HTA Programme, project number 93/25/05.

We wish to thank Lynne George, Alison Smith, Wayne Huttly, Cecily Cromby, Sandeep Patel, Tiffany Rowbotham, Claire Hearld, Lu Gau and Tianhua Huang for assisting with the administration of the study; Tiesheng Wu for computing assistance; Janette Mackie for editorial assistance; Maggie Fitchett and the Association of Clinical Cytogeneticists, and Eva Alberman and David Mutton of the Down's syndrome register for their assistance. We thank Jack Canick, James Haddow, Simon Rish, Malcolm Law, Glenn Palomaki and Alicja Rudnicka for comments on various drafts of the manuscripts. We thank Perkin Elmer Life Sciences (Cambridge, UK) for the gift of the reagents for the AFP, uE_3, PAPP-A, free β-hCG and total hCG assays, Oxford Bioinnovation and Diagnostic Systems Laboratories Incorporated (Oxford, UK) for the gift of the reagents for the inhibin-A assay, and Quest Diagnostics (New Jersey, USA) for performing the urine assays. We are also indebted to the referees for their perseverance in reading the report and the quality of their comments.

We acknowledge the assistance of the following at the participating centres:

- Aberdeen Maternity Hospital: Kate Duncan, radiographer; Kay Jamieson, midwife; Pam Hewson, radiographer; Julie Smith, midwife; Chris Clarke, radiographer; Grace Porter, midwife; Susan Lethbridge, midwife; Kay Chubb, radiographer
- Barnsley District Hospital: A P Huggard, superintendent ultrasonographer; L Bird, sonographer; P Sykes, sonographer; C Mulligan, sonographer; G Bateman, midwife; J Day, midwife
- Billinge Hospital: S Atkinson, sonographer; D Bushell, sonographer; V Purnell, sonographer; C Henshall, sonographer; J Jones, sonographer; A Mundy, sonographer; C Houghton, sonographer; J Phoenix, coordinator and sonographer
- Birmingham Women's Hospital: Lynn North, coordinator and radiographer; Marguerite Usher-Smyers, radiographer; Lynn Greenfield, radiographer; Lida Debono, radiographer; Nicola Somerset, radiographer; Steve Hussey, radiographer; Jane Stus, radiographer; Liam Allen, radiographer; Ellen Poole, radiographer
- Birmingham Heartlands Hospital: P Wyldes, consultant obstetrician; P Jordan, midwife; A Harborne, midwife; J Agnew, midwife
- Derby City Hospital: Colleen Hardwick, sonographer; Pauline Noonan, sonographer
- Kettering General Hospital: Lynne King, sonographer; Susan Howes, data collection; Pat Ward, coordinator; Margaret Greer, coordinator
- Llandough Hospital: Helena Clarke, radiographer; Pam Henderson, radiographer; Allison Rogers, radiographer; Molly Garbett, radiographer; Emma Regan, radiographer
- Leicester General Hospital: Janet Brookes, sonographer; Julie Broughton, sonographer; Claire Roberts, sonographer; Debbie Conway, data collection ; Chris Gunn, antenatal manager
- Liverpool Women's Hospital: Mary McDonald, coordinator
- Nevill Hall Hospital: Myrna McGrath, coordinator; Karen Bertie, coordinator; Nicki Carey, coordinator; Kay James,

sonographer; Ceri Burrows, sonographer; Ruth Fouracre, sonographer; Cheryl Palmer, sonographer; Susan Walters, radiographer; John Tovey, biochemist; Angela Esfahani, sonographer

- Alexandra Hospital, Redditch : Kim Dalley, coordinator; Chris Upcott, sonographer
- The Royal Gwent Hospital: A Coburn, coordinator; A Estfahani, sonographer; R Jenkins, sonographer; H Hughes, biochemistry
- The General Hospital, Jersey: Susan Poingoestre, coordinator and sonographer; Laura Westwater, sonographer; Pat Smith, sonographer; Deirdre Hill, coordinator; Tim Sims, serum and urine biochemist; Emily Olive, coordinator
- Donauspital am SMZO, Vienna: Marion Rankine, sonographer; Astrid Landgraf, sonographer; Eva Liebhart, sonographer; Gunter Kahles, sonographer; Inge Anderl, sonographer
- Whittington Hospital: Julie Logue, coordinator; Jigna Patel, radiographer; Marlene Minhas, radiographer; Irene Vacheras, radiographer; Mari Semanchelan, radiographer; Elizabeth Hoo, radiographer; Raj Dave, radiographer
- Stepping Hill Hospital: Julie Holowko, coordinator; Malcolm Nye, sonographer.

The views expressed in this report are those of the authors and not necessarily those of the NHS R&D HTA Programme. Any errors are the responsibility of the authors.

AUTHORS' AFFILIATIONS

N J Wald, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, London, UK
C Rodeck, Department of Obstetrics and Gynaecology, University College and Middlesex School of Medicine, University College, London, UK
A K Hackshaw, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, London, UK
J Walters, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, London, UK
L Chitty, Department of Obstetrics and Gynaecology, University College and Middlesex School of Medicine, University College, London, UK
A M Mackinson, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, London, UK

Declared competing interests of the authors: St Bartholomew's and the Royal London School of Medicine and Dentistry together with the Foundation for Blood Research and the Women and Infants' Hospital (Providence, Rhode Island, USA) hold a patent for unconjugated oestriol measurement as a marker for Down's syndrome screening. Professor Wald is a director of Logical Medical Systems Ltd which produces αlpha™, a commercial interpretive software package for Down's syndrome screening using ultrasound and serum markers. He is also a director of Intema Ltd which holds a patent application for the integrated test.

This report is also published as: Wald NJ, Rodeck C, Hackshaw AK, Chitty L, Mackinson AM. First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and ultrasound Screening Study (SURUSS). *Health Technol Assess* 2003;**7**(11).

LIST OF INVESTIGATORS (THE SURUSS RESEARCH GROUP)

Professor N J Wald, Mr A K Hackshaw, Ms J Walters, Ms A M Mackinson, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, UK.
Professor C Rodeck, Dr L Chitty, University College and Middlesex School of Medicine, London, UK.
Ms Heulwen Morgan, Ms Gaye Henson, Whittington Hospital, London, UK.
Mr Neil MacLachlan, The General Hospital, St Helier, Jersey, UK.
Mr Bryan Beattie, Mrs Pam Manfield, University Hospital of Wales, Cardiff, UK.
Mrs Jo Weiner, Royal Gwent Hospital, Newport, UK.
Dr Pat Smith, Ms Nicola Anderson, Ms Rosey Foubister, Aberdeen Maternity Hospital, Aberdeen, and Peterhead Hospital, Peterhead, UK.
Mr Bill Mackenzie, Ms Gina Dutton, Birmingham Heartland's Hospital, Birmingham, UK.
Mr Charles Stewart, Ms Debbie Conway, Ms Chris Gunn, Leicester General Hospital, Leicester, UK.
Dr Claire Candelier, Ms Julie Holowko, Stepping Hill Hospital, Stockport, UK.
Ms Alison Fowlie, Ms Coleen Hardwick, Ms Pauline Noonan, Derby City Hospital, Derby, UK.
Ms Alexandra Rees, Ms Brenda Richards, Llandough Hospital, Penarth, UK.
Mr Paul Wood, Ms Pat Ward, Kettering General Hospital, Kettering, UK.
Mr S J Whittaker, Ms Sue Wade, Ms Gillian Bateman, Barnsley District General Hospital, Barnsley, UK.
Mrs P A Lewis, Billinge Hospital, Wigan, UK.
Dr Simon Jenkinson, Ms Kim Dalley, Alexandra Hospital, Redditch, UK.
Mrs Margaret Morgan, Mr Rob Llewelyn, Ms Sharon Butcher, Singleton Hospital, Swansea, UK.
Mr Andrew Dawson, Nevill Hall Hospital, Abergavenny, UK.
Dr Katrina Schuchter, Mr Martin Metzenbauer, Donauspital, Vienna, Austria.
Dr Boon Lim, Ms Sarah Johnson, Dr Gray's Hospital, Elgin, UK.
Mr Rob Bradley, Royal Sussex County Hospital, Brighton, UK.
Professor Martin Whittle, Ms Lynn North, Birmingham Women's Hospital, Birmingham, UK.
Mr Steve Walkinshaw, Ms Nell Shaw, Liverpool Women's Hospital, Liverpool, UK.
Mr Colin Partington, Ms Mollie Smith, St John's Hospital, Chelmsford, UK.
Mr David Griffin, Watford General Hospital, Watford, UK.
Professor A McGuire, London School of Economics, UK.

GLOSSARY

Affected pregnancy A pregnancy with a fetus affected with Down's syndrome.

Combined test First trimester test based on combining nuchal translucency measurement with free β-human chorionic gonadotrophin (β-hCG), pregnancy-associated plasma protein A (PAPP-A) and maternal age.

Cut-off level The level chosen to define a positive result and distinguish it from a negative result. With a single marker this will be a specified level of the marker. With tests based on a combination of markers it will be a risk estimate.

Double test Second trimester test based on the measurement of α-fetoprotein (AFP), hCG, (either free β-hCG or total hCG), together with maternal age.

False-positive Unaffected pregnancy with a positive test result.

Fetal loss A pregnancy that has miscarried or resulted in an intra-uterine death or a stillbirth.

Integrated test The integration of measurements performed during the first and second trimester of pregnancy into a single test result. Unless otherwise qualified, "integrated test" refers to the integration of nuchal translucency and PAPP-A measurements in the first trimester with the quadruple test markers in the second.

Nuchal translucency (NT) measurement The width of an area of translucency at the back of the fetal neck, usually measured at about 10–13 weeks of pregnancy using ultrasound.

Quadruple test Second trimester test based on the measurement of AFP, unconjugated oestriol (uE_3), free β-hCG (or total hCG), and inhibin-A together with maternal age.

Serum integrated test A variant of the integrated test using serum markers only (PAPP-A in the first trimester and the quadruple test markers in the second trimester).

Triple test Second trimester test based on the measurement of AFP, uE_3, and hCG (either total hCG or free β-hCG) together with maternal age.

True-positive Affected pregnancy with a positive test result.

Unaffected pregnancy A pregnancy with a fetus without a chromosomal defect.

ABBREVIATIONS

AFP	α-fetoprotein
BPD	biparietal diameter
CI	confidence interval
CRL	crown–rump length
CVS	chorionic villus sampling
DR	detection rate: the proportion of women with Down's syndrome (affected) pregnancies who have positive results
FASTER	a study entitled "First And Second Trimester Evaluation of Risk"
FPR	false-positive rate: the proportion of women with unaffected pregnancies who have positive results
hCG	human chorionic gonadotrophin
IRS	Integrated Risk Screening (project)
ITA	invasive trophoblast antigen (also called hyperglycosylated hCG)
IU	international unit
LMP	last menstrual period
MoM	multiple of median among unaffected pregnancies
NT	nuchal transclucency
OAPR	odds of being affected given a positive result: true-positives to false-positives
SD	standard deviation
PAPP-A	pregnancy-associated plasma protein A
PCR	polymerase chain reaction
SURUSS	Serum Urine and Ultrasound Screening Study
uE_3	unconjugated oestriol

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.

REFERENCES

1 Wald NJ, Cuckle HS, Densem JW, *et al.* Maternal serum screening for Down's syndrome in early pregnancy. *BMJ* 1988;**297**:883–7.
2 Wald NJ, Kennard A, Hackshaw A, McGuire A. Antenatal screening for Down's syndrome. *J Med Screen* 1997;**4**:181–247.
3 Wald NJ, Watt HC, Hackshaw AK. Integrated screening for Down's syndrome based on tests performed during the first and second trimesters. *N Engl J Med* 1999;**341**:461–7.
4 Cuckle HS, Canick JA, Kellner LH on behalf of the Collaborative Study Group. Collaborative study of maternal urine β-core hCG screening for Down syndrome. *Prenat Diagn* 1999;**19**:911–17.
5 Cole LA, Shahabi S, Oz UA, *et al.* Hyperglycosylated hCG (invasive trophoblasts antigen) immunoassay: a new basis for gestational Down syndrome screening. *Clin Chem* 1999;**45**:2109–19.
6 Wald NJ, Cuckle HS, Densem JW, *et al.* Maternal serum screening for Down's syndrome: the effect of routine ultrasound scan determination of gestational age and adjustment for maternal weight. *Brit J Obstet Gynaecol* 1992;**99**:144–9.
7 Morris JK, Mutton DE, Alberman E. Revised estimates of maternal age specific live birth prevalence of Down's syndrome in the absence of antenatal screening and selective termination. *J Med Screen* 2002;**9**:2–6.
8 Wald N, Hackshaw A. Tests using multiple markers. In: Wald NJ, Leck I, editors. Antenatal and neonatal screening, 2nd ed., Oxford: Oxford University Press; 2000. p. 23–57.
9 Office for National Statistics. Birth Statistics. England and Wales; 1996, 1997, 1998 (FM1, Nos 25–27). London: The Stationery Office; 1998.
10 Royston P, Thompson SG. Model-based screening by risk with application to Down's syndrome. *Stat Med* 1992;**11**:257–68.
11 Morris JK, Wald NJ, Watt HC. Fetal loss in Down syndrome pregnancies. *Prenat Diagn* 1999;**19**:142–5.
12 Gosden C, Tabor A, Leck IL, *et al.* Amniocentesis and chorionic villus sampling. In: Wald NJ, Leck I, editors. *Antenatal and neonatal screening*, 2nd ed., Oxford: Oxford University Press; 2000, p 470–516.
13 Gilbert RE, Augood C, Gupta R, *et al.* Screening for Down's syndrome: effects, safety, and cost effectiveness of first and second trimester strategies. *BMJ* 2001;**323**:423–5.
14 Nicolaides KH, Snijders KM, Cuckle HS. Correct estimation of parameters for ultrasound translucency screening. *Prenat Diagn* 1998;**18**:519–21.
15 Hyett JA, Sebire NJ, Snijders, Nicolaides KH. Intrauterine lethality of trisomy 21 fetuses with increased nuchal translucency thickness. *Ultrasound Obstet Gynecol* 1996;**7**:101–3.
16 Grudzinskas JG. Endocrinological and metabolic assessment of early pregnancy. In: Chamberlain G, editor. *Turnbull's obstetrics.* 2nd ed., Edinburgh: Churchill Livingstone; 1995, p 185–93.
17 Macri J, Kasturi RV, Krantz BS, *et al.* Maternal serum Down syndrome screening: Free β-protein is a more effective marker than human chorionic gonadotropin. *Am J Obstet Gynecol* 1990;**163**:1248–53.
18 Cuckle HS, Holding S, Jones R, *et al.* Combining inhibin A with existing second-trimester markers in maternal serum screening for Down's syndrome. *Prenat Diagn* 1996;**16**:1095–100.
19 Aitken DA, Wallace EM, Crossley JA, *et al.* Dimeric inhibin-A as a marker for Down's syndrome in early pregnancy. *N Engl J Med* 1996;**334**:1231–6.
20 Forest JC, Masse J, Rousseau F, *et al.* Screening for Down's syndrome during the first and second trimesters: impact of risk estimation parameters. *Clin Biochem* 1995;**28**:443–9.
21 Spencer K. Evaluation of an assay of the free b-subunit of choriogonadotropin and its potential value in screening for Down's syndrome. *Clin Chem* 1991;**37**:809–14.
22 Knight GJ, Palomaki GE, Neveux LM, *et al.* hCG and the free b-subunit as screening tests for Down syndrome. *Prenat Diagn* 1998;**18**:235–45.
23 Wald NJ, Densem JW, Smith D, Klee GG. Four-marker serum screeeing for Down's syndrome. *Prenat Diagn* 1994;**14**:707–16.
24 Spencer K, Macri JN. Early detection of Down's syndrome using free beta human chorionic gonadotrophin. *Ann Clin Biochem* 1992;**29**:349–50.
25 Crossley JA, Aitken DA, Connor JM. Prenatal screening for chromosome abnormalities using maternal serum chorionic gonadotrophin, alpha-fetoprotein, and age. *Prenat Diagn* 1991;**11**:83–101.
26 Norgaard Pedersen B, Alfthan H, Arends J, *et al.* A new simple and rapid dual assay for AFP and free β-hCG in screening for Down syndrome. *Clin Genet* 1994;**45**:1–4.
27 Ryall RG, Staples AJ, Robertson EF, Pollard AC. Improved performance in a prenatal screening programme for Down's syndrome incorporating serum-free-hCG subunit analyses. *Prenat Diagn* 1992;**12**:251–61.
28 Macri JN, Anderson RW, Krantz DA, *et al.* Prenatal maternal dried blood screening with AFP and free β–hCG for open neural tube defects and Down's syndrome. *Am J Obstet Gynecol* 1996;**174**:566–72.
29 Wenstrom KD, Owen J, Chu DC, Boots L. Free β-hCG subunit versus intact hCG in Down's syndrome screening. *Obstet Gynecol* 1997;**90**:370–4.
30 Spencer K, Coombes EJ, Mallard AS, Milford-Ward A. Free beta human choriogonadotropin in Down's syndrome screening: a multicentre study of its role compared with other biochemical markers. *Ann Clin Biochem* 1992;**29**:506–18.
31 Stone S, Henley R, Reynolds T, John R. A comparison of total and free β-subunit assays in Down syndrome screening. *Prenat Diagn* 1993;**13**:535–7.
32 Spencer K. Second trimester prenatal screening for Down's syndrome using AFP and free β-hCG: a 7 year review. *Br J Obstet Gynecol* 1999;**106**:1287–93.
33 Macri JN, Spencer K, Anderson R. Dual analyte immunoassay – a new approach to neural tube defect and Down syndrome screening. *Ann Clin Biochem* 1992;**29**:390–6.
34 Hsu JJ, Hsieh TT, Soong YK, Spencer K. Comparison of Down's syndrome screening strategies in Asians. Combining serum free β-hCG and AFP with maternal age. *Prenat Diagn* 1997;**17**:707–16.
35 Chao AS, Chung CL, Wu CD, *et al.* Second trimester maternal serum screening using AFP, free β-hCG and maternal age. Result of chromosomal abnormalities detected in screen positive for Down syndrome in an Asian population. *Acta Obstet Gynecol Scand* 1999;**78**:393–7.
36 Macri JN, Anderson RW, Kasturi RV, *et al.* Enhanced Down syndrome detection with free beta (hCG). *Am J Hum Genet* 1992;**51**:A418.

37 Spencer K, Souter V, Tul N, *et al.* A screening programme for trisomy 21 at 10–14 weeks using fetal NT, maternal serum free β-hCG and PAPP-A. *Ultrasound Obstet Gynecol* 1999;**13**:231–7.
38 Wald NJ, Hackshaw AK, George LM. Assay precision of serum (a fetoprotein in antenatal screening for neural tube defects and Down's syndrome. *J Med Screen* 2000;**7**:74–7.
39 Wald NJ, George L, Smith D, *et al.* (On behalf of the International Prenatal Screening Research Group). Serum screening for Down's syndrome between 8 and 14 weeks of pregnancy. *Br J Obstet Gynaecol* 1996;**103**:407–12.
40 Cicero S, Curcio P, Papageorghiou A, *et al.* Absence of nasal bone in fetuses with trisomy 21 at 11–14 weeks of gestation: an observational study. *Lancet* 2001;**358**:1665–7.
41 De Biasio P, Venturini PL. Absence of nasal bone and detection of trisomy 21. *Lancet* 2002;**359**:1344.
42 Ward BE, Gersen SL, Carelli MP, *et al.* Rapid prenatal diagnosis of chromosomal aneuploidies by fluorescence in situ hybridisation: clinical experience with 4500 specimens. *Am J Hum Genet* 1993;**52**:854–65.
43 Levett IJ, Liddle S, Meredith R. A large scale evaluation of amnio-PCR for the rapid prenatal diagnosis of fetal trisomy. *Ultrasound Obstet Gynecol* 2001;**17**:115–18.
44 Mann K, Fox SP, Abbs SJ, *et al.* Development and implementation of a new rapid aneuploidy diagnostic service within the UK National Health Service and implications for the future of prenatal diagnosis. *Lancet* 2001;**358**:1057–61.
45 Lawson HW, Frye A, Altrash HK, *et al.* Obstetrics: abortion mortality, United States, 1972 through 1987. *Am J Obstet Gynecol* 1994;**171**:1365–72.
46 Chitty L, Statham H, Sololou W, Green J. Grief after termination for fetal abnormality – is there a gestational age effect? *Ultrasound Obstet Gynecol* 1998;**12** Suppl 1:3.
47 Bischoff FZ, Lewis D, Simpson JL, *et al.* Isolating fetal cells from maternal blood: strategies to increase sensitivity for fetal aneuploidy detection. *Am J Hum Genet* 1999;**65** Suppl:A943.
48 Bianchi DW, Simpson JL, Jackson LG, *et al.* Fetal gender and aneuploidy detection using fetal cells in maternal blood: analysis of NIFTY I data. *Prenat Diagn* 2002;**22**:609–15.
49 Holzgreve W. Prenatal diagnosis using fetal cells and free fetal DNA in maternal blood. *Clin Perinatol* 2001;**28**:353–65.
50 Wald N, Huttly W, Hackshaw A. Antenatal screening for Down's syndrome with the quadruple test: 5 year results from a screening programme. *Lancet* 2003;**361**:835–6.

APPENDIX

Supplementary figures and tables

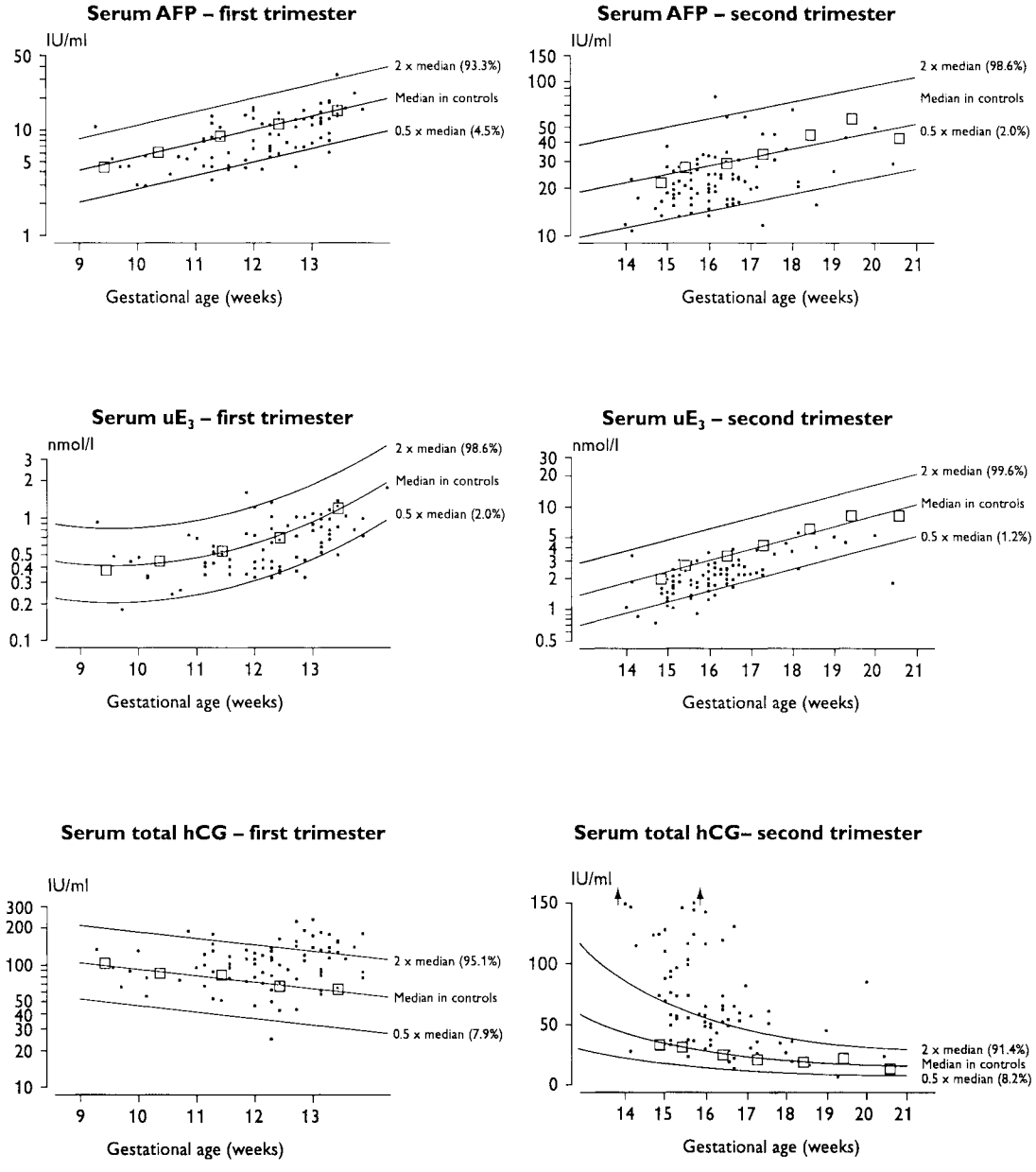

Figure 11 Relationship between gestational age and serum AFP, uE_3 and total hCG; first (9–13 completed weeks) and second (14–20 completed weeks) trimesters. Open squares are medians in controls. Dots are cases. The numbers of unaffected pregnancies were 19, 32, 125, 181 and 133 at 9, 10, 11, 12 and 13 completed weeks respectively, and 18, 187, 194, 45, 15, 14, and 19 at 14, 15, 16, 17, 18, 19 and 20 completed weeks respectively. The centiles corresponding to twice the median and half the median for unaffected pregnancies are shown in brackets.

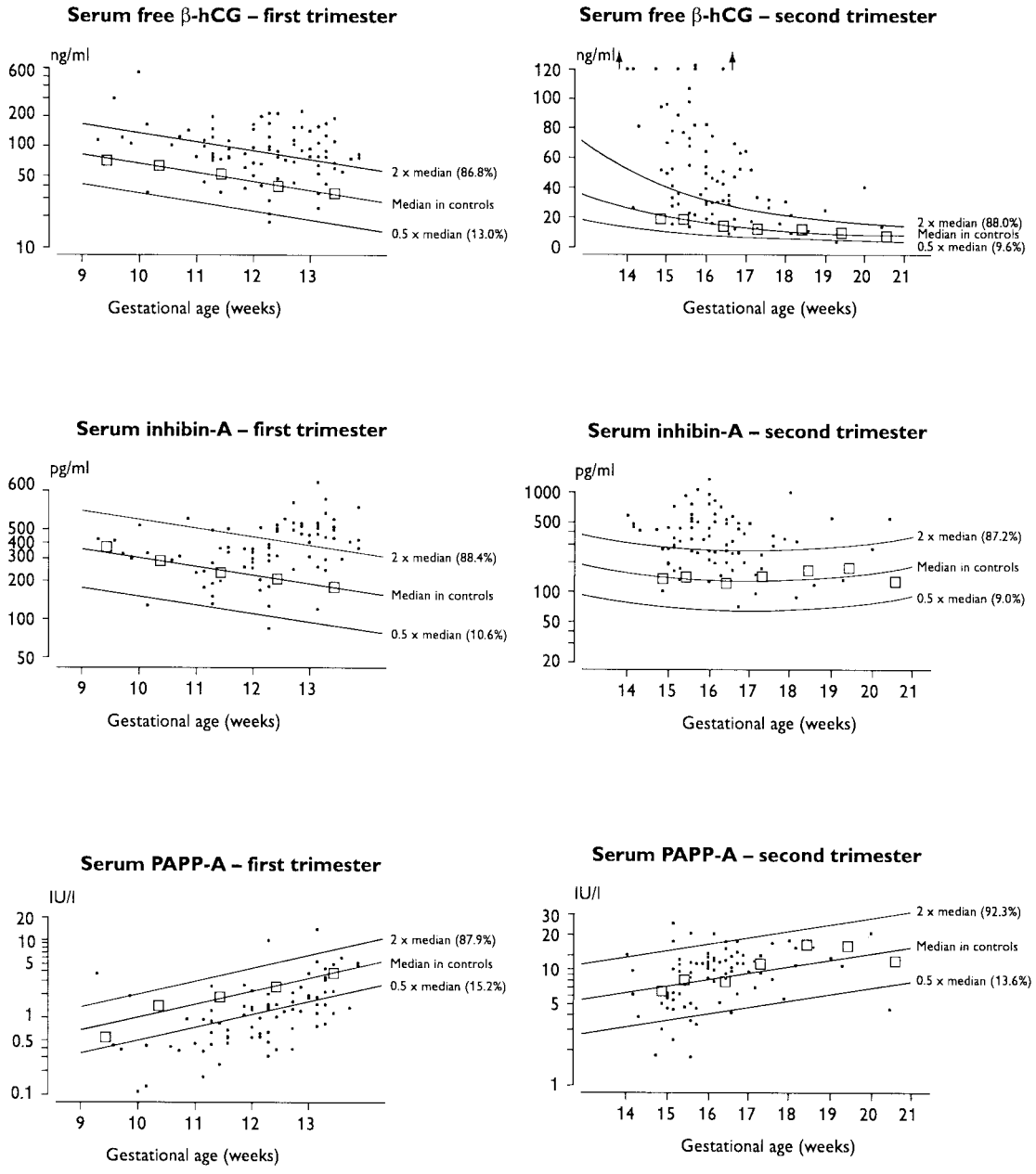

Figure 12 Relationship between gestational age and serum free β-hCG, inhibin-A and PAPP-A; first (9–13 completed weeks) and second (14–20 completed weeks) trimesters. Open squares are medians in controls. Dots are cases. The numbers of unaffected pregnancies were 19, 32, 125, 181 and 133 at 10, 11, 12 and 13 completed weeks respectively, and 18, 187, 194, 45, 15, 14, and 19 at 14, 15, 16, 17, 18, 19 and 20 completed weeks respectively. The centiles corresponding to twice the median and half the median for unaffected pregnancies are shown in brackets. The seven high free β-hCG concentrations in the second trimester (indicated by the arrows) are 359, 191, 163, 132, 486, 486 and 135 ng/ml from left to right.

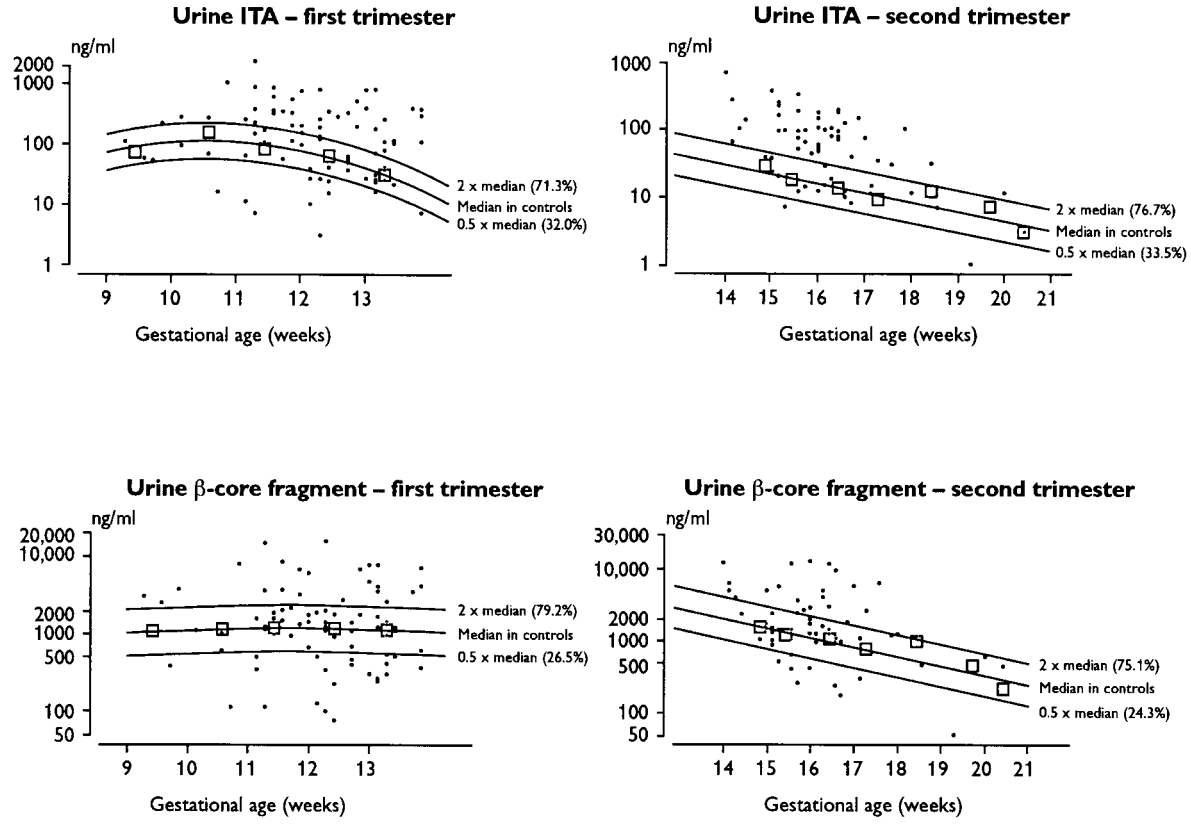

Figure 13 Relationship between gestational age and urine ITA, β-core fragment; first (9–13 completed weeks) and second (14–20 completed weeks) trimesters. Open squares are medians in controls. Dots are cases. The numbers of affected pregnancies were 19, 38, 120, 179 and 130 at 10, 11, 12, and 13 completed weeks respectively, and 18, 163, 183, 42, 14, 25 and 45 at 14, 15, 16, 17, 18 19, and 20 completed weeks respectively. Centiles corresponding to twice the median and half the median for unaffected pregnancies are shown in brackets.

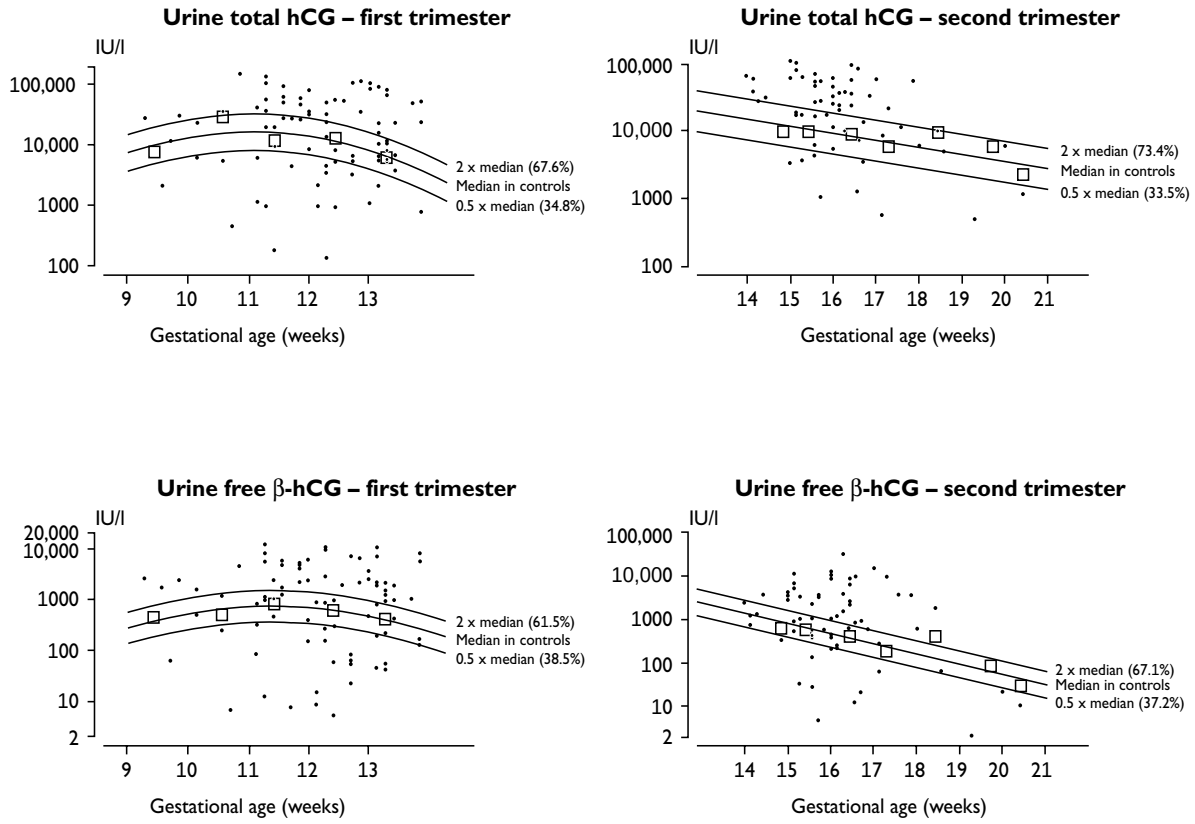

Figure 14 Relationship between gestational age of the urine sample and total hCG and free β-hCG; first (9–13 completed weeks) and second (14–20 completed weeks) trimesters. Open squares are medians in controls. Dots are cases. The numbers of affected pregnancies were 19, 38, 120, 179 and 130 at 10, 11, 12, and 13 completed weeks respectively, and 18, 163, 183, 42, 14, 25 and 45 at 14, 15, 16, 17, 18 19, and 20 completed weeks respectively. Centiles corresponding to twice the median and half the median for unaffected pregnancies are shown in brackets.

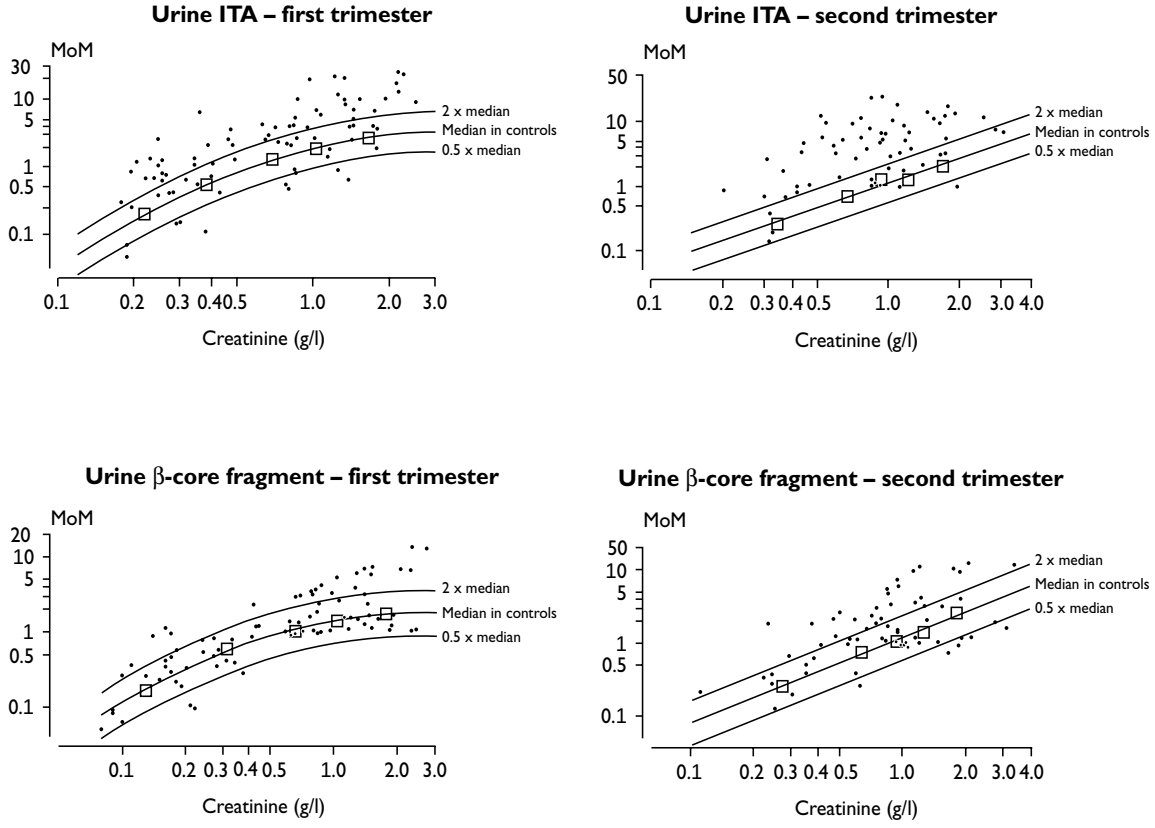

Figure 15 Relationship between creatinine concentration and MoM values for urinary ITA and β-core fragment. ITA values were adjusted for the creatinine concentration, as suggested by Cole and co-workers.[5] Open squares are medians in controls, dots are cases.

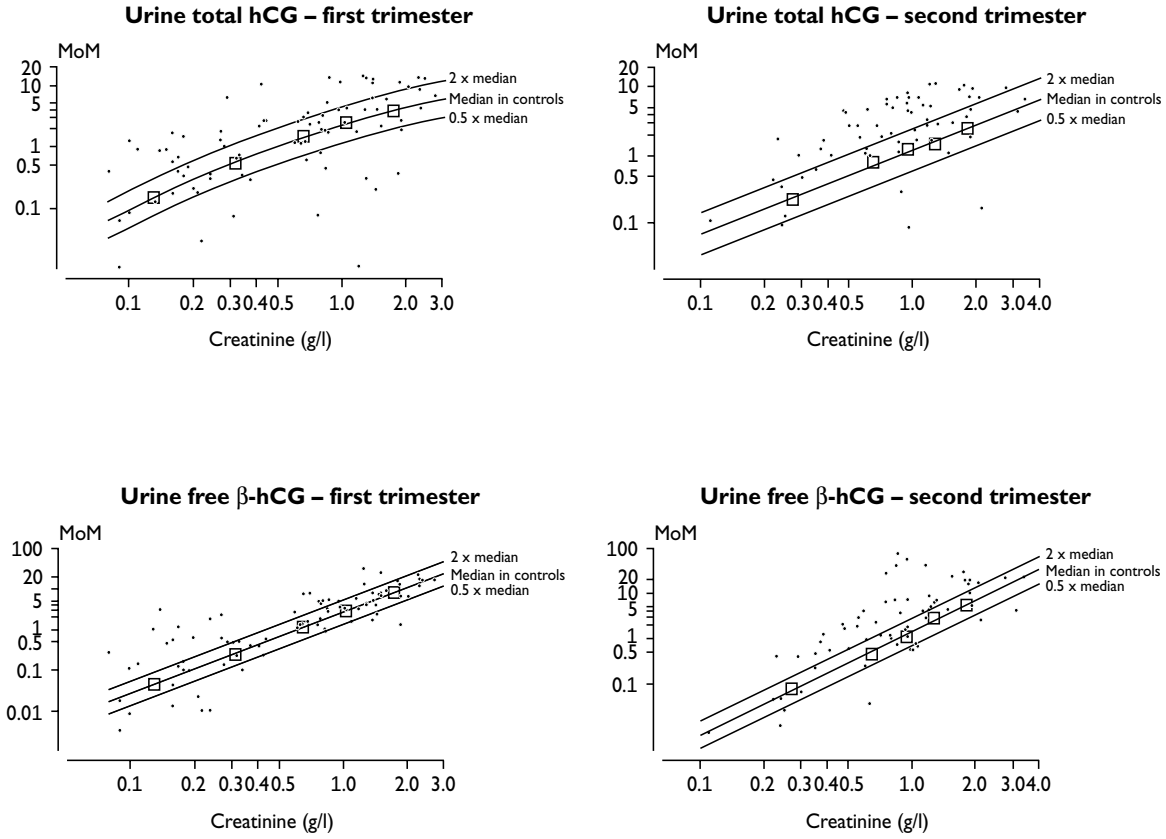

Figure 16 Relationship between creatinine concentration and MoM values for urinary total hCG and free β-hCG. Open squares are medians in controls, dots are cases.

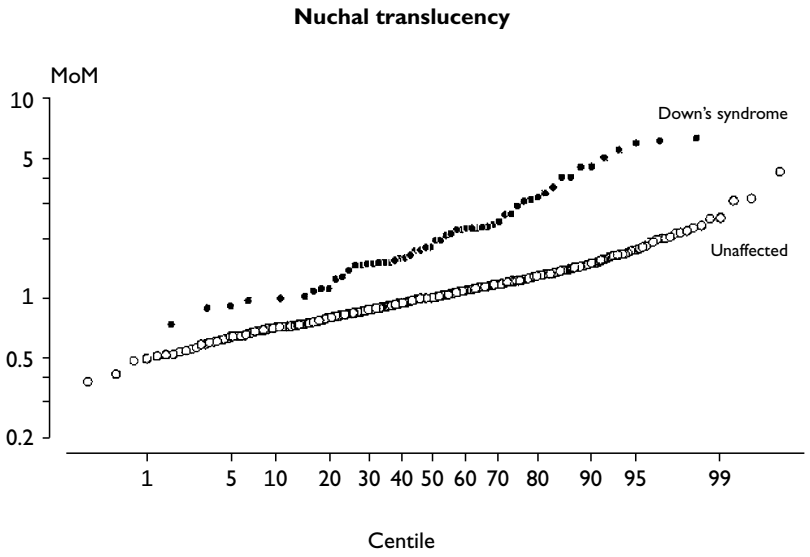

Figure 17 Probability plot for first trimester NT (MoM) in affected and unaffected pregnancies.

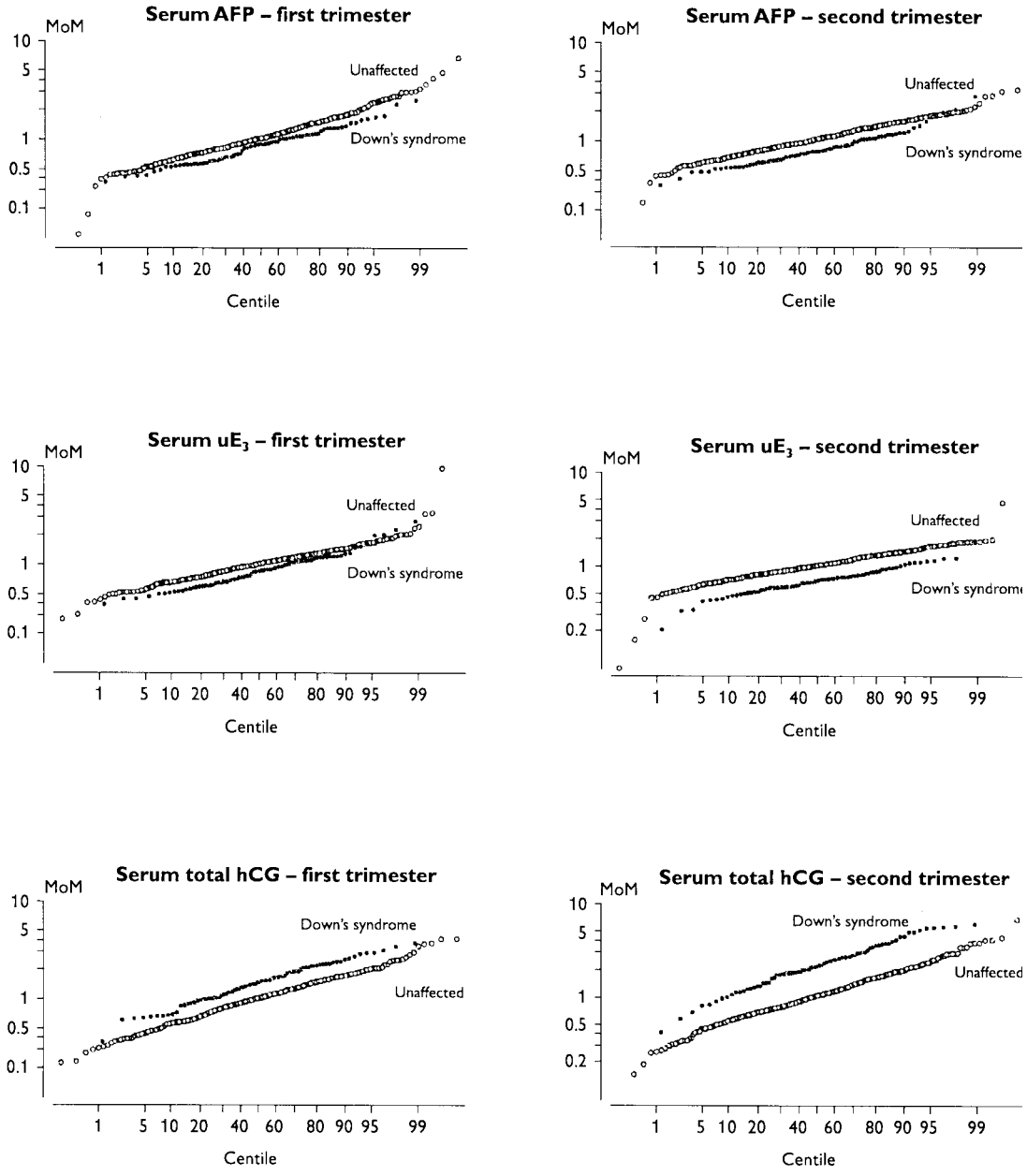

Figure 18 Probability plot for serum AFP, uE_3 and total hCG (MoM) in affected and unaffected pregnancies, in the first and second trimesters of pregnancy.

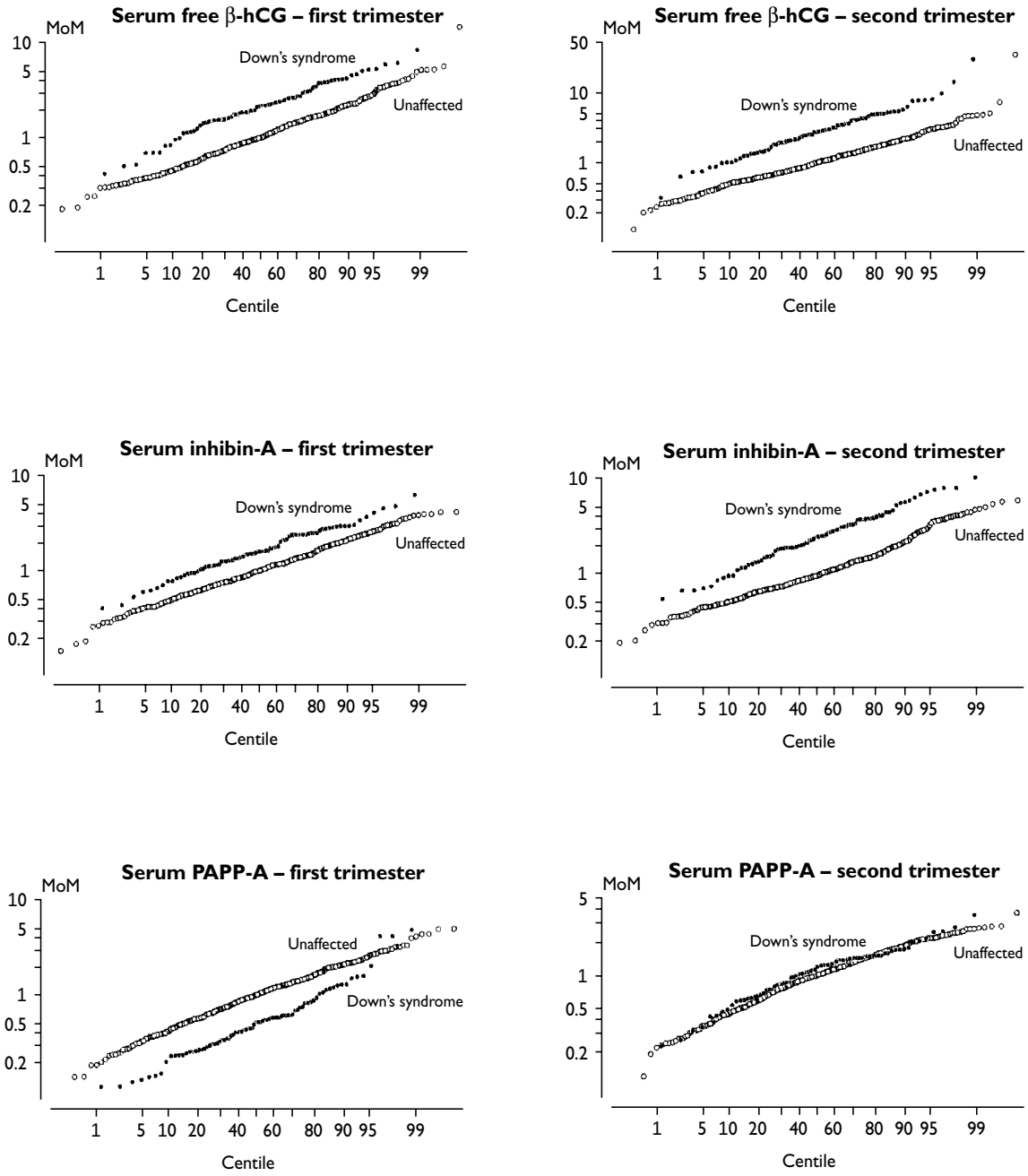

Figure 19 Probability plot for serum free β-hCG, inhibin-A and PAPP-A (MoM) in affected and unaffected pregnancies, in the first and second trimesters of pregnancy.

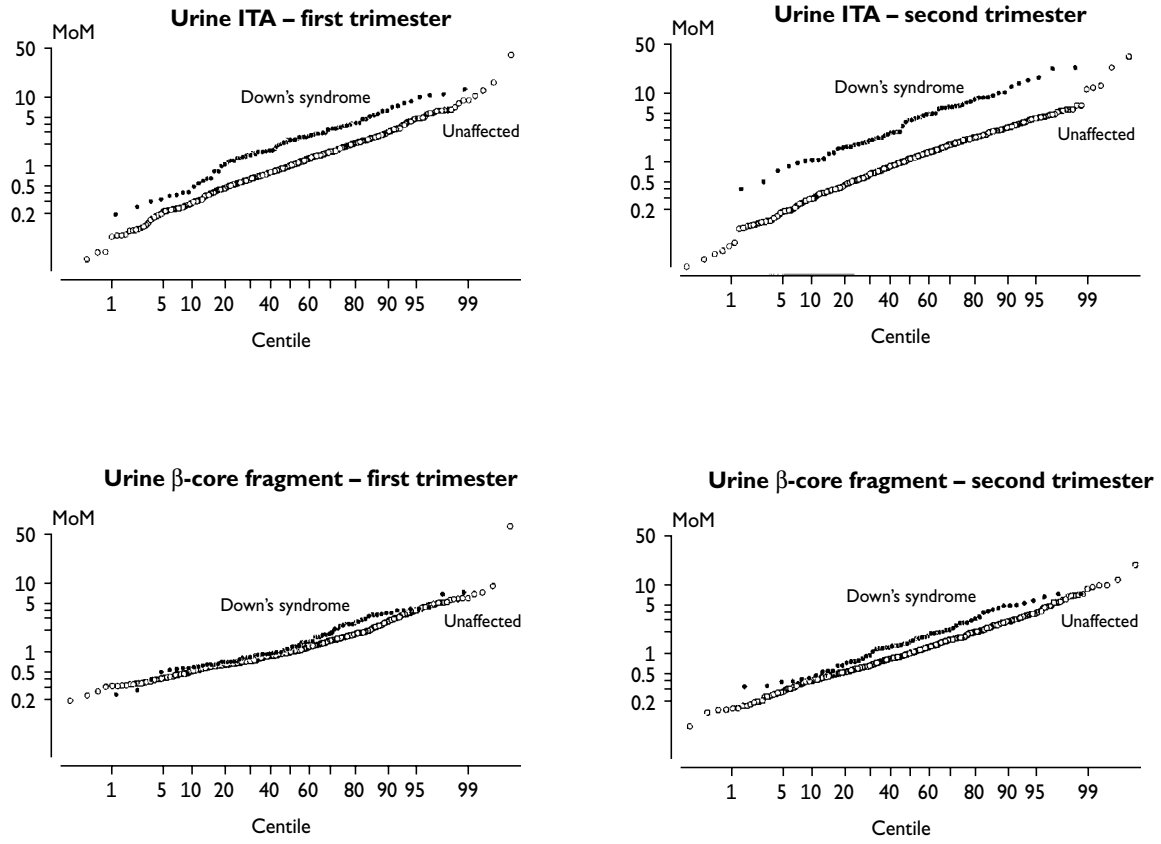

Figure 20 Probability plot for urinary ITA and β-core fragment (MoM) in affected and unaffected pregnancies, in the first and second trimesters of pregnancy.

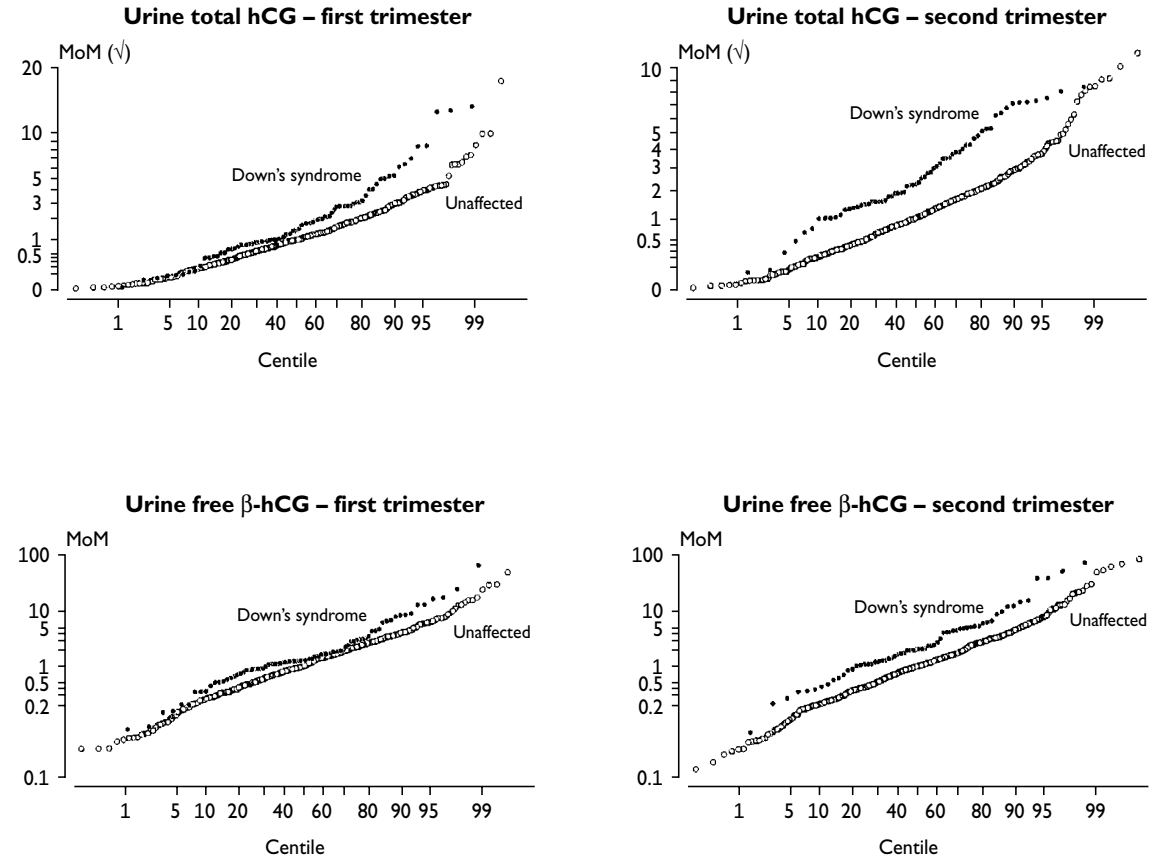

Figure 21 Probability plot for urinary total hCG and free β-hCG (MoM) in affected and unaffected pregnancies, in the first and second trimesters of pregnancy.

Table 33 Median marker level (log_{10}MoM) in Down's syndrome pregnancies according to marker and gestational age

	Completed week of pregnancy				
	10	11	12	13	14–22
NT		0.2922 (10–13 weeks)			–
AFP[a]		–0.0655 (10–13 weeks)			–0.1308
uE_3[a]	–0.0044	–0.0605	–0.1024	–0.1427	–0.1549
Total hCG[a]	–0.0177	0.1038	0.1875	0.2742	0.3118
Free β-hCG[a]	0.2095	0.2878	0.3404	0.3944	0.4249
Inhibin-A[a]	–0.0269	0.1303	0.2380	0.3384	0.3384
PAPP-A[a]	–0.4685	–0.3768	–0.3010	–0.2366	0.0453

[a] Regression equations used to estimate the median in affected pregnancies in the first trimester are: median $uE_3 = 10^{0.4137 - 0.00595 \times \text{gestational age in days}}$; median total hCG = $10^{-0.8662 + 0.01213 \times \text{gestational age in days}}$; median free β-hCG = $10^{-0.3271 + 0.00768 \times \text{gestational age in days}}$; median inhibin-A = $10^{-1.1134 + 0.01554 \times \text{gestational age in days}}$; median PAPP-A = $10^{-1.1444 + 0.00965 \times \text{gestational age in days}}$, so, for example, if a woman is 12 weeks and 3 days, her gestational age is 87 days. The median PAPP-A level used to estimate her risk is then $10^{-1.1444 + 0.00965 \times 87} = 0.50$ MoM which can be log transformed to –0.3010 ($log_{10}0.50$).

Table 34 SDs (log_{10}) for the screening markers in each trimester of pregnancy

	Gestational age based on			
	Dates (LMP)		Scan	
Maternal weight correction:	No	Yes	No	Yes
Down's syndrome				
First trimester (10–13 weeks)				
NT[a]	–	–	0.2313	0.2313
AFP	0.1881	0.1850	0.1752	0.1672
uE_3	0.1893	0.1892	0.1723	0.1720
Total hCG	0.2202	0.2111	0.2151	0.2069
Free β-hCG	0.2651	0.2637	0.2587	0.2569
Inhibin-A	0.2428	0.2368	0.2402	0.2343
PAPP-A	0.3161	0.2959	0.3006	0.2802
Second trimester (14–20 weeks)				
AFP	0.1497	0.1416	0.1485	0.1398
uE_3	0.1377	0.1370	0.1251	0.1238
Total hCG	0.2424	0.2397	0.2422	0.2395
Free β-hCG	0.3020	0.2996	0.2987	0.2965
Inhibin-A	0.2723	0.2679	0.2723	0.2679
PAPP-A	0.2114	0.1878	0.2102	0.1872
Unaffected				
First trimester (10–13 weeks)				
NT[a]				
10 completed weeks	–	–	0.1732	0.1732
11 completed weeks	–	–	0.1439	0.1439
12–13 completed weeks	–	–	0.1329	0.1329
AFP	0.2012	0.1983	0.1892	0.1818
uE_3	0.1440	0.1438	0.1208	0.1204
Total hCG	0.2091	0.1994	0.2037	0.1950
Free β-hCG	0.2731	0.2718	0.2669	0.2651
Inhibin-A	0.2282	0.2218	0.2254	0.2191
PAPP-A	0.2893	0.2670	0.2722	0.2495
Second trimester (14–20 weeks)				
AFP	0.1498	0.1417	0.1486	0.1399
uE_3	0.1292	0.1284	0.1156	0.1142
Total hCG	0.2307	0.2279	0.2305	0.2276
Free β-hCG	0.2639	0.2612	0.2602	0.2577
Inhibin-A	0.2135	0.2078	0.2135	0.2078
PAPP-A	0.2393	0.2187	0.2382	0.2181

[a] NT measurement was not corrected for maternal weight because there was no association between the NT and maternal weight
The truncation limits in the first trimester were: NT (0.5–2.5), AFP (0.4–3.0), uE_3 (0.4–2.0), total hCG (0.3–3.0), free β-hCG (0.3–5.0), inhibin-A (0.3–5.0), PAPP-A (0.2–3.0). The truncation limits in the second trimester were: AFP (0.4–3.0), uE_3 (0.4–2.0), total hCG (0.4–5.0), free β-hCG (0.3–5.0), inhibin-A (0.3–5.0), PAPP-A (0.2–3.0).

Table 35 SDs (log_{10} except square root in the case of total hCG) for the urine markers, based on a scan estimate of gestational age with maternal weight correction of the marker levels

	First trimester (10–13 completed weeks)	Second trimester (14–20 completed weeks)
Down's syndrome		
ITA	0.3821	0.3994
β-core fragment	0.3214	0.3720
Total hCG	0.6732	0.6360
Free β-hCG	0.3989	0.5195
Unaffected		
ITA	0.3676	0.3945
β-core fragment	0.2646	0.3297
Total hCG	0.4672	0.4771
Free β-hCG	0.4561	0.5333

The truncation limits in the first trimester were: ITA (0.2–10.0), β-core fragment (0.4–5.0), total hCG (0.2–4.0), free β-hCG (0.2–10.0). The truncation limits in the second trimester were ITA (0.2–10.0), β-core fragment (0.2–5.0), total hCG (0.1–5.0), free β-hCG (0.2–10.0).

Table 36 Correlation coefficients for the serum markers in Down's syndrome pregnancies. Gestational age based on dates (LMP). Marker levels not adjusted for maternal weight

	First trimester						Second trimester				
	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester											
uE_3	0.2686										
Total hCG	–0.1341	–0.1391									
Free β-hCG	–0.1735	–0.2710	0.5115								
Inhibin-A	–0.0792	–0.1923	0.4092	0.3049							
PAPP-A	0.3923	0.4463	0.1330	–0.0456	0.1245						
Second trimester											
AFP	0.5175	0.1389	0.1416	0.1063	0.0811	0.1945					
uE_3	0.1503	0.7736	–0.1656	–0.3446	–0.2393	0.4586	–0.0017				
Total hCG	–0.1727	–0.1917	0.6809	0.4695	0.2808	–0.1761	0.2121	–0.3796			
Free β-hCG	–0.1936	–0.2126	0.5605	0.7675	0.2390	–0.2484	0.2060	–0.4172	0.8160		
Inhibin-A	–0.0270	–0.1058	0.2618	0.3031	0.6243	–0.1282	0.1976	–0.2989	0.4296	0.4394	
PAPP-A	0.1050	0.2476	0.1818	0.0638	0.1459	0.8557	0.1675	0.5001	–0.0451	–0.0868	–0.0160

Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 37 Correlation coefficients for the serum markers in unaffected pregnancies. Gestational age based on dates (LMP). Marker levels not adjusted for maternal weight

	First trimester						Second trimester				
	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester											
uE_3	0.2919										
Total hCG	–0.0029	0.0280									
Free β-hCG	–0.0158	–0.0207	0.7023								
Inhibin-A	0.0556	–0.0608	0.5719	0.5036							
PAPP-A	0.2185	0.3018	0.2067	0.1315	0.2305						
Second trimester											
AFP	0.5670	0.1853	0.1085	0.0571	0.1384	0.2461					
uE_3	0.2629	0.6708	0.0044	–0.0479	–0.0997	0.2820	0.2595				
Total hCG	0.0367	–0.0115	0.7050	0.5667	0.4084	0.0816	0.1781	–0.0825			
Free β-hCG	0.0067	–0.0264	0.6968	0.7494	0.4212	0.0692	0.1168	–0.0904	0.8607		
Inhibin-A	0.1437	–0.0572	0.3295	0.3094	0.6939	0.0652	0.2293	–0.0876	0.4412	0.4122	
PAPP-A	0.1752	0.1506	0.3858	0.2812	0.3833	0.7012	0.2802	0.1875	0.2873	0.2732	0.2728

Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 38 Correlation coefficients for the serum markers in Down's syndrome pregnancies. Gestational age based on dates (LMP). Marker levels adjusted for maternal weight

	First trimester						Second trimester				
	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester											
uE_3	0.2660										
Total hCG	–0.1938	–0.1524									
Free β-hCG	–0.2221	–0.2787	0.4929								
Inhibin-A	–0.1288	–0.2045	0.3919	0.2762							
PAPP-A	0.3526	0.4666	0.0749	–0.1112	0.0712						
Second trimester											
AFP	0.4877	0.1372	0.0879	0.0554	0.0272	0.1211					
uE_3	0.1360	0.7756	–0.1921	–0.3646	–0.2643	0.4671	–0.0288				
Total hCG	–0.2206	–0.2001	0.6758	0.4418	0.2535	–0.2487	0.1725	–0.4008			
Free β-hCG	–0.2333	–0.2193	0.5551	0.7494	0.2170	–0.3170	0.1738	–0.4366	0.8046		
Inhibin-A	–0.0621	–0.1124	0.2442	0.2824	0.6215	–0.1879	0.1661	–0.3182	0.4148	0.4284	
PAPP-A	0.0284	0.2659	0.1222	–0.0052	0.0869	0.8963	0.0762	0.5320	–0.1243	–0.1602	–0.0799

Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 39 Correlation coefficients for the serum markers in unaffected pregnancies. Gestational age based on dates (LMP). Marker levels adjusted for maternal weight

	First trimester						Second trimester				
	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester											
uE_3	0.2877										
Total hCG	–0.0540	0.0193									
Free β-hCG	–0.0565	–0.0287	0.6953								
Inhibin-A	0.0120	–0.0728	0.5653	0.4812							
PAPP-A	0.1642	0.3123	0.1490	0.0763	0.1799						
Second trimester											
AFP	0.5442	0.1834	0.0481	0.0048	0.0857	0.1728					
uE_3	0.2510	0.6726	–0.0162	–0.0654	–0.1234	0.2776	0.2473				
Total hCG	–0.0045	–0.0201	0.7010	0.5401	0.3832	0.0214	0.1336	–0.1008			
Free β-hCG	–0.0304	–0.0340	0.6967	0.7292	0.4011	0.0127	0.0722	–0.1089	0.8475		
Inhibin-A	0.1087	–0.0670	0.3095	0.2852	0.6934	0.0030	0.1896	–0.1083	0.4225	0.3957	
PAPP-A	0.1210	0.1504	0.3596	0.2447	0.3567	0.7052	0.2184	0.1755	0.2524	0.2411	0.2385

Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 40 Correlation coefficients for NT and the serum markers in Down's syndrome pregnancies. Gestational age based on scan. Marker levels not adjusted for maternal weight

	First trimester							Second trimester				
	NT	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester												
AFP	0.0484											
uE_3	0.1234	0.1676										
Total hCG	–0.0804	–0.1092	–0.1042									
Free β-hCG	0.1063	–0.1576	–0.2466	0.5272								
Inhibin-A	–0.1272	–0.0585	–0.1784	0.4169	0.3123							
PAPP-A	–0.1420	0.2967	0.3481	0.1894	0.0004	0.1672						
Second trimester												
AFP	0.0744	0.5248	0.1120	0.1627	0.1250	0.0937	0.1580					
uE_3	0.0679	0.0190	0.7321	–0.1410	–0.3387	–0.2245	0.3788	–0.0670				
Total hCG	0.0451	–0.1619	–0.1594	0.6960	0.4869	0.2886	–0.1552	0.2276	–0.3540			
Free β-hCG	0.1454	–0.1767	–0.1907	0.5803	0.7955	0.2503	–0.2293	0.2294	–0.4106	0.8279		
Inhibin-A	0.1818	–0.0166	–0.0882	0.2668	0.3117	0.6294	–0.1217	0.2069	–0.3035	0.4319	0.4485	
PAPP-A	0.0160	0.0471	0.2075	0.2041	0.0909	0.1643	0.8548	0.1452	0.4679	–0.0191	–0.0619	–0.0063

The correlations with NT are based on sonographer-specific medians and satisfactory NT images. Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 41 Correlation coefficients for NT and the serum markers in unaffected pregnancies. Gestational age based on scan. Marker levels not adjusted for maternal weight

	First trimester									Second trimester				
	NT (10 completed weeks)	NT (11 completed weeks)	NT (12–13 completed weeks)	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester														
AFP	0.0196	0.0236	0.0255											
uE_3	0.0542	0.0652	0.0706	0.1730										
Total hCG	–0.0626	–0.0753	–0.0815	0.0341	0.1129									
Free β-hCG	–0.0334	–0.0402	–0.0436	0.0127	0.0557	0.7281								
Inhibin-A	–0.0662	–0.0797	–0.0863	0.0870	–0.0200	0.5869	0.5181							
PAPP-A	–0.0417	–0.0503	–0.0544	0.0975	0.1184	0.2795	0.1961	0.2891						
Second trimester														
AFP	–0.0097	–0.0116	–0.0126	0.5752	0.1630	0.1298	0.0736	0.1531	0.2105					
uE_3	0.0476	0.0573	0.0621	0.1536	0.5798	0.0572	–0.0023	–0.0647	0.1533	0.2219				
Total hCG	–0.0556	–0.0669	–0.0725	0.0621	0.0633	0.7225	0.5858	0.4189	0.1217	0.1939	–0.0193			
Free β-hCG	–0.0507	–0.0610	–0.0660	0.0429	0.0425	0.7258	0.7781	0.4399	0.1186	0.1417	–0.0357	0.8764		
Inhibin-A	–0.0414	–0.0499	–0.0540	0.1675	–0.0172	0.3366	0.3179	0.7003	0.0877	0.2411	–0.0627	0.4441	0.4242	
PAPP-A	–0.0499	–0.0601	–0.0651	0.1330	0.0972	0.4136	0.3105	0.4049	0.6999	0.2520	0.1287	0.3134	0.3050	0.2851

The correlations with NT are based on sonographer-specific medians and satisfactory NT images. Correlations are estimated after excluding outliers ±3.5 SDs from the median.

Table 42 Correlation coefficients for NT and the serum markers in Downs syndrome pregnancies. Gestational age based on scan. Marker levels adjusted for maternal weight[a]

	First trimester							Second trimester				
	NT	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester												
AFP	0.0526											
uE_3	0.1240	0.1626										
Total hCG	–0.0819	–0.1808	–0.1200									
Free β-hCG	0.1080	–0.2201	–0.2588	0.5053								
Inhibin-A	–0.1292	–0.1169	–0.1942	0.3970	0.2830							
PAPP-A	–0.1506	0.2374	0.3562	0.1284	–0.0692	0.1119						
Second trimester												
AFP	0.0809	0.5003	0.1036	0.1075	0.0697	0.0374	0.0660					
uE_3	0.0695	–0.0099	0.7356	–0.1741	–0.3666	–0.2557	0.3712	–0.1093				
Total hCG	0.0466	–0.2186	–0.1703	0.6912	0.4598	0.2629	–0.2295	0.1920	–0.3808			
Free β-hCG	0.1471	–0.2262	–0.1999	0.5735	0.7797	0.2291	–0.3004	0.1981	0.4356	0.8178		
Inhibin-A	0.1854	–0.0554	–0.0969	0.2493	0.2909	0.6269	–0.1842	0.1770	–0.3276	0.4197	0.4384	
PAPP-A	0.0202	–0.0550	0.2129	0.1432	0.0199	0.1049	0.8798	0.0424	0.4837	–0.0912	–0.1319	–0.0677

[a] Except NT, which is not associated with maternal weight.

Table 43 Correlation coefficients for NT and the serum markers in unaffected pregnancies. Gestational age based on scan. Marker levels adjusted for maternal weight[a]

	First trimester									Second trimester				
	NT (10 completed weeks)	NT (11 completed weeks)	NT (12–13 completed weeks)	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A	PAPP-A	AFP	uE_3	Total hCG	Free β-hCG	Inhibin-A
First trimester														
AFP	0.0227	0.0273	0.0296											
uE_3	0.0550	0.0662	0.0717	0.1632										
Total hCG	–0.0630	–0.0758	–0.0821	–0.0232	0.1010									
Free β-hCG	–0.0325	–0.0391	–0.0423	–0.0346	0.0422	0.7178								
Inhibin-A	–0.0663	–0.0798	–0.0865	0.0400	–0.0374	0.5771	0.4958							
PAPP-A	–0.0429	–0.0516	–0.0559	0.0114	0.1009	0.2198	0.1395	0.2380						
Second trimester														
AFP	–0.0079	–0.0095	–0.0103	0.5587	0.1515	0.0675	0.0167	0.0982	0.1160					
uE_3	0.0495	0.0596	0.0645	0.1318	0.5803	0.0306	–0.0255	–0.0942	0.1213	0.1981				
Total hCG	–0.0549	–0.0661	–0.0716	0.0199	0.0512	0.7191	0.5606	0.3956	0.0624	0.1535	–0.0416			
Free β-hCG	–0.0502	–0.0604	–0.0654	0.0031	0.0308	0.7236	0.7605	0.4208	0.0627	0.0974	–0.0585	0.8651		
Inhibin-A	–0.0415	–0.0499	–0.0540	0.1344	–0.0307	0.3167	0.2937	0.7003	0.0237	0.2033	–0.0875	0.4293	0.4092	
PAPP-A	–0.0520	–0.0625	–0.0677	0.0640	0.0815	0.3853	0.2725	0.3775	0.6895	0.1918	0.0983	0.2838	0.2752	0.2530

The correlations with NT are based on sonographer-specific medians and satisfactory NT images. Correlations are estimated after excluding outliers ±3.5 SDs from the median. [a] Except NT which is not associated with maternal weight.

Table 44 Screening performance of selected first trimester markers at 10–13 completed weeks of pregnancy (with maternal age)

	With NT									Without NT								
	DR (%) for FPR of:			FPR (%) for DR of:						DR (%) for FPR of:			FPR (%) for DR of:					
	1%	3%	5%	70%	75%	80%	85%	90%	95%	1%	3%	5%	70%	75%	80%	85%	90%	95%
10 completed weeks																		
PAPP-A	59	73	79	2.3	3.5	5.3	8.4	13.9	26	44	59	67	6.0	8.3	11.7	16.5	25	39
PAPP-A, inhibin	60	74	80	2.2	3.3	5.0	7.9	13.3	25	44	60	68	5.8	8.0	11.5	16.2	24	39
PAPP-A, free β-hCG	66	78	83	1.5	2.3	3.7	6.1	10.8	22	52	67	74	3.8	5.5	8.0	12.1	19	33
PAPP-A, total hCG	60	74	80	2.2	3.3	5.1	8.0	13.4	26	44	60	68	5.8	8.0	11.3	16	24	39
PAPP-A, free β-hCG, inhibin	66	78	83	1.4	2.2	3.5	5.9	10.4	21	53	68	74	3.6	5.2	7.6	11.5	18	32
PAPP-A, total hCG, inhibin	61	75	81	2.0	3.1	4.7	7.5	12.7	24	45	61	69	5.4	7.6	10.7	15	23	38
11 completed weeks																		
PAPP-A	58	71	77	2.6	4.1	6.5	10.5	18	33	32	48	57	10.7	14.2	19	26	36	52
PAPP-A, inhibin	65	77	82	1.6	2.5	4.0	6.8	12.2	25	41	57	64	7.1	9.9	13.7	19	28	45
PAPP-A, free β-hCG	68	79	83	1.2	2.1	3.5	6.0	11.1	23	47	62	70	5.1	7.2	10.3	15	23	38
PAPP-A, total hCG	64	76	81	1.7	2.8	4.5	7.5	13.3	26	39	55	63	7.6	10.5	14.4	20	29	46
PAPP-A, free β-hCG, inhibin	70	80	85	1.0	1.7	2.9	5.0	9.4	20	49	64	72	4.5	6.4	9.2	13.5	21	35
PAPP-A, total hCG, inhibin	67	78	83	1.3	2.1	3.5	6.0	10.8	22	43	59	67	6.2	8.6	12.1	17	26	41
12 completed weeks																		
PAPP-A	58	70	75	3.1	4.9	8.0	13.0	22	39	25	40	49	16	20	26	34	47	62
PAPP-A, inhibin	70	80	85	1.0	1.7	2.8	5.1	9.6	21	43	58	66	6.5	9.1	12.7	18	27	43
PAPP-A, free β-hCG	69	74	83	1.1	1.9	3.4	6.0	11.2	24	44	60	67	5.9	8.3	11.7	17	25	40
PAPP-A, total hCG	67	78	83	1.3	2.1	3.6	6.3	11.7	24	39	55	63	7.8	10.6	14.6	20	30	46
PAPP-A, free β-hCG, inhibin	73	83	87	0.7	1.2	2.1	3.8	7.4	17	50	65	73	4.2	5.9	8.5	12.5	19	32
PAPP-A, total hCG, inhibin	72	82	87	0.8	1.3	2.3	4.1	7.8	17	46	62	69	5.2	7.3	10.3	14.9	22	37
13 completed weeks																		
PAPP-A	56	67	73	3.9	6.2	9.8	16	26	45	20	34	42	21	26	33	42	54	70
PAPP-A, inhibin	74	84	88	0.6	1.1	1.9	3.4	6.6	15	49	64	71	4.5	6.5	9.3	13.7	21	35
PAPP-A, free β-hCG	69	79	84	1.1	1.9	3.2	5.8	10.8	23	44	60	67	5.9	8.2	11.4	16	24	38
PAPP-A, total hCG	71	81	86	0.9	1.6	2.6	4.6	8.7	19	43	60	67	5.9	8.1	11.4	16	24	39
PAPP-A, free β-hCG, inhibin	77	86	90	0.4	0.8	1.4	2.5	5.0	11.9	55	70	77	2.9	4.2	6.2	9.2	14.5	25
PAPP-A, total hCG, inhibin	77	87	90	0.4	0.8	1.3	2.4	4.8	11.3	54	69	76	3.2	4.6	6.6	9.9	15	27
10–13 completed weeks[a]																		
PAPP-A	58	70	76	3.0	4.8	7.6	12.4	21	37	28	43	52	14.4	18	24	31	42	58
PAPP-A, inhibin	68	79	84	1.2	2.0	3.2	5.6	10.1	21	44	59	67	6.2	8.6	12.1	17	26	49
PAPP-A, free β-hCG	68	79	84	1.2	2.0	3.4	6.0	11.0	23	46	61	69	5.4	7.6	10.8	16	24	38
PAPP-A, total hCG	66	78	83	1.4	2.3	3.8	6.5	11.7	24	40	57	64	7.1	9.7	13.4	19	28	44
PAPP-A, free β-hCG, inhibin	72	82	87	0.8	1.4	2.3	4.1	7.8	17	51	66	74	3.9	5.6	8.1	12	18	32
PAPP-A, total hCG, inhibin	70	81	86	1.0	1.6	2.7	4.7	8.6	18	47	63	70	5.1	7.1	10.1	14	22	36

[a] Weighted average of the estimates at each week; weighted by the percentage of women who book at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%) (as observed in SURUSS). These estimates of DR will vary according to local centre; distribution will vary.

Table 45 Performance of integrated screening using first (12 completed weeks) and second trimester (14–20 completed weeks) markers (with maternal age)

Maternal age with:		With NT								
		DR (%) for FPR of:			FPR (%) for DR of:					
First trimester	Second trimester	1%	3%	5%	70%	75%	80%	85%	90%	95%
	AFP, total hCG	70	81	86	1.0	1.6	2.7	4.6	8.4	18
	AFP, free β-hCG	72	82	87	0.8	1.3	2.3	4.1	7.8	17
	AFP, uE_3, total hCG	77	86	90	0.5	0.8	1.4	2.6	5.1	11.8
	AFP, uE_3, free β-hCG	78	87	90	0.4	0.7	1.3	2.4	4.9	11.7
	AFP, uE_3, total hCG, inhibin-A	81	89	92	0.3	0.5	0.9	1.7	3.7	9.4
	AFP, uE_3, free β-hCG, inhibin-A	82	89	92	0.2	0.4	0.8	1.6	3.5	9.2
PAPP-A	AFP, total hCG	74	84	88	0.6	1.0	1.8	3.4	6.6	14.9
PAPP-A	AFP, free β-hCG	76	85	89	0.5	0.9	1.6	3.0	6.1	14.4
PAPP-A	AFP, uE_3, total hCG	79	87	91	0.4	0.6	1.1	2.1	4.4	10.8
PAPP-A	AFP, uE_3, free β-hCG	80	88	91	0.3	0.5	1.0	2.0	4.2	10.7
PAPP-A	AFP, uE_3, total hCG, inhibin-A	83	90	92	0.2	0.4	0.7	1.4	3.2	8.6
PAPP-A	AFP, uE_3, free β-hCG, inhibin-A	83	90	93	0.2	0.3	0.6	1.3	3.0	8.4

The tests are ordered according to screening performance at an 85% DR in groups categorised by number of markers used. Screening performance for NT measurement is based on sonographer-specific medians and technically satisfactory images. FPRs ≥ 15% are rounded to the nearest whole percentage.

Table 46 FPRs for specified DRs for selected first trimester and integrated tests according to the time of the first trimester serum sample

	DR (%)			
Test	75	80	85	90
10 completed weeks				
NT	12.9	18.0	25.1	36.6
Combined	2.3	3.7	6.1	10.8
Combined + inhibin-A	2.2	3.5	5.9	10.4
Serum integrated	0.8	1.5	2.7	5.3
Integrated	0.3	0.6	1.2	2.6
11 completed weeks				
NT	9.8	14.4	21.6	33.7
Combined	2.1	3.5	6.0	11.1
Combined + inhibin-A	1.7	2.9	5.0	9.4
Serum integrated	1.3	2.2	3.9	7.4
Integrated	0.3	0.6	1.2	2.8
12 completed weeks				
NT	8.6	13.0	20.0	32.4
Combined	1.9	3.4	6.0	11.2
Combined + inhibin-A	1.2	2.1	3.8	7.4
Serum integrated	1.8	2.9	4.9	8.9
Integrated	0.3	0.6	1.3	3.0
13 completed weeks				
NT	8.6	13.0	20.0	32.4
Combined	1.9	3.2	5.8	10.8
Combined + inhibin-A	0.8	1.4	2.5	5.0
Serum integrated	2.1	3.4	5.6	9.8
Integrated	0.4	0.7	1.5	3.3
10–13 completed weeks[a]				
NT	9.5	14.0	21.1	33.3
Combined	2.0	3.4	6.0	11.0
Combined + inhibin-A	1.4	2.3	4.1	7.8
Serum integrated	1.6	2.6	4.5	8.2
Integrated	0.3	0.6	1.3	3.0

[a] Weighted average of the individual estimates at each completed week, using the distribution of women in the SURUSS who booked at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%).

Table 47 Number of procedure-related unaffected fetal losses in 100,000 women screened for selected first trimester and integrated tests according to the time of the first trimester serum sample

	DR (%)			
Test	75	80	85	90
10 completed weeks				
NT	93	129	180	263
Combined	17	27	44	78
Combined + inhibin-A	16	25	42	75
Serum integrated	6	11	19	38
Integrated	2	4	9	19
11 completed weeks				
NT	70	103	155	242
Combined	15	25	43	80
Combined + inhibin-A	12	21	36	68
Serum integrated	9	16	28	53
Integrated	2	4	9	20
12 completed weeks				
NT	62	93	144	233
Combined	14	24	43	80
Combined + inhibin-A	9	15	27	53
Serum integrated	13	21	35	64
Integrated	2	4	9	22
13 completed weeks				
NT	62	93	144	233
Combined	14	23	42	78
Combined + inhibin-A	6	10	18	36
Serum integrated	15	24	40	70
Integrated	3	5	11	24
10–13 completed weeks[a]				
NT	68	100	152	240
Combined	15	24	43	79
Combined + inhibin-A	10	17	30	57
Serum integrated	11	19	32	59
Integrated	2	4	9	22

[a] Weighted average of the individual estimates at each completed week, using the distribution of women in the SURUSS who booked at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%).

Table 48 Cost (£ millions) of screening (including diagnosis and termination of pregnancy) 100,000 women for selected first trimester and integrated tests according to the time of the first trimester serum sample

	DR (%)			
Test	75	80	85	90
10 completed weeks				
NT	4.5	6.0	8.0	11.3
Combined	2.2	2.6	3.3	4.6
Combined + inhibin-A	2.3	2.7	3.4	4.7
Serum integrated	2.1	2.3	2.6	3.2
Integrated	2.4	2.5	2.6	3.0
11 completed weeks				
NT	3.7	5.0	7.0	10.4
Combined	2.2	2.6	3.3	4.7
Combined + inhibin-A	2.2	2.5	3.1	4.4
Serum integrated	2.2	2.4	2.9	3.7
Integrated	2.4	2.5	2.6	3.0
12 completed weeks				
NT	3.3	4.6	6.6	10.1
Combined	2.1	2.5	3.3	4.7
Combined + inhibin-A	2.0	2.3	2.8	3.8
Serum integrated	2.3	2.6	3.1	4.1
Integrated	2.4	2.5	2.7	3.1
13 completed weeks				
NT	3.3	4.6	6.6	10.1
Combined	2.1	2.5	3.2	4.6
Combined + inhibin-A	1.9	2.1	2.4	3.1
Serum integrated	2.4	2.7	3.3	4.3
Integrated	2.4	2.5	2.7	3.2
10–13 completed weeks[a]				
NT	3.6	4.9	6.9	10.4
Combined	2.1	2.5	3.3	4.7
Combined + inhibin-A	2.1	2.4	2.9	3.9
Serum integrated	2.3	2.5	3.0	3.9
Integrated	2.4	2.5	2.7	3.1

[a] Weighted average of the individual estimates at each completed week, using the distribution of women in the SURUSS who booked at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%).

Table 49 Cost (£1000s) of screening (including diagnosis and termination of pregnancy) per Down's syndrome pregnancy detected for selected first trimester and integrated tests according to the time of the first trimester serum sample

	DR (%)			
Test	75	80	85	90
10 completed weeks				
NT	29.8	36.8	46.3	61.5
Combined	14.5	16.0	19.0	25.2
Combined + inhibin-A	15.3	16.6	19.6	25.4
Serum integrated	13.7	13.9	14.8	17.5
Integrated	15.9	15.4	15.3	16.3
11 completed weeks				
NT	24.1	30.6	40.6	57.0
Combined	14.1	15.7	18.9	25.7
Combined + inhibin-A	14.3	15.6	18.1	23.9
Serum integrated	14.5	15.0	16.5	20.2
Integrated	15.9	15.4	15.3	16.6
12 completed weeks				
NT	21.9	28.2	38.0	55.0
Combined	13.7	15.5	18.9	25.8
Combined + inhibin-A	13.4	14.2	16.1	20.8
Serum integrated	15.3	16.0	17.9	22.2
Integrated	15.9	15.4	15.5	16.9
13 completed weeks				
NT	21.9	28.2	38.0	55.0
Combined	13.7	15.2	18.5	25.2
Combined + inhibin-A	12.7	13.0	14.0	17.1
Serum integrated	15.6	16.8	18.9	23.4
Integrated	16.0	15.5	15.7	17.3
10–13 completed weeks[a]				
NT	23.5	29.9	39.8	56.4
Combined	13.9	15.6	18.8	25.6
Combined + inhibin-A	13.7	14.6	16.6	21.4
Serum integrated	15.0	15.6	17.3	21.3
Integrated	15.9	15.4	15.5	16.8

[a] Weighted average of the individual estimates at each completed week, using the distribution of women in the SURUSS who booked at 10 completed weeks (12% including those at 9 completed weeks), 11 completed weeks (29%), 12 completed weeks (37%), and 13 completed weeks (22%).