Contingent screening for down syndrome is an efficient alternative to non-disclosure sequential screening

Wright, D., Bradbury, I., Benn, P., Cuckle, H., & Ritchie, K. (2004). Contingent screening for down syndrome is an efficient alternative to non-disclosure sequential screening. Prenatal Diagnosis, 24(10), 762-766. DOI: 10.1002/pd.974

Published in:
Prenatal Diagnosis

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Contingent screening for Down syndrome is an efficient alternative to non-disclosure sequential screening

Dave Wright1*, Ian Bradbury2, Peter Benn3, Howard Cuckle4 and Karen Ritchie5

1Department of Mathematics and Statistics, University of Plymouth, UK
2School of Biomedical Sciences, University of Ulster, Coleraine, UK
3Department of Genetics and Developmental Biology, University of Connecticut, Health Center, Farmington, CT, USA
4Reproductive Epidemiology, School of Medicine, University of Leeds, UK
5NHS Quality Improvement, Scotland, UK

Objective To present a first and second trimester Down syndrome screening strategy, whereby second-trimester marker determination is contingent on the first-trimester results. Unlike non-disclosure sequential screening (‘the Integrated test’), which requires all women to have markers in both trimesters, this allows a large proportion of the women to complete screening in the first trimester.

Methods Two first-trimester risk cut-offs defined three types of results: positive and referred for early diagnosis; negative with screening complete; and intermediate, needing second-trimester markers. Multivariate Gaussian modelling with Monte Carlo simulation was used to estimate the false-positive rate for a fixed 85% detection rate. The false-positive rate was evaluated for various early detection rates and early test completion rates. Model parameters were taken from the SURUSS trial.

Results Completion of screening in the first trimester for 75% of women resulted in a 30% early detection rate and a 55% second trimester detected rate (net 85%) with a false-positive rate only 0.1% above that achievable by the Integrated test. The screen-positive rate was 0.1% in the first trimester and 4.7% for those continuing to be tested in the second trimester. If the early detection rate were to be increased to 45% or the early completion rate were to be increased to 80%, there would be a further 0.1% increase in the false-positive rate.

Conclusion Contingent screening can achieve results comparable with the Integrated test but with earlier completion of screening for most women. Both strategies need to be evaluated in large-scale prospective studies particularly in relation to psychological impact and practicability. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: Down syndrome; prenatal screening; sequential screening; contingent; integrated

INTRODUCTION

Second-trimester screening for Down syndrome using maternal serum α-fetoprotein (AFP), human chorionic gonadotrophin (hCG), or free β-hCG, with or without unconjugated estriol (uE3) and inhibin has become established practice. The discovery that ultrasound nuchal translucency (NT) measurement is a powerful Down syndrome marker has led to the development of first-trimester screening using NT and concurrent maternal serum pregnancy-associated placent protein-A (PAPP-A) and free β-hCG. This has the advantages over second-trimester screening of earlier diagnosis and reassurance, and also a detection rate that is higher for a given false-positive rate. In the pooled results of the six prospective intervention studies published to date, the detection rate (DR) is 89% and false-positive rate (FPR) is 6.6% (Krantz et al., 2000; Schuchter et al., 2002; von Kaisenberg et al., 2002; Bindra et al., 2002; Spencer et al., 2003; Wapner et al., 2003).

A logical extension is to attempt to obtain even better results by a combination of first- and second-trimester markers. Wald et al. (1999) have suggested using the best markers in each trimester, NT and PAPP-A in the first together with AFP, free β-hCG, uE3 and inhibin in the second, to derive a single risk figure for counselling (‘the Integrated test’). Model predictions based on parameters from the Serum, Urine and Ultrasound Screening Study (SURUSS) suggest that this can achieve a DR of 85% for an FPR as low as 1.2% (Wald et al., 2003). However, this level of performance requires that the first-trimester results be withheld from women until the second-trimester markers have been measured.

This type of strategy, referred to as non-disclosure sequential screening (Cuckle, 2002), remains controversial. Herman et al. (2002) have suggested that clinical staff may find non-disclosure of high-risk first-trimester results unacceptable, believing that the interests of individual patients may be better served by immediate counselling for invasive prenatal diagnosis. Similarly, Copel and Bahado-Singh (1999) argue that patients may put considerable pressure on clinicians to disclose intermediate findings since for social and personal reasons they would prefer to terminate an abnormal fetus as early as possible. These concerns have motivated some to
propose a policy of routine disclosure (Herman et al., 2002) despite the concomitant considerable increase in false-positive rate for any given detection rate (Wald et al., 2003; Cuckle and Arbuzova, 2004).

One possible compromise solution is a contingent strategy similar to that used in the design of group sequential clinical trials (Pocock, 1982). Such an approach has been considered in relation to first-trimester screening, whereby only women with borderline NT results are offered concurrent serum marker determination (Cuckle and Arbuzova, 2002, 2004) or vice versa (Christiansen and Larsen, 2002). Recently, Maymon et al. (2004) used the same strategy in sequential first and second-trimester screening. They chose a high first-trimester risk cut-off above which a complete Integrated test is unlikely to become negative to define a group for early karyotyping, and used logistic regression to define a very low risk group who need no further testing. The remainder would have their risk reassessed after second trimester—marker determination.

In this article, we develop contingent sequential screening further by using statistical modelling to define optimal first trimester upper and lower risk cut-offs. We demonstrate that only a small proportion of women need to have second-trimester markers measured, with little loss of performance compared with the Integrated test.

METHODS

Policy

Contingent screening in general is defined in terms of three risk cut-offs (Figure 1). In the first trimester, those women at very high risk (above 1 in c1) would be offered immediate invasive prenatal diagnosis. Those women with very low risks (below 1 in c2) would be provided with their risk estimate and require no additional testing. Those with intermediate risks (between 1 in c1 and 1 in c2) would receive second trimester marker testing. Their risk would be re-evaluated after incorporating all first and second-trimester markers and compared with a third risk cut-off (1 in d). This formulation includes, as special cases, all women receiving only concurrent first-trimester screening (c1 = c2) or all women receiving the Integrated test (c1 = 1, c2 = ∞). Other special cases are when no pregnancies in the first trimester are screen negative (c2 = ∞) or screen positive (c1 = 1).

In the primary analysis reported below, contingent screening was applied to the components of the Integrated test as specified in the SURUSS: NT and PAPP-A at 10 weeks’ gestation; AFP, free ß-hCG, ßE3 and inhibin at 14 to 20 weeks. In subsidiary analyses, free ß-hCG is moved to 10 weeks and in addition hCG is included at 14 to 20 weeks. All risks were expressed in terms of the live-birth incidence of Down syndrome.

Model

All markers were expressed as multiples of the gestation specific median (MoMs) for unaffected pregnancies and log10 transformed. A multivariate Gaussian model was used to describe the distributions of marker profiles in Down syndrome and in unaffected pregnancies (Royston and Thompson, 1992). Estimates of the means, standard deviations and correlation coefficients in affected and unaffected pregnancies were obtained from the SURUSS (Wald et al., 2003). The study included NT measurements on 85 Downs syndrome and 39 898 unaffected pregnancies with serum marker levels for 98 cases and 490 controls.

The maternal age distribution was taken to be that of England and Wales for the period 1996 to 1998 (Office for National Statistics, 1997–1999). The maternal age related risk of Down syndrome was obtained from the model of Wright and Bray (2000).

Rates

Detection and false-positive rates were estimated using Monte Carlo methods to sample from the modelled distributions. Samples of 500 000 observations were drawn; likelihood ratios were computed for each observation and used to derive maternal age specific rates, which were then combined. The precision of the various measures reported was assessed both theoretically and empirically and 500 000 pregnancies in each group give sufficient precision. However, variability due to the model parameter estimation remains. Documentation and copies of the S-Plus functions used in the estimation of detection and false-positive rates are available on request from the corresponding author.

Following the SURUSS (Wald et al., 2003), a reference detection rate of 85% was adopted to facilitate comparisons. c1 was varied to derive early detection rates of 30, 35, 40, 45 and 50%; for each value of c1, c2 was varied to derive early screening completion rates of 65, 70, 75, 80 and 85%. For all the combinations of c1 and c2, d was then determined, and the overall false-positive rate was estimated and compared with the 1.2% estimated for the Integrated test.
RESULTS

Table 1 shows the increase in false-positive rate over that achieved with the Integrated test, according to the contingent screening protocol. For any given early detection rate, the false-positive rate is not greatly affected by increases in the early completion rate under 80% but rises rapidly thereafter (see also Figure 2). For example, it would require a 0.5 to 0.6% increase in the false-positive rate to obtain an early completion rate of 85% for all the tabulated early detection rates. Figure 2 also shows that at all early completion rates the false-positive rate rises rapidly for early detection rates over 50%.

One attractive contingent screening policy was associated with a 0.1% increase in the false-positive rate and yielded a 30% early detection rate and 75% early completion rate. Either the early detection rate could be increased to 45% or the early completion rate could be increased to 80% for a false-positive rate that was 0.2% more than the Integrated test. To achieve both would require an increased false-positive rate of 0.3%. Any further improvement in the early detection rate or early completion rate would require an increased false-positive rate of 0.4 or greater.

Table 2 shows the values of c₁, c₂ and d together with the predictive values corresponding to three contingent screening policies with a 0.1 or 0.2% false-positive increase. In each case, only a small proportion of women would need to be referred for first trimester invasive prenatal diagnosis and they would form an extremely high risk group, about 25 to 65 times higher than the group risk for those referred in the second trimester. The group risk in those needing second-trimester screening is also high, although the range of risks is very wide, and the false-positive rate is 3 to 5 times higher than that for the whole screened population.

The values of c₁ and c₂ are governed by the choice of early detection and completion rates respectively. In the current analysis, to facilitate comparison with the Integrated test, we have fixed the overall detection rate at 85%. This means that for any given values of c₁ and c₂ the value of d will be completely determined. For example, with a 30% early detection rate, increasing the early completion rate from 75 to 80% means lowering the second trimester risk cut-off from 1 in 126 to 1 in 155 in order to maintain the 85% overall detection rate (see Table 2). The net effect of these changes is to increase the false positive by an estimated 0.1%. Had the cut-off remained at 1 in 126 the overall detection rate would have decreased to 84%.

Contingent screening could be improved further by adding free β-hCG to PAPP-A in the first trimester and replacing it in the second trimester with intact hCG. This will allow additional early detection or early completion without a marked increase in the false-positive rate. Thus, early detection and completion rates of 50 and 75% or 30 and 85% respectively could be obtained by a false-positive rate that was just 0.3% greater than the standard Integrated test. Alternatively, the same early detection and completion rates as in Table 2 could be obtained with a false-positive rate below that of the Integrated test.

DISCUSSION

We have demonstrated that, at the cost of a small reduction in overall performance, an appropriate contingent screening policy can permit clinicians to disclose the first-trimester results of a sequential screening programme. Women who are at an extremely high risk of
Table 2 — Cut-off risks (at term) and performance characteristics for three contingent screening policies with a 0.1 or 0.2% false-positive increase

<table>
<thead>
<tr>
<th>First trimestera</th>
<th>Screen Positive</th>
<th>30 &amp; 75%</th>
<th>45 &amp; 75%</th>
<th>30 &amp; 80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper cut-off (1 in $c_1$)</td>
<td>1 in 8.9</td>
<td>1 in 23.4</td>
<td>1 in 8.9</td>
<td></td>
</tr>
<tr>
<td>Positive rate (%)</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Group risk</td>
<td>1 in 2.4</td>
<td>1 in 4.9</td>
<td>1 in 2.4</td>
<td></td>
</tr>
<tr>
<td>Screen Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower cut-off (1 in $c_2$)</td>
<td>1 in 1980</td>
<td>1 in 1980</td>
<td>1 in 1410</td>
<td></td>
</tr>
<tr>
<td>Group risk</td>
<td>1 in 7500</td>
<td>1 in 7500</td>
<td>1 in 6100</td>
<td></td>
</tr>
<tr>
<td>Continue testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group risk</td>
<td>1 in 210</td>
<td>1 in 270</td>
<td>1 in 170</td>
<td></td>
</tr>
<tr>
<td>Second trimesterb</td>
<td>Screen Positive</td>
<td>30 &amp; 75%</td>
<td>45 &amp; 75%</td>
<td>30 &amp; 80%</td>
</tr>
<tr>
<td>Cut-off (1 in $d$)</td>
<td>1 in 126</td>
<td>1 in 124</td>
<td>1 in 155</td>
<td></td>
</tr>
<tr>
<td>Positive ratec (%)</td>
<td>4.7</td>
<td>4.2</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>Group risk</td>
<td>1 in 13</td>
<td>1 in 16</td>
<td>1 in 15</td>
<td></td>
</tr>
<tr>
<td>Screen Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group risk</td>
<td>1 in 1300</td>
<td>1 in 1300</td>
<td>1 in 1300</td>
<td></td>
</tr>
</tbody>
</table>

a NT and PAPP-A.
b $\alpha$-fetoprotein, free $\beta$-hCG, uE3 and inhibin.
c Among those with a first trimester 'Continue testing' result; among all women the rates are 1.3, 1.3 and 1.4% respectively.

carrying an affected fetus can have an early diagnosis, with up to half of the Down syndrome pregnancies detected by the whole programme, identified early. Furthermore, second-trimester screening can be avoided in up to 80% of the unaffected pregnancies.

The marker parameters used in this analysis were from the SURUSS, which was an intervention study and so the Down syndrome means are subject to viability bias. This arises because a proportion of those having termination of pregnancy because of extreme marker levels would have been destined to miscarry anyway, whereas non-viable cases with normal levels will not be known to the investigators. In the SURUSS, intervention was generally delayed until the second trimester so that bias, although present, would have been reduced somewhat. The SURUSS investigators suggest that a degree of bias is acceptable by adopting second trimester risk cut-offs rather than term risks as we have done, and is the norm in many countries.

When the analysis was repeated using an unbiased set of parameters (Cuckle and Arbuzova, 2004), based on a meta-analysis including several hundred Down syndrome cases, the findings were essentially unaltered. The only other difference is that unlike the SURUSS, these parameters assume that PAPP-A is measured at 10 weeks but NT is determined at 11 to 13 weeks when it is more reliable. As more data on between-trimester correlations emerges, the meta-analysis can be updated. However, this is only likely to change the specific first- and second-trimester cut-off risks, not the general conclusions.

To allow comparison with the SURUSS, we chose to fix the detection rate at 85% and examined the effect on the false-positive rate. For a wide range of early detection and completion rates, the difference between the Integrated test and contingent screening was marginal. Fixing the false-positive rate and allowing the detection rate to vary would have led to the same conclusion. For example, for a 0.6% false-positive rate, the Integrated test detection rate is 80%, whilst contingent screening with early detection and completion rates of 30 and 80% has a detection rate of 79%.

The discriminatory power of PAPP-A is greatest at 10 weeks’ gestation, which is why the Integrated test is designed to begin the sequential process at this gestational age. However, in practice, errors in gestational dating will dictate that for many women PAPP-A is tested at 11 to 13 weeks when it is less useful. In contrast to PAPP-A, the discriminatory power of free $\beta$-hCG improves as the first-trimester advances, and when both markers are measured, the two effects are almost counterbalanced (Cuckle and Arbuzova, 2004). Since this combination is less sensitive to gestational age error, it might be the preferred option. Our model predictions are that contingent screening with NT, serum PAPP-A and free $\beta$-hCG at 10 weeks, and AFP, hCG, uE3 and inhibin at 14 to 20 weeks, the early detection rate and the early completion rate can be further increased with few extra false-positives.

Contingent screening offers considerable psychological and clinical advantages over the Integrated test of early diagnosis, and if necessary, termination of pregnancy, and early reassurance for large numbers of women. It also overcomes the difficulties for health professionals, clearly expressed in the literature, imposed by non-disclosure over a long waiting period. On the other hand, as many as 25% of women will have a
similarly long wait before their tentative first-trimester risk assessment can be resolved by the further markers. Another problem, though probably less serious, is explaining to women why different risk cut-offs are being applied to define high risk in the two trimesters. It remains to be seen what the actual preferences of women are.

Statistical modelling has proven to be a useful tool in the evaluation of competing Down syndrome screening strategies. But modelling rests on assumptions that do not always apply in practice. Even if these assumptions hold, it tends to overestimate efficacy (Aitchison and Dunsmore, 1975). Both the current analysis and the Integrated test modelling in the SURUSS (Wald et al., 2003) are reliant on the assumption that a Gaussian fit is reasonable in the tails of a multivariate distribution. These efficient methods require the Gaussian model to hold further into the extreme tails of the multivariate distribution than less efficient strategies.

The practicality of sequential screening policies has yet to be established. One issue is the extent to which non-compliance to the protocol will lead to additional early invasive testing. This arises for the Integrated test when the NT is extremely large and for contingent screening when the first-trimester risk is not high enough to warrant intervention but higher than the cut-off for a non-sequential test. In order to examine the relative psychological impact, statistical validity and practicality of both contingent screening and the Integrated test, it will be important to carry out prospective intervention studies.

REFERENCES


