Engels MAJ, Pajkrt E, Groot DT, Schats R, Twisk JWR, van Vugt JMG.

Validation of correction factors for serum markers for first-trimester Down syndrome screening in singleton pregnancies conceived with assisted reproduction

Abstract

Objective
To validate previously computed correction factors for free-β human chorionic gonadotrophin (fβ-hCG) and pregnancy-associated plasma protein-A (PAPP-A) in in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) pregnancies with hormone treatment and to determine the effect on false-positive rate (FPR).

Methods
Retrospective study on 249 IVF and 250 ICSI cases and 20,190 controls. Correction factors 1.42 (PAPP-A), 1.17 (fβ-hCG) in IVF; 1.56 (PAPP-A) in ICSI were applied on the absolute serum concentrations. Analysis was done on log10 transformed multiples of medians (MoMs).

Results
In the controls mean PAPP-A and fβ-hCGMoM were 1.004 and 1.062. Before correction mean PAPP-A MoM was significantly lower in IVF (0.757; P < 0.001) and in ICSI (0.671; P < 0.001) and after correction comparable (1.071; P = 0.053 in IVF; 1.048; P = 0.178 in ICSI). Before correction mean fβ-hCGMoM was comparable (1.054; P = 0.59 in IVF and 1.051; P = 0.56 in ICSI) and after correction significantly higher in IVF (1.241; P < 0.001). After correction the likelihood for receiving a false-positive result was 1.03 in IVF pregnancies (95% CI 0.98-1.09; P = 0.248) and 1.02 in ICSI pregnancies (95% CI 0.97-1.07; P = 0.448).

Conclusions
After correction the FPR in IVF and ICSI pregnancies with hormone treatment reduces to the observed FPR in the controls.
Introduction

Prenatal screening for chromosomal abnormalities has become part of routine antenatal care in the Netherlands over the past decade. Since January 2007, all pregnant women, regardless of age, are informed about the first-trimester screening (FTS) options for Down syndrome (DS), combining maternal age, fetal nuchal translucency (NT) thickness and maternal serum free β-human chorionic gonadotrophin (fβ-hCG) and pregnancy-associated plasma protein-A (PAPP-A)[1]. This enables pregnant women and their partners to make an autonomous informed decision about whether or not to participate in this prenatal screening program. As a consequence of the changing pattern of childbirth with Dutch women postponing childbirth until in later life [2], infertility is more common, as is utilization of assisted reproductive techniques (ART) [3,4].

In ART pregnancies maternal serum marker levels change during the first trimester affecting the risk assessment for DS [5-26]. Counselling for prenatal screening of women carrying an ART pregnancy is complicated due to contradictory reports in FTS performance in ART pregnancies. Several studies reported higher false-positive rate (FPR) compared to spontaneously conceived pregnancies [5-11], some studies found higher FPR even after correction for maternal age [12-15] and others reported similar FPR [16-20]. Whether the prevalence of fetal chromosomal abnormalities in ART pregnancies is different from naturally conceived pregnancies is however unclear. Bettio et al. (2008) [27] reported no higher prevalence for chromosomal abnormalities whereas other studies showed a higher prevalence in ART pregnancies mainly related to a higher number of sex chromosome aneuploidies [28-31].

In our previous case-control study we computed correction factors for the maternal serum parameters in selected singleton in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) pregnancies (all cases with fresh embryo transfer and hormone treatment and normal pregnancy outcome) compared to matched controls. Cases with diabetes and chromosomal abnormalities (except for DS) were excluded. The computed correction factors correct for the mode of conception (fresh embryo transfer with hormone treatment), and exclude the effect on serum parameters due to other confounding factors. The effect on screening performance cannot be determined, because of limited numbers of ART pregnancies affected with DS. Determining the FPR is the first step to evaluate whether correction for serum markers is appropriate in ART pregnancies.

The exact effect of correction on the FPR cannot be predicted, because the correlation of risks before and after adjustment of maternal serum markers is not of the same magnitude as the correlation of marker change. It is also not known in advance whether a pregnancy (ART or non-ART), will be complicated. The effect of applying a correction on the FPR should therefore be studied in non-selected (complicated and uncomplicated) ART pregnancies in comparison to an overall screening population.

The aim of this study was to validate the correction factors for PAPP-A and fβ-hCG, in non-selected ART pregnancies based on fresh embryo transfer with hormone treatment in comparison to an overall screening population and to determine the effect on the FPR.
Materials and methods

This retrospective cohort study included all the ART data from the VU University Medical Center (VUMC) and its IVF referral centres from January 1st, 2008 until January 1st, 2011 and the Academic Medical Center (AMC) and its IVF referral centres from January 1st, 2008 until January 1st, 2010. Moreover, all the data on the FTS for DS performed in the VUMC in the period from January 1st, 2008 until January 1st, 2011 were used as an overall screening population, with exclusion of the data of ART pregnancies.

We collected detailed information on the ART cycles, including the type of ART treatment, whether embryos were fresh or frozen-thawed, the use of donor gametes or embryos and the number of ongoing pregnancies. ART cases based on oocyte donation, frozen-thawed embryo conception or intrauterine insemination were excluded from this study. For all the IVF and ICSI cases, a standard IVF/ICSI procedure was carried out. Women aged 38 years or younger underwent controlled ovarian hyperstimulation with a “long” protocol with gonadotropin-releasing-hormone agonist. In women older than 38 years, a “short” stimulation protocol was applied [32].

In all the cases serum was sampled at 9 – 14 weeks of gestation. Serum samples of the ART cases of the AMC were analyzed at the endocrine laboratory of the AMC and serum of the ART cases of the VUMc and of the overall study population were analyzed at the endocrine laboratory of the VUMc. In both laboratories the Delfia Xpress (PerkinElmer, WallacOy, Turku, Finland) was used for analysis. Both laboratories fulfilled the quality control demands of UK NEQAS (United Kingdom National External Quality Assessment Service) during the study period. For all the cases of the AMC DS risk was recalculated, using the risk software program Ellips / Lifecycle 2.2 (PerkinElmer, WallacOy, Turku, Finland). Gestational age (GA) at serum sampling was based on a first-trimester dating scan. The NT measurements were carried out in accordance with the FMF protocol [33]. The VUMc NT reference curve [34] was used allowing NT measurements at a fetal crown rump length (CRL) between 45 and 79 mm. The NT and the CRL measurements as well as information on an earlier pregnancy with Down syndrome, smoking habits, and maternal weight were taken into account for the risk assessment on DS. For increased risk a cut-off value of 1:200 was used (mid-term). Data on the diagnostic procedures (i.e. chorion villi sampling and amniocentesis) were provided by the cytogenetic laboratories of the VUMc and AMC. Follow-up was evaluated by questionnaires and delivery room records.

The following correction factors for the absolute serum concentrations were used: in the IVF group 1.42 for PAPP-A and 1.17 for fβ-hCG and in the ICSI group 1.56 for PAPP-A [15]. The computed correction factor of 1.05 for fβ-hCG in the ICSI group was not statistically significant and therefore it was not used in this study. The absolute PAPP-A and fβ-hCG concentrations were multiplied by the correction factors and DS risk was recalculated based on the corrected absolute serum values together with the other parameters necessary for the risk assessment (all with the exact same values as before correction).

Statistical analyses were conducted using SPSS software (version 20). Independent samples t-tests and Chi-square tests were performed to test the differences between the baseline characteristics of
the ART groups and the overall screening population. Because of skewed distributions the data of \( \beta \)-hCG and PAPP-A were log-transformed. Results were then back-transformed and are presented as geometric means. Linear regression analysis was done to adjust for relevant confounders. Two-sided \( P < 0.05 \) was considered to reflect statistical significance. The statistical significance of differences in false-positive rates were expressed as odds ratios (ORs) with 95% confidence intervals (CIs).

The study was approved by the Ethics Committee of the VUMC.

Results

We included 249 IVF and 251 ICSI pregnancies; the overall screening population in the selected years consisted of 20,190 women. The selection of the IVF and the ICSI pregnancies is shown in Figure 1. Of the ART cases 20 (9 IVF and 11 ICSI) cases were lost to follow-up. Data on maternal weight were missing in 5 IVF and 7 ICSI cases, data on smoking in 3 IVF and 3 ICSI cases and data on ethnicity in 8 IVF and 11 ICSI cases. Follow-up on the overall screening population was 80%.

Mean maternal age in both the IVF group (36.2 years) and the ICSI group (34.6 years) was significantly higher at the time of screening compared to the overall screening population (33.1 years). In the ART cases serum samples were taken at an earlier gestational age and the measured CRL was smaller than in the overall study population. The distribution of smokers was comparable between the ART groups and the overall study group. All study populations were predominantly of Caucasian origin (Table 1).

Table 1. Baseline characteristics of the ART populations and the overall study population. \( P \)-values are based on the comparison of the IVF or ICSI group with the overall screening population. All values are reported as means (SD), except for smoking and ethnicity.

<table>
<thead>
<tr>
<th></th>
<th>IVF</th>
<th>P-value</th>
<th>ICSI</th>
<th>P-value</th>
<th>Overall screening population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 249</td>
<td></td>
<td>N = 251</td>
<td></td>
<td>N = 20,190</td>
</tr>
<tr>
<td>Maternal age (in years)</td>
<td>36.2 (± 3.8)</td>
<td>&lt;0.001</td>
<td>34.6 (± 3.9)</td>
<td>&lt;0.001</td>
<td>33.1 (± 4.2)</td>
</tr>
<tr>
<td>Maternal weight (kg)</td>
<td>69.3 (± 13.6)</td>
<td>0.06</td>
<td>68.9 (± 11.4)</td>
<td>0.125</td>
<td>67.8 (± 11.9)</td>
</tr>
<tr>
<td>GA at serum sampling (days)</td>
<td>78.1 (± 7.7)</td>
<td>0.021</td>
<td>77.6 (± 7.9)</td>
<td>&lt; 0.001</td>
<td>79.3 (± 7.9)</td>
</tr>
<tr>
<td>CRL (in mm)</td>
<td>59.2 (± 7.4)</td>
<td>&lt; 0.001</td>
<td>59.0 (± 7.0)</td>
<td>&lt; 0.001</td>
<td>61.0 (± 6.8)</td>
</tr>
<tr>
<td>Number of smokers</td>
<td>9 (3.6)</td>
<td>0.81</td>
<td>8 (3.2)</td>
<td>0.56</td>
<td>790 (4.6)</td>
</tr>
<tr>
<td>Ethnicity (% Caucasian)</td>
<td>88.7</td>
<td>0.55</td>
<td>89.2</td>
<td>0.73</td>
<td>89.9</td>
</tr>
</tbody>
</table>
In Table 2 the geometric mean PAPP-A MoM, fβ-hCG MoM values (before and after correction) and NT MoM values in the ART pregnancies are shown compared to the overall study population. Before correction PAPP-A values were significantly lower in the ART groups compared to the controls (0.757 versus 1.004; \( P < 0.001 \) in IVF; 0.671 versus 1.004; \( P < 0.001 \) in ICSI) and after correction mean PAPP-A MoM values were comparable (1.071 versus 1.004; \( P = 0.053 \) in IVF; 1.048 versus 1.004; \( P = 0.178 \) in ICSI). Before correction mean fβ-hCG MoM was comparable to the controls (1.054 versus 1.062; \( P = 0.59 \) in IVF; 1.051 versus 1.062; \( P = 0.56 \) in ICSI) and after correction mean fβ-
hCGMoM was significantly higher in IVF (1.241 versus 1.062; \( P < 0.001 \)). The differences in NT were not significantly different. After adjusting for maternal age, maternal weight and GA at serum sampling similar results were found.

**Effect on false-positive results**

In the overall screening population an increased risk for DS was found in 1549 cases including 81 cases with DS and 91 cases with other chromosomal abnormalities. All fetal aneuploidies detected with FTS were counted as true-positives. The FPR in the overall screening group was 6.8% (1377 false-positive results of 20,190). Before correction 51 of 249 IVF pregnancies showed an increased DS risk, including 3 cases with DS, one case of trisomy 18, one case with mosaicism and one case with Klinefelter syndrome. In the ICSI group 49 of the 251 cases had an increased risk. There were no cases with chromosomal abnormalities in this group. Adjustment of the serum values had no consequences for the detection of the cases with DS, trisomy 18 and mosaicism in the IVF, nor did it lead to new cases with a screen-positive result in the ART groups.

Table 3 shows the comparison between the ART populations and the overall screening population for the proportion of women receiving a false-positive result with FTS. After adjusting for maternal age, maternal weight and GA at serum sampling, the FPR of women conceiving with ART after correction for the serum parameters was comparable to that of the overall screening population (in IVF OR 1.03, 95% CI 0.98-1.09; \( P < 0.248 \) and in ICSI OR 1.02, 95% CI 0.97-1.07; \( P < 0.448 \)). Because the correction for \( fβ\)-hCG seemed redundant, risk was recalculated for the IVF cases without applying a correction for \( fβ\)-hCG. Six of the IVF cases no longer had a screen-positive result. This would result in a further decrease of the FPR in the IVF group to 13.6% (Table 3).
Table 2. Geometric mean PAPP-A MoM, fβ-hCGMoM(before and after correction) and NT MoM values in ART pregnancies compared to the overall study population. Linear regression analysis was done to adjust for differences in maternal age, maternal weight and GA at serum sampling. *P*-values are based on the comparison of the IVF and ICSI group (before and after correction) with the overall screening population.

<table>
<thead>
<tr>
<th></th>
<th>PAPP-A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Geometric mean MoM</td>
<td>Linear regression analysis</td>
<td></td>
<td>fβ-hCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td></td>
<td></td>
<td>Linear regression analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Geometric mean MoM</td>
<td>P-value</td>
<td></td>
<td></td>
<td>Geometric mean MoM</td>
<td>P-value</td>
</tr>
<tr>
<td>Overall study population</td>
<td>1.004</td>
<td>&lt; 0.001</td>
<td>1.062</td>
<td>0.59</td>
<td>1.047</td>
<td>0.321</td>
</tr>
<tr>
<td>IVF (before correction)</td>
<td>0.757</td>
<td>&lt; 0.001</td>
<td>1.54</td>
<td>0.201</td>
<td>0.321</td>
<td></td>
</tr>
<tr>
<td>IVF (after correction)</td>
<td>1.071</td>
<td>0.053</td>
<td>1.24</td>
<td>&lt; 0.001</td>
<td>Not corrected</td>
<td></td>
</tr>
<tr>
<td>ICSI (before correction)</td>
<td>0.671</td>
<td>&lt; 0.001</td>
<td>1.05</td>
<td>0.446</td>
<td>0.703</td>
<td></td>
</tr>
<tr>
<td>ICSI (after correction)</td>
<td>1.048</td>
<td>0.178</td>
<td>0.154</td>
<td>Not corrected</td>
<td>Not corrected</td>
<td></td>
</tr>
</tbody>
</table>

*Not corrected*
Table 3. Comparison of false-positive results of FTS between the ART populations and the overall screening population before and after correction. *P*-values are based on the comparison of the ART groups with the reference population.

<table>
<thead>
<tr>
<th></th>
<th>FTS Number</th>
<th>False-positive results</th>
<th>Univariate analysis</th>
<th>Multivariate analysis adjusting for maternal age, weight and GA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>%</td>
<td>OR</td>
</tr>
<tr>
<td>Overall screening population</td>
<td>20 190</td>
<td>1391</td>
<td>6.8</td>
<td>Reference</td>
</tr>
<tr>
<td>IVF before correction</td>
<td>249</td>
<td>45</td>
<td>18.0</td>
<td>3.05</td>
</tr>
<tr>
<td>IVF after correction</td>
<td>249</td>
<td>39</td>
<td>15.7</td>
<td>1.11</td>
</tr>
<tr>
<td>IVF after correction for only PAPP-A</td>
<td>249</td>
<td>34</td>
<td>13.6</td>
<td>1.08</td>
</tr>
<tr>
<td>ICSI before correction</td>
<td>251</td>
<td>49</td>
<td>19.5</td>
<td>1.76</td>
</tr>
<tr>
<td>ICSI after correction</td>
<td>251</td>
<td>27</td>
<td>10.8</td>
<td>1.05</td>
</tr>
<tr>
<td>Total ART before correction</td>
<td>500</td>
<td>94</td>
<td>18.8</td>
<td>3.07</td>
</tr>
<tr>
<td>Total ART after correction</td>
<td>500</td>
<td>60</td>
<td>12.0</td>
<td>1.08</td>
</tr>
<tr>
<td>Total ART after correction for only PAPP-A</td>
<td>500</td>
<td>55</td>
<td>11.0</td>
<td>1.07</td>
</tr>
</tbody>
</table>
Discussion

To our knowledge, this is the first study that demonstrates that applying a correction for the absolute serum values of PAPP-A in ART pregnancies results in a significant decrease in the FPR comparable to the FPR in an overall screening population. In a previous case-control study of uncomplicated IVF and ICSI pregnancies based on fresh embryo transfer with hormone treatment correction factors were computed. Cases with adverse pregnancy outcome (pregnancy-induced hypertension, intrauterine growth restriction, intrauterine fetal death, preterm delivery (<37 weeks) and diabetes), were excluded because of the association of these conditions with low PAPP-A levels [35-40]. The selection of uncomplicated ART cases based on fresh embryo transfer with hormone treatment for computing the correction factors is supported by the study of Amor et al. (2009) [13]. They showed that in ART pregnancies with adverse outcome (stillbirth, neonatal death, prematurity, birthweight<2500 g) and/or with obstetric complications (preeclampsia, pregnancy-induced hypertension, gestational diabetes) the PAPP-A levels were significantly lower than in ART pregnancies with normal outcome. In ART (fresh embryo transfer and frozen-thawed embryo transfer) pregnancies with hormone treatment the PAPP-A levels were decreased and in ART (fresh embryo transfer and frozen-thawed embryo transfer) pregnancies without hormone treatment the PAPP-A levels were not significantly different from those of non-ART pregnancies. Because the computed correction factors adjust solely for the mode of conception, they are equally applicable in complicated and non-complicated ART pregnancies based on fresh embryo transfer with hormone treatment. Therefore the effect of applying a correction on the FPR is studied in non-selected ART pregnancies as compared to an overall screening population. Bias of cases with adverse pregnancy outcomes, affecting the PAPP-A levels, is not to be expected, because the PAPP-A values in complicated ART pregnancies and complicated non-ART pregnancies are comparable [13].

The effect of ART on PAPP-A and fβ-hCG levels in IVF and ICSI pregnancies have yielded contradictory results. Most studies showed decreased PAPP-A levels in ART pregnancies combined with unaltered levels of fβ-hCG [9,13,14,16,17,20,21] increased fβ-hCG levels [12,22-24] or even decreased fβ-hCG levels [15,25]. Two studies reported only an increase in fβ-hCG levels [10,26], and some studies reported no changes in maternal serum values in ART pregnancies [11,18,19]. These conflicting results are probably due to small sample sizes and/or the use of unselected (complicated as well as non-complicated) ART pregnancies and without distinguishing between pregnancies with and without hormone treatment.

This study shows significant lower PAPP-A levels in ART pregnancies compared to an overall screening population, whereas fβ-hCG values were not significantly different. The fact that the fβ-hCG values in the selected IVF pregnancies in the previous case-control study were significantly different (correction factor 1.17) from the values in this study, might be explained by a selection bias. In all cases of the case-control study serum was analyzed at the laboratory of the VUMC, whereas in the current study serum of the ART cases was analyzed in one of the two laboratories. Small differences
Validation of correction factors for serum markers for first-trimester Down syndrome screening in singleton pregnancies conceived with assisted reproduction

in median serum MoM values cannot be ruled out with immunoassay analyses. There were no changes in the stimulation protocol that was applied.

Application of the proposed correction factors for PAPP-A resulted in expected PAPP-A values in the ART groups, whereas correction for fβ-hCG seemed redundant. The computed correction factors for PAPP-A are appropriate and almost equal. For the use in daily practice it is more convenient to use the same correction factor for IVF and the ICSI cases. It should be noted that the computed correction factors in this study might not be appropriate for other screening populations. The degree of correction is dependent on the local median PAPP-A MoM value. Therefore each screening center should calculate a correction factor based on their own uncomplicated ART population with hormone treatment.

A recent study demonstrated that significant gestational age effects should be taken into account when applying a correction for ethnicity and IVF [26]. In contrary to our study they found increased fβ-hCG levels and no difference in PAPP-A levels between IVF pregnancies and controls, but no distinction was made in complicated and non-complicated ART pregnancies nor in pregnancies with and without hormone treatment. Therefore further investigation needs to be undertaken to determine the effect of gestational age on correction for the different subgroups of ART pregnancies.

After correction a distinct decrease in the FPR in IVF and ICSI pregnancies was shown, comparable to the observed FPR in the overall screening population when adjusted for maternal age and GA at serum sampling. An effect of ethnicity in our study populations on serum marker levels is not to be expected, because the distribution of ethnicity is comparable between the ART groups and the overall screening population and the populations are predominantly of Caucasian origin.

Limitations of our study are the small sample sizes of the IVF and the ICSI populations and the limited number of affected pregnancies with DS in the ART populations. Therefore screening performance cannot be determined. We assume that even after correction for PAPP-A, the differences in the serum marker values and the NT values between affected and unaffected pregnancies will still be discriminatory enough to detect the DS cases. It is to be expected that the odds of being affected given a positive result (OAPR) will improve and the number of procedure-related miscarriages will decrease. The uptake of the FCT in the ART pregnancies is low at about 30%, but it is comparable to the overall uptake in our country (approximately 25%) (41) and therefore it is not expected that this could have caused a potential selection bias.

The exact pathogenesis of the differences in maternal serum PAPP-A concentrations between ART pregnancies and spontaneously conceived pregnancies is still unclear. Endocrine changes early in pregnancy are the result of complex interactions between the corpus luteum, endometrium, placenta and the embryo. It is hypothesized that Inhibin A secreted by corpora lutea inhibits the secretion of PAPP-A [16]. Thus ovarian stimulation might play an important role in the biological cascade leading to lower PAPP-A values in ART pregnancies. It has been shown that the hormone treatment that accompanies ART cycles is strongly associated with the decreased PAPP-A levels in ART pregnancies. A model is proposed whereby the hormone treatment with embryo transfer results in abnormal levels of ovarian steroid hormones and other factors causing a reduction in PAPP-A levels, likely mediated by the effect of hormones on the endometrium [13]. It has also been suggested that
lower PAPP-A values in ART pregnancies might be the result of metabolic impairments related to infertility in the mother such as ovulatory disorders and defective endometrial development [17]. Lower PAPP-A levels are also associated with adverse pregnancy outcome [35-39]. Lower PAPP-A levels in ART pregnancies might implicate that these pregnancies are possibly at a higher risk of the adverse pregnancy outcome that is known to be more common in ART populations [17,28].

It can be concluded that a correction factor for PAPP-A in ART pregnancies (fresh embryo transfer with hormone treatment) is justified and that correction for ß-hCG is unnecessary. The FPR in ART pregnancies decreases to the observed FPR of the overall screening population. The OAPR is expected to decrease. Our results underline the importance of implementation of a correction for PAPP-A, because it will improve the reliability of the pre- and post-test counselling for women carrying an ART pregnancy and it will lower the uptake of invasive diagnostic testing. As for the ART pregnancies conceived after frozen-thawed embryo transfer with hormone treatment, it is to be expected that a correction should be applied, but further research is necessary to determine the effect of applying a correction.

Acknowledgements

We thank A.W.M. Nieuwint for her support to this study.

Author’s role

J.M.G.V. and M.A.J.E. were responsible for the research design, performed data analysis, interpreted the results and wrote the paper. E.P. provided the data of ART pregnancies of the AMC, attributed to the design, assisted with the interpretation of the results and edited the paper. D.G. contributed to the acquisition of the data (performed all the risk recalculations and data linkage), assisted with the interpretation of the results and edited the paper. R.S. provided the data of ART pregnancies of the VUmc, assisted with the interpretation of the results and edited the paper. J.W.R.T. performed the statistical analyses, interpreted the results and wrote the paper.
References


