Serial first- and second-trimester Down’s syndrome screening tests among IVF- versus naturally-conceived singletons

R. Maymon\textsuperscript{1,3} and A. Shulman\textsuperscript{2}

\textsuperscript{1}Departments of Obstetrics and Gynecology, Assaf Harofeh Medical Center, Zerifin and \textsuperscript{2}Sheba Medical Center, Tel Hashomer, both affiliated with the Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

\textsuperscript{3}To whom correspondence should be addressed. E-mail: intposgr@post.tau.ac.il

BACKGROUND: It has been reported that second-trimester serum markers may be affected by assisted reproduction, leading to a higher false-positive rate. METHODS: A total of 285 naturally and 71 IVF-conceived singletons which underwent a serial disclosure Down’s syndrome screening programme were compared. The study protocol included first-trimester combined [nuchal translucency (NT), free $\beta$-HCG and pregnancy-associated plasma protein-A (PAPP-A)] testing. The second-trimester triple serum screening included $\alpha$-fetoprotein (AFP), intact HCG and unconjugated estriol (uE$\textsubscript{3}$). After excluding aneuploidies, miscarriages, anatomical anomalies and cases with incomplete follow-up, the serum samples of normal cases were assessed and correlated. RESULTS: NT measurement was not significantly changed in either group. However, the IVF group had lower PAPP-A [0.96 versus 1.05 multiples of normal median (MoM)] and higher AFP (1.13 versus 1.07 median MoM). Both groups had similar rates of first-trimester false-positive results (FPR; 7 and 9% respectively), but the IVF group had a significantly higher mid-gestation FPR rate (10 versus 5%; Pearson $\chi^{2}$, $P = 0.029$). This has contributed to amniocentesis uptake rates of 15 and 13% for the IVF and natural conception pregnancies respectively. CONCLUSIONS: The IVF group tended to have a significantly higher second-trimester FPR rate. To counterbalance this phenomenon, integrated first- and second-trimester screening tests or the use of NT alone might be a reasonable option that deserves further investigation.

Key words: assisted reproduction/combined test/Down’s syndrome/triple test

Introduction

Routine multi-marker screening for Down’s syndrome (DS) is now an established practice in many countries. In the second trimester, the most common markers are maternal serum HCG or its free $\beta$-subunit (F$\beta$HCG), $\alpha$-fetoprotein (AFP) and unconjugated estriol (uE$\textsubscript{3}$). Large studies using combinations of HCG or F$\beta$HCG and either or both of the other markers have confirmed model predictions that about two-thirds of DS-affected pregnancies can be detected with a false-positive result (FPR; 7 and 9% respectively), but the IVF group had a significantly higher mid-gestation FPR rate (10 versus 5%; Pearson $\chi^{2}$, $P = 0.029$). This has contributed to amniocentesis uptake rates of 15 and 13% for the IVF and natural conception pregnancies respectively. CONCLUSIONS: The IVF group tended to have a significantly higher second-trimester FPR rate. To counterbalance this phenomenon, integrated first- and second-trimester screening tests or the use of NT alone might be a reasonable option that deserves further investigation.

Key words: assisted reproduction/combined test/Down’s syndrome/triple test
(Liao et al., 2001) found decreased PAPP-A and increased FβHCG among their IVF cases. In contrast, the other group (Wojdeman et al., 2001) found those serum markers not to be significantly changed.

The aim of the current study was to assess among a preselected group of unaffected IVF pregnancies the marker profile which constitutes the combined test (Wald and Hackshaw, 1997), followed by the second-trimester triple test (TT) and the FPR rate of each test. These results were compared with data from a group of women who conceived naturally, and underwent the same investigation.

Materials and methods

This longitudinal study was carried out between January 1999 and September 2000 in a private-practice facility. The pregnant women attending were invited to participate in a sequential disclosure DS screening approach. Those who fulfilled the entry criteria and gave their consent comprised the study population. This assessment was carried out among women of a similar age (±3 years). Exclusion criteria included more than one gestational sac seen at the first-trimester scan, cases with chromosomal aneuploidy, early to mid-gestation pregnancy loss (≤24 completed weeks), anatomical abnormalities and those cases for which essential information was not available. After they were given a leaflet which explained the nature and implications of this test, the parturient women underwent NT screening. Following the scan, blood for monitoring of the first-trimester serum markers (FβHCG and PAPP-A) was withdrawn, and the women were offered mid-gestation TT for assessing AFP, total HCG and uE3. The women were informed of the results for each test upon completion of each screening test (i.e. combined first trimester and TT). Those women who had one positive screening test result were advised to have mid-gestation amniocentesis for fetal karyotyping.

The results of the screening tests are presented in Israel as risk for DS live birth, and the threshold for recommending fetal karyotyping in both tests is ≥1:380. This is the equivalent risk of a 35-year-old gravida at term pregnancy, and is the common practice in the Israeli national healthcare system (Legum et al., 1994).

**IVF protocol**

The procedure for the assisted conception in this programme was carried out in accordance with a previously reported protocol (Shulman et al., 1999; Maymon and Shulman, 2001).

**Screening tests**

Only singleton fetuses with a crown–rump length (CRL) of 38–84 mm were included. This corresponds with a gestational age of between 10 weeks and 3 days, and 13 weeks and 7 days. Measurement of CRL to estimate gestational age was performed concomitantly with NT assessment, and was carried out according to a previously described method (Pandya et al., 1995). The scans were performed by experienced examiners (R.M. and A.S.). First-trimester PAPP-A and FβHCG assays were measured as previously reported (Maymon et al., 2001). Calculation of first-trimester DS risk was based on the multivariate Gaussian distribution (Wald and Hackshaw, 1997), using the ‘alpha’ software (Logical Medical Systems Ltd, London, UK). Second-trimester TT were assessed as reported previously (Maymon and Shulman, 2001). Second-trimester DS risk was derived from the combination of triple serum markers and maternal age, and was calculated using commercial software. The serum samples in both sessions were tested in a routine analytical run together with regular maternal serum samples, all at the same antenatal DS screening programme in Zer Medical Laboratories (ISO 9002 UK, certified and authorized by the Ministry of Health, Israel). Testing was carried out in a fashion blinded to group classification; that is, samples from the study cases and those of other pregnant women were assessed simultaneously in the same laboratory. The measured marker levels were expressed as multiples of the gestation-specific normal medians (MoM). Median values for each serum analyte were calculated against completed menstrual weeks and adjusted for maternal weight. Reference MoM values were calculated from the local population as established in Zer Medical Laboratories during the same time period. Gestational age in the IVF cases was calculated from the date of oocyte retrieval minus 2 weeks (or embryo transfer minus 16 days) to convert the menstrual dating. The gestational age in all cases was confirmed by first-trimester scan, as this was performed simultaneously with the NT measurement. Data on the sequential DS screening tests results and pregnancy outcomes were recorded prospectively on a database.

**Statistical analysis**

Standardized kurtosis showed that the data were derived from a normal distribution, and were expressed as mean (± SD). Frequencies were expressed as percentages. The DS screening test was estimated based both on NT values and five serum markers (data were transformed using natural logarithms) and maternal age. The data all fitted a Gaussian distribution after log10 transformation. After excluding all women who did not fulfil the study criteria, the remaining parturient women with normal pregnancy outcome were further subdivided according to first- versus second-trimester screening results. Differences between various variables of pregnancy follow-up were compared using the Pearson χ2-test. Student’s t-test was applied in order to compare the first- and second-trimester markers in MoM values between the IVF- and naturally conceived pregnancies. A P-value < 0.05 was considered significant in both tests. Statistical analysis was performed by the Tel Aviv University statistical department using Statistics Package for Social Sciences software.

**Results**

The present study was limited to normal pregnancies from a general obstetric population. This included 285 naturally and 71 IVF-conceived singletons for whom the results of both screenings tests and complete data on pregnancy outcome were available.

The mean log10 MoMs between each marker within each screening tests were compared, and median MoM values are listed in Table I. The mean age of women in the IVF group was greater than that of the natural conception group, and their fetuses had a smaller mean CRL value. The IVF group also had a significantly lower PAPP-A in the first trimester, and increased AFP levels in the second trimester (Table I). Based on maternal age alone (≥35 years), 24 and 14% of the IVF and natural conception parturient women respectively were defined as screen positive (Table II). The FPR rate among the IVF parturient women increased from 7% during the first trimester to 10% in the second (Table II). Similar screen-positive results were found throughout both tests (1.4%) in both groups (Table II). The amniocentesis uptake due to screen-positive test results (avoiding repetition) was 15 and 13% for the IVF- and natural conceptions respectively (Table II).
Naturally conceived versus IVF pregnancy: serial DS screening

Table I. Comparison of maternal age, crown–rump length (CRL) and first- and second-trimester screening tests\(^a\) between natural versus IVF-conceived singleton pregnancies

<table>
<thead>
<tr>
<th>Conception</th>
<th>Maternal age (years)(^b)</th>
<th>CRL (mm)</th>
<th>Gestational age (weeks)(^b)</th>
<th>NT(^c)</th>
<th>FβHCG(^c)</th>
<th>PAPP-A(^c)</th>
<th>AFP(^c)</th>
<th>HCG(^c)</th>
<th>Unconjugated estriol(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural (n = 285)</td>
<td>30 ± 4</td>
<td>59 ± 11</td>
<td>12.2 (range 11–13)</td>
<td>1.06 ± 0.7</td>
<td>1.06</td>
<td>1.05</td>
<td>1.07</td>
<td>1.08</td>
<td>1.04</td>
</tr>
<tr>
<td>IVF (n = 71)</td>
<td>31.5 ± 5</td>
<td>54</td>
<td>11.5 (range 11–12)</td>
<td>1.16 ± 0.4</td>
<td>1.16</td>
<td>0.96</td>
<td>1.13</td>
<td>1.12</td>
<td>0.94</td>
</tr>
<tr>
<td>(P)</td>
<td>&lt; 0.05</td>
<td>&lt; 0.05</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.005</td>
<td>&lt; 0.05</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

\(^a\)Comparison of the first- and second-trimester log\(_{10}\) markers.
\(^b\)Values are mean ± SD.
\(^c\)Values are median MoM.

\(\alpha\)-fetoprotein; FβHCG = free βHCG; NS = not significant; NT = nuchal translucency; PAPP-A = pregnancy-associated plasma protein-A.

Table II. Comparison of first and second trimester screening tests results and pregnancy follow-up between natural versus IVF-conceived singleton pregnancies. Values in parentheses are percentages

<table>
<thead>
<tr>
<th>Conception</th>
<th>Maternal age &gt;35 years</th>
<th>First trimester screen-positive (&gt;1:380)</th>
<th>Second trimester screen-positive in both tests</th>
<th>Screen-positive in both tests</th>
<th>Amniocentesis uptake(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural (n = 285)</td>
<td>43 (15)</td>
<td>26 (9)</td>
<td>14 (5)</td>
<td>4 (1.4)</td>
<td>36 (13)</td>
</tr>
<tr>
<td>IVF (n = 71)</td>
<td>17 (24)</td>
<td>5 (7)</td>
<td>7 (10)</td>
<td>1 (1.4)</td>
<td>11 (15)</td>
</tr>
<tr>
<td>Pearson (\chi^2)</td>
<td>(P &lt; 0.05)</td>
<td>NS</td>
<td>(P &lt; 0.05)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\(^a\)Amniocentesis uptake because of screen-positive results in either test (avoiding repetition). NS = not significant.

There were no fetal or neonatal deaths in either group. No prenatal anomalies or traits that would have justified karyotyping were detected post-natally. Long-term follow-up (mean 8 months; range 3–18) was available for all infants included in this study and for whom there was an uneventful pregnancy and delivery outcome. No other genetic abnormalities were detected, and the motor and mental development of the infants appeared normal. The condition of all infants was checked by an experienced paediatrician.

Discussion

The current study addresses serial first- and second-trimester serum screening as applied to preselected IVF-achieved pregnancies. With regard to first-trimester assessment, the IVF fetuses had a smaller CRL, though this may have reflected an earlier uptake of testing. The NT measurements among IVF cases were slightly increased, but not in any statistically significant manner. These findings were compatible with cross-sectional data which demonstrated an increase in NT measurement between 9 and 12 weeks gestation, followed by a decrease at 13–14 weeks (Braithwaite et al., 1996). NT measurements were all within the normal range for gestation (Herman et al., 2000a) in both groups of the study population; NT was consistently not influenced by ART and this is in agreement with previous reports (Maymon et al., 1999b; Lara et al., 2000; Liao et al., 2001; Wojdemian et al., 2001). The most marked difference between the naturally conceived pregnancies and those obtained after IVF was hormonal treatment and multiple corpora lutea, all of which persisted throughout the first trimester of gestation. Thus, it might be expected that marker changes induced by this treatment would be more marked during the first trimester, and be reflected mainly by high HCG levels (Wojdemian et al., 2001).

Surprisingly, no significant changes were found in this marker in both trimesters; likewise, with regard to the first trimester this was in agreement with previously published data (Wojdemian et al., 2001). In addition, low levels of PAPP-A were found, in agreement with other data (Liao et al., 2001). The lower PAPP-A might be attributed to early testing in the IVF group, as it has been reported that PAPP-A levels increase significantly with gestation (Ong et al., 2000). Another theory, however—and one which deserves further investigation—is that the lower PAPP-A level in IVF cases is secondary to metabolic impairment, which also reflects infertility cases. Lower maternal serum PAPP-A levels at 10–14 weeks gestation were recently reported as being associated with the subsequent development of various complications of pregnancy (Ong et al., 2000). The differential diagnosis between such complications, or the constitutional characteristic of the IVF cases, should be further investigated. A FPR rate of 7–9% was found at a threshold of 1:380 for first-trimester screening in both groups. This was lower than the 16% FPR rate found for the same test previously reported elsewhere among IVF cases (Liao et al., 2001), but was higher when compared with data obtained at other centres (Wojdemian et al., 2001) or when the NT screening test was used alone in ART pregnant women (Maymon et al., 1999c).

Interestingly, a significantly higher rate of second-trimester FPR (10%) was found among IVF cases in comparison with the natural conception group (5%). Again, this observation was in agreement with a previous report which identified...
a similar mid-gestation FPR rate among ART pregnancies (Maymon et al., 1999b,c; Maymon and Shulman, 2001) whereas Bar Hava et al., have reported 18% mid-gestations FPR from a different centre in Israel (Bar-Hava et al., 2001).

After having corrected for gestational age by dating scan, and having checked for maternal age (which did not change from the first- to the second-trimester test), it was contended that the high second-trimester FPR rate in IVF patients could be attributed mainly to other factors that are directly associated with the pregnancy itself. The reasoning behind this was that such pregnancies in infertile couples might have another underlying pathology that may lead to various (as yet unknown) metabolic disturbances, and these might also affect the serum screening results. It has been suggested (Wald et al., 1999a) that a high HCG level might be explained by progesterone levels remaining high in IVF pregnancies. In contrast, the findings of a previous report (Maymon and Shulman, 2001) and data from the present study are in agreement with the outcome of a recent in-vivo study which demonstrated an elevated maternal progesterone serum level, but without a concomitant increase in HCG concentration in early first-trimester IVF pregnancies (Costea et al., 2000). The present data are in agreement with two other studies on IVF cases (Muller et al., 1993; Lam et al., 1999) in which the mid-gestation serum HCG marker levels were similar to those of the natural conception group. Three studies which assessed mid-gestation serum markers among IVF cases have reported, in addition to a higher HCG level, a higher rate of complicated obstetric outcome (Maymon and Shulman, 2001), including proteinuric pre-eclampsia (Heinonen et al., 1996), premature deliveries, and smaller infants (Ribbert et al., 1996). In this respect, it has been found that pregnant women who later in gestation developed pre-eclampsia or have growth-impaired infants tend also to have both higher HCG and AFP levels (Ogle et al., 2000). Thus, it is possible that the high mid-gestation HCG level found among IVF pregnancies provides a warning for such complications rather than for a DS risk.

A significantly increased AFP level was also found among the current IVF pregnancy cases, and this was even higher than shown previously in this population (Maymon and Shulman, 2001). There was no clear explanation for this finding. Several problems emerged during the present study. One problem involved counselling the patient in whom the results were conflicting, and another problem involved the high number of invasive procedures, partly because the screen-positive results could not be adjusted by other screen-negative reports. This summation of the two screening test results led to a suboptimal interpretation of the sequential disclosure screening results. Currently, 15 and 13% of the IVF and natural conception pregnancies respectively were found to be screen-positive, and the women subsequently underwent amniocentesis. Based on previously published data, it has been reported recently (Hackshaw and Wald, 2001a) that ~0.07% of all women found screen-positive on the combined test remain screen-positive after the integrated test. In the present study, higher values of 1.4% screen-positive tests were found in both groups which underwent the disclosure screening tests.

The advantage of a method which combines the results of different screening tests in order to reach a single value is clear. For example, it may significantly reduce the FPR rate and even increase the detection rate. Previously, a risk of \( \geq 1:120 \) was used as the threshold to define a positive result on the integrated first- and second-trimester screening test (Wald et al., 1999b). The model used by these authors consisted of NT and PAPP-A from the first- trimester test, and uE3, HCG, AFP and inhibin A from the second-trimester test. Accordingly, the detection rate was expected to be 85% with a FPR rate of 0.9% (Wald et al., 1999b). This approach seems very promising, mainly for IVF-conceived pregnancies which bear a high second-trimester FPR rate. As yet, it is not possible to comment on the detection rate as the present study was focused only on normal cases, and it is not possible to provide integrated results as both tests included either FBHCG or HCG. It has been shown previously that the two markers correlate well, both in unaffected and DS cases (Maymon et al., 2001). A recent report (Hackshaw and Wald, 2001b), using published data, suggested that if the correct age-specific risk and screening marker distribution were used, then risk estimation would be accurate. However, the use of two tests would be less efficient than integrating all the information into a single test. Therefore, if an integrating screening programme is planned, based on our previous experience (Herman et al., 2000b) it is contended that a combination of NT (which is unaffected by IVF pregnancies) and mid-gestation TT offers a reasonable and simple alternative.

Otherwise, first-trimester screening using either NT alone or combined with serum screening seems to be an efficient option (Liao et al., 2001; Wejdeman et al., 2001), and this approach—at least in our experience—yields the lower FPR rate. Larger prospective studies are required to verify the best screening policy for IVF pregnancies.

Acknowledgements

The authors thank the Zer Medical Laboratories (Israel) for their cooperation and support, and Ms Esther Eshkol for editorial assistance.

References


Submitted on November 27, 2000; resubmitted on August 24, 2001; accepted on November 21, 2001.