NCCHTA

09 May 2007

1. Title: Cross trimester repeated measures testing for Down's Syndrome screening

2. Changes since the outline proposal was submitted

Our outline application went beyond the HTA commissioning brief by incorporating repeated measures within first trimester and including the assessment of the new marker Adam 12. As advised, we have removed these components and have restricted our bid to cross trimester repeated measures. Consequently, the cost has been reduced from £247,000 to £83,145 and the duration has been reduced from 18 to 10 months.

We have secured additional data from Dr Anne Summers, Chief and Medical Director, Genetics Program, North York General Hospital, Toronto, Canada. She is now included as a co-applicant.

At the time we made the outline application, Dr Ian Bradbury, co-author of the paper cited in the HTA call, was working in the pharmaceutical industry. He has recently taken up a post as Senior Lecturer in Statistics at Queen's University in Belfast and is now a co-applicant on this proposal.

3. Planned Investigation

3.1 Research objectives

1. The primary objective is to provide estimates of the performance (detection rates and false positive rates) of screening tests, by maternal age and for appropriately selected maternal age distributions. Results will be presented for tests involving repeated measures of combinations of pregnancy-associated plasma protein-A (PAPP-A), human chorionic gonadotrophin (hCG) and unconjugated oestriol (uE3). Contingent screening tests with repeated measures, including variants in which nuchal translucency (NT) is measured contingently on first trimester biochemical results, will also be considered. Statistical uncertainty will be quantified in terms of 95% confidence intervals.

Secondary objectives are

2. To review the available evidence on the parameters underlying the likelihood ratio calculation for repeated measures tests.

3. To synthesise available evidence on parameters for risk calculation in Down's syndrome screening.

Other outputs of the research will include estimates of the frequency of amniocentesis and chorionic villus sampling (CVS). The effects of other factors, including the uptake of screening, uptake of diagnostic tests, compliance, patient choice and differential fetal loss, will be assessed.

3.2 Existing Research

Repeated Measures

The choice of markers in multi-marker screening tests has been influenced by the extent to which they provide 'independent information' as characterised by low correlations between markers and the properties of markers when viewed individually. The prevailing view has been that combining makers with low correlations that individually have good discriminatory power represents the best approach to screening test design. Against this background, the integrated test (1) was obtained by combining the best markers from the first and second trimesters. The SURUSS report of Wald *et al*. (2) confirmed that the integrated test offers the most effective and safe current method of screening.

However, from the statistical perspective, the thinking behind the combination of markers in the integrated test is misguided. This was demonstrated in the paper of Wright and Bradbury (3) which showed that highly correlated repeated measures of markers, some of which, individually, have poor discriminatory power, have substantial benefits over the established combinations of markers used in the integrated test. Wald *et al*. (4) have carried out further work on repeated measures testing and reach the same general conclusions about its benefit over the integrated test. The principal difference between the two approaches adopted is that Wright and Bradbury (3) deal with individual pairs of multiples of the median (MoM_1 and MoM_2), whilst Wald *et al*. (4) deal with either MoM_1 or MoM_2 and the cross-trimester ratio ($CT=MoM_2/MoM_1$).

The two approaches are equivalent formulations of the same multivariate Gaussian likelihood ratio for log transformed MoM values. With the same underlying parameters they will produce identical risks (5). The differences between the results in Wright and Bradbury (3) and those in Wald *et al*. (4) arise due to substantial differences between the underlying parameter estimates in the CT ratios paper and previously published parameter estimates from SURUSS.

The existence of these differences illustrates the problem that the development of screening methods has to some extent relied on subjective decisions on which data points to exclude, with the consequence that results may be difficult to reproduce. An advantage of the methodology we propose to use is that it will produce results that are reproducible by any group of workers.

As Wright and Bradbury (3) emphasise, there is a need for further research because of uncertainty in parameter estimates, departures from model assumptions and inherent optimistic bias in the methods used to assess screening performance. A further source of uncertainty is that the assays used in the SURUSS study were not designed to work across the full range of marker values encountered in the first and second trimesters. This is a particular problem with PAPP-A in the second trimester. The National Screening Committee has recognised this need, prompting the HTA call addressed by this proposal.

Existing Validation Studies

We have collaborated in a validation study on repeated measures of PAPP-A (6). The parameters associated with second trimester PAPP-A were estimated using evidence from meta-analysis and from a consecutive series of 838 women using appropriately adjusted assays. An independent validation data set comprising 34 Down's syndrome pregnancies and 514 unaffected pregnancies was used. Using repeated measures of PAPP-A alone, 21 of 34 cases were detected (62%, 95% CI 44% to 78%) with 5% false positives. At an observed 2% false positive rate, the detection rates for the quadruple (69%) and serum integrated (69%) tests were lower than for the repeated measures test (75%). The data for this study were

obtained from the Canadian population where, in contrast to other countries including the UK and US, outcome data is routinely linked to data from screening. In this paper, we also introduced an alternative to the use of univariate truncation limits applied to each marker separately.

3.3 Research Methods

Data

In accordance with the commissioning brief this study will be based on an analysis of data from stored serum samples. These data arise from two sources; the FASTER repeated measures study and a case control study carried out on data from North York General Hospital in Toronto. In addition we shall include data on 1,000 routinely screened women attending North York General Hospital to provide additional evidence on repeated measures autocorrelations and cross-correlations. A summary of the data from these sources is presented in Table 1. We already have the repeated measures data from the FASTER study. We also have data on repeated measures of PAPP-A and the integrated test markers for a case study sample of 34 cases and 514 controls together with data on 1000 consecutively screened women attending North York General Hospital. These data were used in the validation study of Palomaki *et al.* (6). As part of this project further assays will be performed to obtain repeated measures data on hCG and uE3 to supplement the repeated measures data on PAPP-A. As more outcome data has become available since the original case control study was conducted, we expect to be able to augment this data set with a number of Down's syndrome pregnancies and appropriately selected controls. For both data sets we have information on demographic and pregnancy-related information, including maternal age, gestational age, maternal weight, smoking status and ethnicity. The additional assays will be carried out early in 2007 providing raw data on a total of over 112 Down's syndrome cases and over 900 controls. We shall also have data on the markers in the integrated test for these pregnancies.

	Faster repeats (nested case-control)		***North York repeats (case-control)***		***North York routine***	
	Unaffected	**Down's**	**Unaffected**	**Down's**	**Unaffected**	**Down's**
First trimester						
PAPP-A	390	78	508(+?)	34 (+?)	1000	0
hCG	390	78	(508+?)	(34+?)	(1000)	0
uE3	390	78	(508+?)	(34+?)	(1000)	0
NT	390	78	(508+?)	(34+?)	1000	0
Second Trimester						
PAPP-A	390	78	508(+?)	34 (+?)	1000	0
hCG	390	78	508(+?)	34 (+?)	1000	0
uE3	390	78	508(+?)	34 (+?)	1000	0

Table 1: Data available, number of observations, for cross trimester repeated measures analysis. Figures without brackets represent the number of samples already available for analysis. Bracketed figures show data we expect to have available by April 2007.

Statistical Methodology

The calculation of risks in Down's syndrome screening is an application of Bayes theorem for combining prior information on the maternal age specific risk with a likelihood ratio obtained from appropriately transformed measurements of analyte concentrations from maternal serum

and possibly ultrasound markers such as the Nuchal Translucency (NT). Almost invariably the transformation involves two steps. Firstly, the measurement is expressed as a multiple of the median (MoM) value for unaffected pregnancies standardising for gestational age and other variables such as maternal weight, smoking status and ethnicity which are known to have an effect on the measurement. Secondly, a log transformation is used to produce a log (MoM) value. The likelihood ratio is calculated under the assumption that log MoM values follow different multivariate Gaussian distributions in unaffected and in Down's syndrome pregnancies. In practice, the unknown parameters defining the multivariate Gaussian distributions are replaced by estimates obtained from fitting multivariate Gaussian models to data such as those collected in the SURUSS study. To deal with departures from the Gaussian form in the tails of the distribution, truncation is applied to values beyond a specified range. The established approach is to apply truncation separately to each dimension. We have adopted a multivariate approach, based on the Mahalanobis distance, truncating values that are atypical of either the Down's syndrome or unaffected distribution (6). This avoids the production of extreme risks for highly atypical pregnancies.

We remark that this standard approach to risk assessment takes no account of the uncertainty in the parameter estimates, but simply assumes these are fixed at the observed values. A more accurate reflection of this uncertainty could be obtained by a Bayesian *predictive* approach (7), essentially averaging over the posterior distribution of parameters. However, to ensure that our results will be directly comparable with previous work, we do not pursue this approach in this proposal.

The properties of screening tests derived from applying cut-off values to the risks calculated from the method above are determined by the population, as opposed to fitted, distributions of log MoM values. Ideally the fitted distributions should perfectly match the true population distributions and screening tests are usually assessed assuming this ideal. In practice, the fitted distributions will differ from the population distributions to some degree because the populations are not perfectly Gaussian and the fitted parameter estimates are subject to sampling error. This means that assessment of screening performance under ideal modelling assumptions is optimistically biased. In this proposal we shall deal with this by assessing the performance of fitted models for risk assessment using independent test data. This will provide realistic estimates of screening performance for existing models. However, it is inefficient in the sense that the independent test data is used solely for assessment and valuable information that could be used to improve the risk model is ignored. We shall address this limitation by fitting a new risk model using all the available evidence including the test data sets. The performance of this model will be assessed using modern computational statistical methods to reduce the bias that can result from using the same data for model fitting and testing.

Parameter estimates

This application focuses on assessing the performance of screening tests incorporating combinations of repeated measures of PAPP-A, hCG and uE3. To date, three sets of parameter estimates, I-III below, have been published that could be used as a basis for such screening tests. We propose to produce a fourth set of parameters estimates (IV) based on synthesis of existing evidence incorporating new data on repeated measures from the North York routine sample. The case control samples, which we shall use as an independent test set, will be excluded from this synthesis.

(I) the original SURUSS parameter estimates including all corrections (2,8,9).

(II) the CT ratios parameters estimates published by Wald *et al*. (4) incorporating any subsequent corrections.

(III) the SURUSS parameter estimates incorporating the modifications associated with measurements of PAPP-A in the second trimester Palomaki *et al.* (6).

(IV) we intend to take the model for means from published meta-analyses of established literature on means. Covariance matrices, or equivalently standard deviations and correlations will be obtained from synthesis of available evidence.

The synthesis for IV will be carried out using a Bayesian analysis implemented using WinBUGS. This approach will enable us to deal with missing data and, through the use a mixture model with contamination, obtain robust estimates of parameters without the need to make arbitrary or subjective decisions about exclusion of outliers (10). We remark that there is extensive published evidence on the mean log(MoM) values and their relationship with gestational age (11) that can be included in this synthesis.

We shall present summaries of the different parameter estimates I-IV in terms of the log MoM and log CT ratios formulation.

Estimation of screening performance using an independent test data

Independent validation studies of the risk algorithms arising from the four parameter sets I-IV will be carried out as follows. For the Canadian case control data set and the FASTER repeats data, we shall compute likelihood ratios using each of the parameter sets I-IV. The likelihood ratios will then be used to estimate the age specific detection rates and false positive rates for a range of cut-off values. The analysis of screening performance will be presented separately for the two test data sets in terms of point estimates and 95% confidence intervals for detection rates and false positive rates. Confidence intervals will be produced using non-parametric bootstrapping (12). The extent to which the individual risks and overall screening performance depend on the choice of parameter estimates will be illustrated graphically and discussed. The results will also be contrasted with those obtained from the Gaussian modelling described below. Graphical summaries, including Receiver Operating Characteristic (ROC) curves will be used to display the results.

Estimating screening performance under Gaussian assumptions

Given parameter estimates for means and covariances, the established modelling approach to estimating detection rates is to use Monte-Carlo methods and draw large samples (typically 500,000) from the joint distribution of log multiples of medians in Down's syndrome and in unaffected pregnancies. Likelihood ratios are then computed for each observation. Given a fixed risk cut-off, detection rates and false positive rates are computed for each maternal age. Alternatively, these rates can be obtained by numerical integration. Overall population detection rates for a reference maternal age distribution are obtained by combining maternal age specific rates according to the maternal age distribution for Down's syndrome. Similarly, false positive rates are obtained using the maternal age distribution of unaffected pregnancies. The risk cut-off above which pregnancies screen positive can be determined to fix the false positive rate or the detection rate. This enables different screening tests to be compared in terms of false positive rate or detection rate for a fixed detection rate or false positive rate respectively. In the SURUSS study, the maternal age distribution of England and Wales 1996-1998 was used as the reference maternal age distribution. We will report estimates of screening performance using the method described above, but, to reduce the bias due to reusing the same sample for building and testing the model, we shall also report estimates and 95% confidence intervals of screening performance obtained using the parametric bootstrap (13). We propose to use the parametric, rather than non-parametric, bootstrap because the raw data required for application of the non-parametric bootstrap are not available to us for the SURUSS study.

Evidence synthesis and development of a new screening algorithm

The assessment of screening performance using an independent test data set will provide evidence of the benefits of repeated measures and the effect of using different parameter estimates. We shall also compare the parameter estimates I – IV with those obtained using the case control data set. We shall then synthesise all available evidence including that from the case control data sets with the aim of producing the best available parameter estimates for use in repeated measures testing. The performance of this will be examined using the standard Gaussian modelling with the parametric bootstrap to remove bias and provide interval estimates. The results of this will be compared with the use of the non-parametric bootstrap (12) which accounts for departures from Gaussian assumptions in the population distributions.

Working methods

A full pre-specified analysis plan will be produced. All analyses will be carried out using documented and validated programs written in S-PLUS, R or WinBUGS enabling our results to be reproduced by other workers. The project will be overseen by a multi-disciplinary advisory group of experts in the field. The group will meet at two key points in the project; the first to finalise the analysis plan and publication policy and the second to finalise the report.

4. Project timetable and milestones

The project timetable below is based on an assumed start date of 1st January 2007. Project milestones are shown by the symbol ♦. The project plan assumes that we will know whether or not the HTA will fund this project by early in December 2006. We will then organise for additional assays on stored samples to commence in January 2007.

Month	**D**	**J**	**F**	**M**	**A**	**M**	**J**	**J**	**A**	**S**	**O**
Finalise arrangements for North York data collection.											
Briefing meeting with HTA		♦									
Additional North York assays											
Transfer of North York data					♦						
Preparation of analysis plan and report structure											
Advisory group meeting to discuss and finalise analysis plan				♦							
Finalised analysis plan and publication policy submitted to FASTER publications committee and other agencies					♦						
Programming, data analysis, production of tables and graphs for main report											
Draft report and papers											
Advisory group meeting to discuss and agree final papers										♦	
Final draft of HTA reports and main papers											♦
Months from project start		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

5. Expertise

Bradbury, Nix and Wright are experienced statisticians with strong methodological backgrounds and keen interests in applications in the areas of risk assessment, screening and diagnosis. They have a strong history of collaboration in these areas; with each other and with medical workers in the field. They will work together on development of the methodological aspects of the analyses plan, programming, data analysis and reporting to ensure that the statistical aspects of the work are of the highest quality. Bradbury and Wright both have experience in working on Health Technology Assessments for the NCCHTA and for NHS Scotland. Nix and Wright work as Statisticians to the National Screening Program through DQASS (http://www.screening.nhs.uk/downs/dqass.htm). Aitken, Crossley, Cuckle and Spencer have played leading roles in the development of Down's syndrome screening in the UK and beyond. Their primary role is to serve on the advisory panel and ensure that the project is conducted in a way that has practical relevance and that the reports and publications arising from the project are accessible to practitioners in the field. Malone and Summers have made the project possible by proving high quality data from the FASTER trial and from the Canadian screening program. They both have outstanding international reputations in the area and their contributions through the advisory group will be invaluable.

6. Service users

There are currently no plans to include service users in this project.

7. Justification for support required

Statistical analysis and reporting
The major component of the project is the statistical analysis. The team of three senior statisticians involved will contribute in total 6 months of a full time equivalent member of staff. This will enable us to complete the project within the 10 months time scale with a high level of rigour. In particular, a full pre-specified analysis plan will be produced, and all analyses will be carried out using well validated and documented programs permitting independent audit of the work.

Laboratory costs
We are seeking support for further assays on stored samples at North York hospital in Toronto, Canada. This will provide data on repeated measures of hCG and PAPP-A to supplement the data we already have on 34 Down's syndrome pregnancies as well as stored data on routinely screened pregnancies for estimations of cross correlations and autocorrelations. The laboratory staff and materials costs for these assays total £15,506.

Advisory group
The project will be overseen by an advisory group of recognised experts in the field bringing expertise in fetal and maternal medicine, reproductive epidemiology, clinical biochemistry, genetics and statistics.

8. References

1. Wald N, Watt HC, and 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.

2. Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, and 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 (Winch) 2003;7:1-88.

3. Wright DE and Bradbury I. Repeated measures screening for Down's Syndrome. BJOG 2005;112:80-3.

4. Wald NJ, Bestwick JP, Morris JK. a. Cross trimester marker ratios in prenatal screening for Down's Syndrome. Prenat Diag 2006a;26:514-23.

5. Wright D, Bradbury I, Benn P, Nix B, Spencer K and Cuckle H. CT Ratios: Parameter estimates are inconsistent with SURUSS publications? 2006; In press: Included in Appendix I.

6. Palomaki GE, Wright DE, Summers AM, Neveux, LM, Meier C, O'Donnell A, Huang T, Knight GJ, and Haddow J Repeated Measurement of Pregnancy-Associated Plasma Protein-A (PAPP-A) in Down Syndrome Screening: a Validation Study. Prenat Diagn 2006; DOI: 10.1002 pd.1497.

7. Aitchison J and Dunsmore IR. Statistical Prediction Analysis. Cambridge University Press; 1980.

8. Wald NJ, Rodeck C, Rudnicka A and Hackshaw A. Nuchal Translucency and gestational age. Prenat Diagn 2004;24:150-3.

9. Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM, Bestwick JP. Correction to the SURUSS Report. J Med Screen 2006b;13:51-2.

10. Berger JO, Ríos Insua D, and Ruggeri F. Bayesian robustness. In: Ríos Insua D and Ruggeri F, editors. Robust Bayesian Analysis. New York: SpringerVerlag; 2000. p1–32.

11. Spencer K, Crossley JA, Aitken DA, Nix ABJ, Dunstan FDJ and Williams K. Temporal changes in maternal serum biochemical markers of trisomy 21 across the first and second trimester of pregnancy. Ann Clin Biochem 2002;39:567-76.

12. Efron B, Tibshirani RJ. Improvements on Cross-Validation: The Bootstrap Method. JASA 1997;92:548-60.

13. Davison AC, Hinkley DV and Young GA. Recent Developments in Bootstrap Methodology. Statistical Science 2003;18:141-57.

Appendix I: Response to Wald *et al.* (2006)

CT Ratios: Parameter estimates are inconsistent with SURUSS publications?

Wright and Bradbury (2005) used mathematical modelling to demonstrate the potential benefits of repeat measurement of the same Down's syndrome screening markers in first and second trimester serum samples. Wald *et al*. (2006a) have now carried out their own modelling exercise and reach the same general conclusions. Whilst both groups used Gaussian modelling with distribution parameters estimated from the SURUSS study, their approaches differed. Wright and Bradbury (2005) modelled the individual pairs of multiples of the median (MoM_1 and MoM_2); Wald *et al*. (2006a) modelled either MoM_1 or MoM_2 and the cross-trimester ratio ($CT=MoM_2/MoM_1$). Wright and Bradbury use previously published parameter estimates from SURUSS; Wald et al. (2006a) also use published SURUSS parameter estimates together with newly presented parameter estimates for CT ratios obtained from the SURUSS data.

The two approaches are equivalent formulations of the same multivariate Gaussian likelihood ratio and with the same underlying parameters they will produce identical risks. This equivalence does not depend, as Wald *et al*. (2006a) suggest, on the assumption that the data are truly Gaussian. Moreover, the means, standard deviations and correlation coefficients for the MoM_1 and MoM_2 configuration can be obtained from the means, standard deviations and correlations of the CT ratios configuration and vice versa. The relationships that enable this transformation of parameters between configurations do not depend on Gaussian assumptions (see for example, Mardia *et al*., 1979).

Nevertheless, there are discrepancies between the two configurations both in terms of estimated detection rates and false positive rates and individual risks. These differences, including the extreme risks and the negative determinant highlighted by Wald *et al*. (2006a), must therefore depend on differences between the estimated parameters underlying the two approaches. We explored these differences by computing parameter estimates in the MoM_1 and MoM_2 configuration from those of CT ratios. We found that the parameter estimates in the CT ratio paper are inconsistent with previously published parameters from SURUSS and established evidence.

Firstly, the mean log MoM for the first trimester marker can be obtained by subtracting the mean log CT ratio from the mean of the second trimester log MoM. Following Wald *et al*.(2006a) and taking data for the CT ratio for AFP in Down's syndrome pregnancies from Table A1 and second trimester mean AFP from Table 33 of the SURUSS report, we obtained mean log MoM AFP values of -0.2380, -0.1839, -0.1264 and -0.0703 at 10, 11, 12 and 13 weeks gestation respectively. The corresponding median MoM values are 0.58, 0.66, 0.75 and 0.85. This is inconsistent with the SURUSS study which presents a single first trimester median MoM of 0.86. Even after accounting for higher first trimester standard deviations, this implies that AFP is a better marker at 10 weeks than it is in the second trimester. This is inconsistent with the large body of published evidence on AFP.

Secondly, the mean of log(MoM_2) can be obtained by adding the mean log(CT) ratio to the mean of log(MoM_1). Taking values for mean log(MoM_1) and mean log (CT ratio) from SURUSS and Wald *et al.* (2006a) for PAPP-A, gives second trimester means for log MoM values of -0.074, -0.0344, -0.0087 and 0.0039 for first trimester sampling times of 10, 11, 12 and 13 weeks. The corresponding second trimester median MoM values are 0.84, 0.92, 0.98, and 1.01. It is, of course, illogical to find that the median MoM in the second trimester should depend on when the first trimester sample was taken.

Lastly, standard deviations and correlation coefficients for the MoM_1 and MoM_2 configuration can be derived from the CT ratios (see, for example, Mardia *et al*., 1979). Applying this to PAPP-A in Down's syndrome, the correlation between first and second trimester log MoM values derived from the CT configuration is 0.57 compared to 0.83 from SURUSS (Wald *et al.*, 2006b). Similar discrepancies can be found for other markers. For example, the correlation between first and second trimester log MoM values for UE3 in Down's syndrome is 0.54 from CT ratios compared to 0.74 in SURUSS. In the opinion of Wald *et al*. (2006a) deviations from Gaussian fit and high cross-trimester correlations imply that the parameters for CT can be more reliably estimated than for MoM_1 and MoM_2. There is no statistical basis for this view.

These discrepancies underline the need for caution in the interpretation of modelled screening performance and raise questions concerning the conclusions of Wald et al. (2006a). The potential value of repeated measures has been demonstrated *in principle* but further data and improved parameter estimates are needed. With the screening policy proposed by Wald et al. (2006a), 45 correlation coefficients, 10 means and 10 standard deviations are required in Down's syndrome pregnancies. The sample of just 74 cases from SURUSS are insufficient to estimate these reliably.

Dave Wright[1], Ian Bradbury[2], Peter Benn[3], Barry Nix[4], Kevin Spencer[5] and Howard Cuckle[6]

[1]School of Mathematics and Statistics, University of Plymouth, UK

[2]Department of Public Health Medicine and Primary Care, Queen's University, Belfast, UK

[3]Department of Genetics and Developmental Biology, University of Connecticut, USA.

[4]Department of Epidemiology, Statistics and Public Health, University of Cardiff, UK.

.

[5]Prenatal Screening Unit, Harold Wood Hospital, Romford, UK
[6]Reproductive Epidemiology, University of Leeds, UK

References

Mardia KV, Kent JT, Bibby JM. 1979. *Multivariate Analysis*. Academic Press, London.

Wald NJ, Bestwick JP, Morris JK. 2006a. Cross trimester marker ratios in prenatal screening for Down's Syndrome. *Prenat Diag* **26**: 514-523.

Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM, Bestwick JP. 2006b. Correction to the SURUSS Report. *J Med Screen* **13**: 51-52.

Wright DE, Bradbury I. 2005. Repeated measures screening for Down's Syndrome. *Br J Obstet Gynaecol* **112**: 80-83.