Is 3D technique superior to 2D in Down syndrome screening? Evaluation of six second and third trimester fetal profile markers

F. I. Vos ¹
M. Bakker ¹
E. A. P. de Jong-Pleij ²
L. S. M. Ribbert ²
E. Tromp ³
C. M. Bilardo ¹

¹ Fetal Medicine Unit, University Medical Centre Groningen, Groningen, the Netherlands.
² Department of Obstetrics and Gynecology, St. Antonius Hospital, Nieuwegein, the Netherlands.
³ Department of Statistics, St Antonius Hospital, Nieuwegein, the Netherlands.

Is 3D technique superior to 2D in Down syndrome screening? Evaluation of six second and third trimester fetal profile markers

F. I. Vos¹, M. Bakker¹, E. A. P. de Jong-Pleij², L. S. M. Ribbert², E. Tromp¹ and C. M. Bilardo¹

¹ Fetal Medicine Unit, University Medical Centre Groningen, Groningen, the Netherlands.
² Department of Obstetrics and Gynecology, St Antonius Hospital, Nieuwegein, the Netherlands.
³ Department of Statistics, St Antonius Hospital, Nieuwegein, the Netherlands.

Objective:
The objective of this article is to investigate whether in the clinical setting of second trimester ultrasound (US) investigations, 3D multiplanar correction prior to the measurement of Down syndrome (DS) facial markers (nasal bone length, prenasal thickness, fetal profile line, maxilla-nasion-mandible angle, prenasal thickness to nasal bone length ratio, and prefrontal space ratio) is superior to subjective judgment of a correct midsagittal plane by 2D technique.

Methods:
Measurements were performed on 2D images and 3D volumes (corrected to the midsagittal plane), acquired during the same scanning session.

Results:
All six markers were measured in 105 datasets (75 of euploid fetuses and 30 of DS fetuses). The maxilla-nasion-mandible angle measured on 2D images was significantly larger than on 3D volumes (p < 0.01). In all other markers, there was no significant difference between measurements performed on 2D images or 3D volumes. No statistical difference was found for any marker between measurements performed on images acquired by either 2D or 3D US in their ability to discriminate between normal and DS fetuses.

Conclusion:
Nasal bone length, prenasal thickness, fetal profile line, prenasal thickness to nasal bone length ratio, and prefrontal space ratio can be confidently used as DS markers in second trimester US examinations performed by 2D US.

Introduction

Specific facial profile features of Down syndrome (US) fetuses have been investigated and used as second and third trimester markers.¹⁻¹³ The nasal bone length (NBL) was the first to be extensively investigated, followed by the prenasal thickness (PT). Recent studies have shown that the ratio between these two markers (PT-NBL ratio) and the prefrontal space ratio (PFSR) yields an even better detection rate.⁶,¹⁰ Furthermore, we have previously investigated the maxilla-nasion-mandible (MNM) angle and fetal profile (FP)
line in both euploid and pathological cases. In countries such as the Netherlands, where participation in first trimester screening is low and many DS fetuses remain undetected until the 20 weeks scan, these markers may be of importance.

Several studies have compared 2D and 3D US imaging during gestation and suggested 3D to be superior by allowing a better identification of anatomical landmarks, a higher accuracy and reproducibility in measurements of structures in the fetal face and profile, including the NBL. In a previous study, we have shown that 2D images judged to be midsagittal in actual fact are not and need 3D multiplanar correction of in average 11.9 degrees (Y-axis) to 4.3 degrees (Z-axis) to become truly midsagittal. Clear landmarks to identify the exact midsagittal plane are missing when only 2D imaging is used, making it difficult to be absolutely sure to be in the exact midsagittal plane. However, it is not clear whether in a clinical setting, 3D imaging has an additional value in terms of an improved detection rate when compared with 2D.

The aim of this study is to compare the differences in 2D and 3D technique in the measurement and detection rate of facial markers in the second and third trimester.

Methods

Eligible cases were collected from the databases of the Fetal Medicine Unit of the University Medical Centre Groningen, which acts as a referral center, and of the Ultrasound Unit of the Saint Antonius Hospital in Nieuwegein, which performs US investigations of high-risk patients.

Images were obtained by a Voluson 730 Expert US machine or E8 system equipped with a RAB4-8L probe (GE Medical Systems, Kretz Ultrasound, Zipf, Austria).

Three-D volumes of euploid fetuses were retrieved from an available dataset used for a previous study of nonsmoking, healthy, low-risk, pregnant Caucasian women with a singleton and uncomplicated pregnancy. The dataset was collected prospectively; after 2D images of the FP were obtained, judged to represent the midsagittal plane and with the fetus facing the transducer, 3D volumes of the fetal face were acquired.

Databases of participating centers were searched for second and third trimester DS fetuses of Caucasian parents, collected between January 2006 and July 2013. All cases had been confirmed by karyotyping. The images were collected during clinical investigations and therefore were, in contrast to the images of the euploid fetuses, gathered in a less systematic fashion. For this study, cases with both a midsagittal 2D image of the FP and a 3D volume, acquired separately in the same scanning session, were included. We excluded 2D images that were obviously not midsagittal by systematically assessing all components of the profile. Images that showed a body of the mandible, a retrognathic appearance of the chin, a nostril, odd appearance of the nose, a frontal process of the maxilla, a sharp or blunt angle between the nasal and frontal bones, a bossing or sloping appearance of the forehead, os sphenoidale or a lateral ventricle, or plexus chorioideus were excluded. Visibility of the vomeral bone was considered a very strong indication of the exact midsagittal plane. A square shape of the mandible, normal appearance of lips, philtrum and nose, a flat or only slightly curved forehead, and visibility of the corpus callosum were indications for a good midsagittal plane.

In order to avoid bias, all the measurements on 2D images were performed first. This
was followed by multiplanar correction of the 3D volumes to the exact midsagittal plane with subsequent measurements.

To assess the interobserver and intraobserver variability of the measurement error, markers were remeasured after a 1-week interval. Markers were measured by two examiners (F. I. V. and E. J. P.), who were blinded to gestational age and to previous measurements, but not to karyotype.

The NBL, PT, FP line, MNM angle, PT-NBL ratio, and PFSR were all measured as described in previous studies of euploid and DS fetuses\(^6,10,14,15,22\) (Figure 1). Data were compared with the reference values derived from previous reports on euploid fetuses:\(^6,10,14,22\) the NBL\(^22\) and PT\(^22\) increased with gestation from 3.3 mm at 15 weeks' gestation to 9.6 mm at 33 weeks (NBL = -6.927 + (0.83 × GA) - (0.01 × GA\(^2\))) and from 2.3 mm at 15 weeks to 6.1 mm at 33 weeks (PT = 0.212 × GA - 0.873), respectively. The MNM angle,\(^14\) PT-NBL\(^6\) ratio, and PFSR\(^10\) were stable throughout gestation, with a mean of 13.5 degrees (95th centile = 16.9), 0.61 degree (95th centile = 0.80), and 0.97 degree (5th centile = 0.55), respectively.

Measurements below the 5th centile (for NBL and PFSR) or above the 95th centile (for MNM angle, PT, and PT-NBL ratio) of the reference ranges were considered abnormal. An FP line that was not ‘zero’ was considered abnormal.\(^15\)

The difference between the 2D and 3D measurement was analyzed in each individual fetus, of which a mean difference was calculated. Differences between measurements
were calculated for the whole group, and a separate analysis was performed in the group of DS fetuses, in order to assess if the use of one of the two techniques (2D vs. 3D) had an impact on the detection rate.
Table 1 – Mean values of 2D and 3D measurements (N=105) in a combined cohort of euploid and Down syndrome fetuses.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Mean 2D</th>
<th>Mean 3D</th>
<th>Mean difference (LOA)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBL (mm)</td>
<td>6.29</td>
<td>6.20</td>
<td>0.08 (-2.48 – 2.57)</td>
<td>0.84</td>
</tr>
<tr>
<td>FP Line</td>
<td>*</td>
<td>**</td>
<td>-</td>
<td>0.68</td>
</tr>
<tr>
<td>MNM (degrees)</td>
<td>14.77</td>
<td>13.75</td>
<td>1.03 (-5.14 – 6.13)**</td>
<td>0.40</td>
</tr>
<tr>
<td>PT (mm)</td>
<td>5.41</td>
<td>5.09</td>
<td>0.33 (-1.50 – 1.83)</td>
<td>0.83</td>
</tr>
<tr>
<td>PT-NBL ratio</td>
<td>0.91</td>
<td>0.89</td>
<td>0.03 (-0.45 – 0.49)</td>
<td>0.83</td>
</tr>
<tr>
<td>PFSR</td>
<td>1.04</td>
<td>1.09</td>
<td>-0.06 (-0.87 – 0.79)</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Corresponding mean difference with limits of agreement (LOA) and intra class correlation coefficients (ICCs) are reported. It was not possible to analyze mean differences between measurements of the FP line, as the outcome was noncontinuous.

* FP line in 2D: 77.5% positive, 22.5% negative.
** FP line in 3D: 79.8% positive, 20.2% negative.
*** Significant difference between 2D and 3D measurements in the MNM angle (p < 0.01).

MNM = maxilla-nasion-mandible angle. PT = prenasal thickness. PT-NBL ratio = prenasal thickness to nasal bone length ratio. PFSR = prefrontal space ratio.

STATISTICAL ANALYSIS

Intraclass correlation coefficients (ICCs) were calculated, and Bland–Altman analysis was performed to analyze intraobserver and interobserver variability. The student’s t-test was used to analyze differences between measurements. A p-value of less than 0.05 was considered statistically significant. Data were analyzed using the statistical software SPSS 20.0 for Windows (SPSS Inc., Chicago, IL, USA) and Excel for windows 2000.

Results

In the analysis, a total of 105 datasets were included: 75 of euploid fetuses (median gestational age 24, range 15-32 weeks) and 30 of DS fetuses (median gestational age 24, range 17-34 weeks). Originally, 32 datasets of DS fetuses were available, but two were excluded, as the 2D images were judged not to be truly midsagittal.

Mean values, mean differences between 2D and 3D measurements, and corresponding ICC of 2D and 3D measurements in a combined cohort of euploid and DS fetuses (N=105) are shown in Table 1. It was not possible to analyze mean differences between measurements of the FP line as the outcome was noncontinuous (positive or zero; no FP line was negative).

For the MNM angle, 2D measurements were significantly larger (p < 0.01), although the mean difference was small (1.0 degree). For the other markers (NBL, PT, FP line, PT-NBL ratio, and PFSR), there was no significant difference in measurements performed in either 2D or 3D US.

Intraobserver and interobserver variability in 2D and 3D US for each marker (except
Table 2 – Mean values for 30 datasets of Down syndrome fetuses with their corresponding detection rate.

<table>
<thead>
<tr>
<th></th>
<th>2D Mean</th>
<th>3D Mean</th>
<th>Detection rate 2D(%)</th>
<th>Detection rate 3D(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBL (mm)</td>
<td>4.37</td>
<td>3.91</td>
<td>83.3</td>
<td>82.6</td>
</tr>
<tr>
<td>FP Line</td>
<td>-</td>
<td>-</td>
<td>37.0</td>
<td>41.4</td>
</tr>
<tr>
<td>MNM (degrees)</td>
<td>14.35</td>
<td>14.07</td>
<td>3.7</td>
<td>10</td>
</tr>
<tr>
<td>PT (mm)</td>
<td>6.31</td>
<td>6.15</td>
<td>83.3</td>
<td>86.7</td>
</tr>
<tr>
<td>PT-NBL ratio</td>
<td>1.40</td>
<td>1.55</td>
<td>96.7</td>
<td>100</td>
</tr>
<tr>
<td>FFSR</td>
<td>0.60</td>
<td>0.51</td>
<td>50</td>
<td>56.7</td>
</tr>
</tbody>
</table>

No statistical significant differences were observed between 2D and 3D measurements. MNM = maxilla-nasion-mandible angle. PT = prenasal thickness. PT-NBL ratio = prenasal thickness to nasal bone length ratio. FFSR = prefrontal space ratio.

Figure 4 – Nasal bone length (NBL) measurements performed on 2D images and 3D volumes in euploid and Down syndrome (DS) fetuses, plotted on normal ranges (mean, 5th centile and 95th centile).

for the FP line), with their corresponding limits of agreement (LOA) and 95% confidence interval, is shown in Figure 2. LOAs were smaller for all 3D measurements, except for the MNM angle.

In the separate analysis of DS fetuses only, no statistical difference was found for any marker between measurements performed in images acquired by either 2D or 3D US in
their ability to discriminate between normal and DS fetuses (Table 2, Figures 3-8).

Discussion

This study demonstrates that when strict criteria are applied, subjective judgment of a good midsagittal plane on 2D images is sufficient to ensure a good performance of facial markers for DS. In a cohort of euploid and DS fetuses, no significant difference was found in NBL, PT, FP line, PT-NBL ratio, and PFSR, measured in midsagittal images obtained by 2D or 3D US. Only for the MNM angle, 2D measurements were slightly, but significantly, larger. Both 2D and 3D technique can perform equally well in identifying DS fetuses, without significant difference between measurements.

The clinical implication of these findings is that these markers can be used effectively in routine screening settings relying on 2D technique using strict criteria, without missing out on the additional benefit of 3D US. This finding has great implications in a moment of worldwide financial constraints and growing medico-legal problems, where the general opinion is that 3D US is superior to 2D US.\[13,18–21\]

In literature, the role of 3D technique in the measurement of fetal facial biometrical parameters has been underlined by many studies. A volume obtained starting from an oblique scanning plane can be corrected to the exact midsagittal plane, allowing accurate and reproducible measurements. Moreover, a stored 3D volume can be analyzed offline retrospectively, possibly shortening the time of investigation.

Suggested disadvantages of using 3D volume corrections are, that it requires costly
Figure 6 – Nasal bone length (NBL) measurements performed on 2D images and 3D volumes in euploid and Down syndrome (DS) fetuses, plotted on normal ranges (mean, 5th centile and 95th centile).

Figure 7 – Nasal bone length (NBL) measurements performed on 2D images and 3D volumes in euploid and Down syndrome (DS) fetuses, plotted on normal ranges (mean, 5th centile and 95th centile).
equipment, specialized personnel, and it may be more time-consuming.\textsuperscript{21} However, other studies found no difference in time\textsuperscript{13,24} or found 3D to be even faster.\textsuperscript{25-27}

Following our previous report,\textsuperscript{13} this is the first study that evaluates the use of 2D versus 3D acquired images in the evaluation of profile markers.

Benoit and Chaoui\textsuperscript{21} demonstrated that, in case of suspicion of an absent nasal bone on 2D images, the nasal bone can be better visualized in 3D volumes. Persico et al.\textsuperscript{9} showed that 3D NBL measurements tend to be larger when the scanning plane is not exactly mid-sagittal, which decreased the detection rate for DS. In our previous study,\textsuperscript{13} we found no difference between measurement modality but reported narrower LOA in 3D performed measurements.

A possible criticism of this study is that for comparison, we performed a selection of 2D pictures likely to represent the true midsagittal plane. All 2D US measurements are taken on planes subjectively judged as correct according to anatomical landmarks. However, in a previous study,\textsuperscript{11} we showed that after volume acquisition, even when the image on the A plane was subjectively judged to be midsagittal, variable degrees of correction by multiplanar mode were required in order to obtain the true midsagittal plane. Based on the results of the present study, the measurements of facial markers in the initial 2D image are highly comparable with those measured in 3D corrected planes.

Limitations of this study are its retrospective nature and the fact that examiners were not blinded to the karyotype. An ideal repeatability and reproducibility study would require reacquisition of the same images by two observers. Because of the retrospective nature of this study, this was not possible; however, reproducibility of the markers is established in the

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure8.png}
\caption{Nasal bone length (NBL) measurements performed on 2D images and 3D volumes in euploid and Down syndrome (DS) fetuses, plotted on normal ranges (mean, 5th centile and 95th centile).}
\end{figure}
original publications.\textsuperscript{6,14,22} Influence of circumstances such as reduced amniotic fluid or the fetus facing down was not taken into account; however, these circumstances would equally affect 2D and 3D performance.\textsuperscript{13,28}

We expressed differences in measurements in ICC, as the markers are quantified by different metric parameters (degrees, mm, and ratios). The MNM angle, FP line and PFSR had a relatively low ICC when 2D and 3D measurements were compared. One explanation may be that the aforementioned markers, in contrast to the PT and NBL, require multiple landmarks that may enhance the variability in the measurement.

For the MNM angle, 2D performed measurements were significantly larger, the ICC of 2D versus 3D measurements was low, and LOAs of intraobserver and interobserver variability were relatively wide. This could be because the reproducibility of the MNM angle is in general more challenging and that, especially in this case, the bony structures used as a landmark for the measurement are better identified by 3D US.

Conversely, the PFSR has a suboptimal ICC when 2D and 3D measurements are compared, but the mean difference between 2D and 3D measurements is not significant. The LOAs of the PFSR are larger when compared with that of the PT-NBL ratio (also expressed as a ratio), especially in 2D. These findings show that the reproducibility of the PFSR (in 2D) is lower, but the actual measurement is not influenced by the technique of image acquisition.

In spite of these findings, no impact of acquisition modality could be found in detection rate and measurements performed in DS fetuses. This is reassuring, as the goal of these measurements is to discriminate between DS and euploid cases.

The best moment for DS screening is undisputable the late first trimester. However, uptake of first trimester screening varies among countries. In the Netherlands, for instance, where the combined test is only covered by insurance beyond 36 years of age, the uptake is low.\textsuperscript{17} In contrast, uptake of second trimester US screening for structural anomalies, which is covered for all women,\textsuperscript{29} is over 95%.\textsuperscript{30} This means that a considerable number of DS pregnancies remain undetected. Also when late termination of pregnancy is not available, a late diagnosis of a chromosomal anomaly is important to prepare future parents and establish the appropriate obstetrical management.

In conclusion, we have shown that, with exception of larger MNM angles in 2D images, no significant differences were found between 2D and 3D imaging in a number of facial DS markers. In particular, NBL, PT, FP line, PT-NBL ratio, and PFSR can be confidently used as DS markers in US examinations performed by 2D US, provided the markers are measured in a midsagittal image, acquired according to strict criteria.

References

22. Vos F. I., De Jong-Pleij E. A., Ribbert L. S. et al. Three-dimensional ultrasound imaging and measurement of nasal bone length, prenasal thickness and frontomaxillary facial


