Ultrasonography of the fetal nose, maxilla, mandible and forehead as markers for aneuploidy

Vos, Fedia

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Nasal bone length, prenasal thickness, prenasal thickness-to-nasal bone length ratio and prefrontal space ratio in second- and third-trimester fetuses with Down syndrome

Vos FI, De Jong-Pleij EA, Bakker M, Tromp E, Pajkrt E, Kagan KO, Bilardo CM

ABSTRACT

Objectives
To evaluate nasal bone length (NBL), prenasal thickness (PT), prenasal thickness-to-nasal bone length (PT-NBL) ratio and prefrontal space ratio (PFSR) as markers for Down syndrome (DS) in the second and third trimesters.

Methods
NBL, PT, PT-NBL ratio and PFSR were measured retrospectively in stored two-dimensional images or three-dimensional volumes (corrected to the mid-sagittal plane) of fetuses with Down syndrome, which were retrieved from the digital databases of participating units. Measurements were performed on the stored images and volumes by two experienced operators, and the values obtained were compared to our previously reported normal ranges for euploid fetuses in order to assess the detection rates for Down syndrome.

Results
A total of 159 fetuses with DS were included in the analysis, six of which were excluded because of inadequate available images. Median maternal age was 36.0 years and median gestational age 23 + 1 weeks. NBL and PT were correlated with gestational age \((P < 0.001)\), but the PT-NBL ratio and PFSR were not. Mean NBL, PT, PT-NBL ratio and PFSR were 4.42 mm, 5.56 mm, 1.26 and 0.34, respectively. The nasal bone was absent in 23 (15.4%) cases. As a marker for Down syndrome, the PT-NBL ratio yielded the highest detection rate (86.2%), followed by PFSR (79.7%), PT (63.4%) and NBL (61.9%). All markers were abnormal in 33.6% of cases, whilst all were normal in 4.7%. At least one of the four markers was abnormal in 95.3%, and either the PT-NBL ratio or PFSR was abnormal in 93.8%. Detection rates were not related to gestational age.

Conclusions
The PT-NBL ratio and PFSR are robust second- and third-trimester markers for Down syndrome. Both provide high detection rates and are easy to use, as the cut-off for normality is constant throughout gestation.
INTRODUCTION

In 1866, Langdon Down first described the typical facial features of those affected by the syndrome that received his name: a flat profile, small nose and redundant skin. These typical features, detectable in the profile of fetuses with Down syndrome, are currently used as ultrasound markers for this condition in first- and second-trimester screening. These profile markers are based on the fact that fetuses with Down syndrome (DS) are characterized by different degrees of mid-facial hypoplasia and skin edema. Hypoplastic or absent nasal bones, reduced convexity of the fetal profile and thickened skin in the nuchal and prefrontal areas have been confirmed in fetuses with DS by postnatal pathological reports and X-ray imaging. Prenatally, these features can be quantified by measuring them as fetal profile parameters. Screening for DS usually is performed in the first trimester of pregnancy, during which nuchal translucency thickness and visualization of the nasal bone are sensitive ultrasound markers. However, in cases in which first-trimester screening is not performed, it is important to define effective second- and third-trimester markers. Fetal profile or neck markers currently are considered by far the most predictive in comparison with other ultrasound markers for Down syndrome, e.g. short femur and humerus, echogenic cardiac focus and hyperechogenic bowel.

Bony markers such as nasal bone length (NBL) and the frontomaxillary facial angle allow assessment of mid-facial hypoplasia, while prenasal thickness (PT) and nuchal fold measurements allow assessment of skin thickness. More recently, combined markers have been proposed, such as the prenasal thickness-to-nasal bone length (PT-NBL) ratio and the prefrontal space ratio (PFSR), which appear to be superior to single markers. The theoretical advantage of the PT-NBL ratio and PFSR is that the measurements used to generate these are affected by both thickness of the skin of the forehead and the respective degrees of nasal and mid-facial hypoplasia. Normal reference ranges for these combined markers have also been assessed.

Despite many reports in the literature concerning single or multiple markers, measured in cohorts of various sizes, a large and comprehensive study in which all profile markers are measured in the same fetus was lacking. The aim of this retrospective study was to assess the performance and interrelation of four fetal profile markers for DS (NBL, PT, PT-NBL ratio and PFSR) measured in the same fetus on images acquired at second- and third-trimester ultrasound examination.

METHODS

Ultrasound records for the study were retrieved from the digital databases of the Fetal Medicine Unit of the following centers: University Medical Centre, Groningen, The Netherlands; Academic Medical Centre, Amsterdam, The Netherlands (until March 2010); University Medical Centre, Utrecht, The Netherlands and the Department of Obstetrics and Gynecology, University Hospital Tübingen, Tübingen, Germany. A search was undertaken for second- and third-trimester midsagittal two-dimensional (2D) ultrasound images or three-dimensional (3D) stored volumes of the profile of fetuses with DS seen between January 2006 and July 2013 at one of the participating institutions.
The diagnosis was confirmed in all cases by prenatal or postnatal karyotyping. Only fetuses of Caucasian parents were included. Ultrasound examinations were performed using GE Voluson 730 Expert ultrasound or E8 systems equipped with a RAB4-8L probe (GE Medical Systems, Kretz Ultrasound, Zipf, Austria). Images and volumes were stored and examined either offline on 4D View software version 7.0 (GE Medical Systems) or on stored images in the GE ultrasound system.

Only good mid-sagittal images of the fetal profile were selected and considered for further analysis; we required that images show the forehead, nose, lips and chin, with the maxilla visible as a single horizontal line without appearance of zygomatic bone. Images with a visible zygomatic bone or ramus of the mandible were excluded. To perform the measurement, stored 2D ultrasound images of the fetal profile were magnified to fill the entire monitor. In cases for which 3D volumes were available, the multiplanar mode was used to depict the exact median plane of the fetal profile.

The NBL was measured from the nasion to the end of the white distal ossification line (B in Figure 1). The nasion was defined as the most anterior point at the junction between the frontal and nasal bones. Care was taken to avoid adding part of the frontal bone to the measurement. In cases in which there was a gap between the nasal and the frontal bones (disjunction), NBL was measured from the distal to the proximal end of the ossification line. To determine the PFSR, the maxilla-mandible line (MM line) was drawn between the midpoint of the anterior edge of the mandible and the anterior edge of the maxilla. The line was then extended cranially towards the forehead. Subsequently, the skin covering the forehead (C in Figure 1) was measured between the anterior edge of the bony forehead and the anterior edge of the skin in a line parallel to the maxilla and traced from the nasion; this measurement is called prenasal thickness (PT). A second measurement (d) (D in Figure 1), was taken from the anterior edge of the skin (where PT measurement ended) to the point of intercept with the MM line. For cases in which the MM line crossed the prenasal skin posteriorly, PT was measured between the frontal bone and the skin, but d was measured between the MM line and the skin and multiplied by -1. The PFSR was determined by dividing d by PT, and the PT-NBL ratio was calculated by dividing PT by NBL.

Figure 1 | Ultrasound image of a fetus with Down syndrome at 21+3 weeks’ gestation, showing the maxilla-mandible line (A), nasal bone length (B), prenasal thickness (C) and the ‘d’ measurement (D). The prefrontal space ratio was calculated by dividing D by C, and the prenasal thickness-to-nasal bone length ratio was calculated by dividing C by B.
Reproducibility of the fetal profile measurements was assessed in all cases, using stored images (or volumes when available). 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. Images were chosen at random at different gestational ages, with at least 3 weeks between the two assessments when performed by the same examiner. Only a single measurement was used for the analysis relating to detection rates for DS. Data were compared to reference values derived from our previous studies on euploid fetuses\textsuperscript{11,12,14}, which found that NBL and PT increased with gestation from 3.3 mm at 15 weeks to 9.6 mm at 33 weeks (NBL\(= -6.927 + (0.83 \times \text{GA}) - (0.01 \times \text{GA}^2)\)) and from 2.3 mm at 15 weeks to 6.1 mm at 33 weeks (PT = 0.212 \times \text{GA} - 0.873), respectively. The PT-NBL ratio and PFSR were stable throughout gestation, with a mean of 0.61 (95\textsuperscript{th} percentile, 0.80) and 0.97 (5\textsuperscript{th} percentile, 0.55), respectively. Measurements of NBL and PFSR below the 5\textsuperscript{th} percentile were considered abnormal, and for PT and PT-NBL ratio values above the 95\textsuperscript{th} percentile were considered abnormal. Multiple of the median (MoM) values were calculated for PT and NBL to correct for gestational age and was considered to be below the 5\textsuperscript{th} percentile.

Data were analyzed using the statistical software SPSS version 21.0 for Windows (SPSS Inc., Chicago, IL, USA) and Excel for Windows 2000. Means with ranges or SD were calculated when appropriate. Correlation was determined by Pearson’s correlation test. A \(P\)-value of less than 0.05 was considered statistically significant. Intra- and interobserver variability was assessed by Bland-Altman analysis and intraclass correlation coefficient (ICC).

RESULTS

Images from a total of 159 fetuses with DS were available for analysis, including 33 3D volumes and 126 2D images. Median maternal age was 36.0 (range, 23.1 – 49.5) years and median gestational age was 23 + 1 (range 14 – 38) weeks. We excluded six images from further analysis, five that were not completely mid-sagittal and one that had an unclear nasion. In an additional 22 cases, with only 2D images available, PFSR was not measured because either the mandible or the maxilla was not displayed clearly enough to allow accurate measurement. In 3 cases gestational age was not known. It was possible to obtain all four measurements in 128 cases. Results relating to intra- and interobserver variability in the measurements, assessed by means of Bland-Altman analysis and ICCs, are reported in Table 1. For both intra- and interobserver analysis, ICC values > 0.9 were found for NBL, PT and the PT-NBL ratio, and a lower value of 0.67 was found for the PFSR.

Measurements of NBL and PT showed a correlation with gestational age (\(r = 0.69; \ p < 0.001\) and \(r = 0.74; \ p < 0.001\), respectively), but the PT-NBL ratio and PFSR did not. The mean (± SD) values of the NBL, PT, PT-NBL ratio and PFSR were 4.42 ± 2.39 mm, 5.56 ± 1.98 mm, 1.26 ± 0.58 and 0.34 ± 0.31, respectively. The nasal bone was absent in 23 (15.4\%) cases. As an absent of nasal bone was significantly more common earlier in pregnancy, as it was negatively correlated with gestational age (\(p < 0.01\)).
In 43 of 128 (33.6%) DS cases, all markers were abnormal, whilst in 6 of 128 (4.7%) cases all markers were normal. The detection rate, false-positive rate, positive likelihood ratio and negative likelihood ratio of each marker are given in Table 2.

Among the markers the PT-NBL ratio and PFSR yielded the highest detection rates for DS, of 86.2% and 79.7%, respectively. Measurements of the markers in fetuses with DS are plotted against the normal ranges for euploid fetuses\textsuperscript{11,12,14} in Figure 2.

Figure 2 | Measurements of facial profile markers on stored prenatal ultrasound images of fetuses with DS plotted against gestational age and in comparison to the mean, 5th and 95 percentiles as previously reported in euploid fetuses\textsuperscript{11,12,14}. Nasal bone length (NBL) (n = 145), prenasal thickness (PT) (n = 145), PT-NBL ratio (145) and prefrontal space ratio (PFSR) (n = 133) fetuses with DS. Fetuses with an absent nasal bone are plotted with an NBL value of zero and a PT-NBL ratio of 4.5. PT and/or NBL measurements could not be obtained in all cases.

At least one of the four markers was abnormal in 95.3% of cases. Abnormality of the PFSR and/or PT-NBL ratio yielded a detection rate of 93.8%. Each individual marker appeared to be equally effective in screening for DS across gestation, since there was no significant correlation between gestational age and the MoM values of NBL and PT or the detection rate of PFSR and PT-NBL ratio.
There was no significant correlation between PFSR and NBL MoM ($p = 0.104$). All other MoM values of individual markers were significantly correlated with each other ($p < 0.01$); significance was lowest for PFSR and PT MoM ($p = 0.045$).

**Table 1 | Intra- and interobserver variability in measurements of facial profile markers on stored ultrasound images and volumes of fetuses with DS. ICC, intraclass correlation coefficient; LOA, limits of agreement; NBL, nasal bone length; PT, prenasal thickness; PFSR, prefrontal space ratio.**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Mean difference (SD)</th>
<th>LOA</th>
<th>ICC</th>
<th>Mean difference (SD)</th>
<th>LOA</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBL</td>
<td>-0.14 (0.40)</td>
<td>-0.93 – 0.65</td>
<td>0.98</td>
<td>0.28 (0.77)</td>
<td>-1.26 – 0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>PT</td>
<td>-0.01 (0.45)</td>
<td>-0.90 – 0.89</td>
<td>0.98</td>
<td>0.08 (0.49)</td>
<td>-0.89 – 1.05</td>
<td>0.97</td>
</tr>
<tr>
<td>PT-NBL ratio</td>
<td>0.04 (0.15)</td>
<td>-0.26 – 0.34</td>
<td>0.94</td>
<td>-0.01 (0.19)</td>
<td>-0.39 – 0.38</td>
<td>0.92</td>
</tr>
<tr>
<td>PFSR</td>
<td>-0.06 (0.27)</td>
<td>-0.60 – 0.48</td>
<td>0.67</td>
<td>-0.17 (0.25)</td>
<td>-0.67 – 0.33</td>
<td>0.67</td>
</tr>
</tbody>
</table>

**Table 2 | Performance of nasal bone length (NBL), prenasal thickness (PT), prenasal thickness to nasal bone length (PT-NBL) ratio and prefrontal space ratio (PFSR) as markers for DS. 95% CI’s are given in parentheses. DR; detection rate, FPR; false positive rate, PLR; positive likelihood ratio, NLR; negative likelihood ratio.**

<table>
<thead>
<tr>
<th>Marker</th>
<th>DR (95% CI)</th>
<th>FPR (95% CI)</th>
<th>PLR (95% CI)</th>
<th>NLR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBL (mm)</td>
<td>61.9%</td>
<td>5.0%</td>
<td>12.32</td>
<td>0.40</td>
</tr>
<tr>
<td>(n = 145)</td>
<td>(53.4% – 69.9%)</td>
<td>(1.7% – 11.3%)</td>
<td>(5.17 – 29.37)</td>
<td>(0.31 – 0.52)</td>
</tr>
<tr>
<td>PT (mm)</td>
<td>63.4%</td>
<td>5.0%</td>
<td>12.73</td>
<td>0.38</td>
</tr>
<tr>
<td>(n = 145)</td>
<td>(53.4% – 73.1%)</td>
<td>(1.7% – 11.3%)</td>
<td>(5.35 – 30.29)</td>
<td>(0.29 – 0.50)</td>
</tr>
<tr>
<td>PT-NBL ratio</td>
<td>86.2%</td>
<td>5.0%</td>
<td>17.37</td>
<td>0.14</td>
</tr>
<tr>
<td>(n = 145)</td>
<td>(79.3% – 91.2%)</td>
<td>(1.7% – 11.3%)</td>
<td>(7.4 – 41.0)</td>
<td>(0.08 – 0.23)</td>
</tr>
<tr>
<td>PFSR</td>
<td>79.7%</td>
<td>5.0%</td>
<td>15.96</td>
<td>0.21</td>
</tr>
<tr>
<td>(n = 133)</td>
<td>(71.6% – 86.0%)</td>
<td>(1.7% – 11.3%)</td>
<td>(6.75 – 37.72)</td>
<td>(0.14 – 0.32)</td>
</tr>
</tbody>
</table>

**DISCUSSION**

This study involves the largest cohort of fetuses with DS thus far, in which all known profile markers have been measured and their detection rate has been established after comparison with normal ranges established by the same study group. The study confirms that screening for DS can be performed effectively in the second and third trimester of pregnancy. The best markers are the PT-NBL ratio and the PSFR, with predicted detection rates of 86% and 80%, respectively. The detection rate further increases to 94% when the PT-NBL ratio and the PFSR are combined, and slightly more (95%) when all facial markers (NBL, PT, PT-NBL ratio and PFSR) are combined. An additional
advantage of using the combined markers (PT-NBL ratio and PFSR) in routine examination is that the 5th (PFSR) and 95th (PT-NBL ratio) percentile cut-offs are constant throughout gestation at 0.55 and 0.80, respectively.

Interest regarding the facial features of individuals with DS dates back to the late 1970’s when the cephalic index was proposed as the first ultrasound screening method for DS\(^{15}\). Sonek et al\(^{16}\) were the first to observe the absence of nasal bones as a marker for DS in 2001, whilst the markers PT, PT-NBL ratio and PFSR have been introduced more recently\(^{13,17}\). It has been proposed that in DS, changes in the extracellular matrix in the skin and abnormalities of lymphatic vessels lead to a variable increase in skin thickness in the neck and prenasal region\(^{6,18}\). Abnormalities in bone growth and development are associated with mid-facial hypoplasia\(^{2,19}\), resulting in an abnormal profile and small nasal bones. This study investigated the efficacy of various methods of quantification of these abnormalities.

A limitation of this study is its retrospective nature and the fact that examiners were not blinded to the karyotype. However, its strength is to have assessed the value of second-trimester ultrasound markers in a large cohort of fetuses with DS. As expected, the previously reported 100% detection rate for both the PT-NBL ratio\(^{12}\) and PFSR\(^{20}\) decreases when the method is applied to a large cohort. However, the combination of both markers leads to a high detection rate (94%), thus far the highest reported in a large study using an algorithm based exclusively on ultrasound measurements.

The PT measurement is part of the PFSR and is referred to as ‘d1’ in the PFSR studies of Sonek et al\(^{13}\), Yazdi et al\(^{14}\) and Chaveeva et al\(^{20}\). However, whereas Sonek et al\(^{13}\) and Yazdi et al\(^{14}\) calculate d1 as the distance between skull and skin in a line parallel to the maxilla, Chaveeva et al\(^{20}\) measure it perpendicular to the MM line. In this study we followed the first method, as we suspect that the position of the MM line would be reflected in the length of d1 (PT) when measured perpendicular to it.

Concerning the interdependency of the NBL MoM, PT MoM, PT-NBL ratio and PFSR, we were not surprised to find the PT-NBL ratio significantly correlated to all other markers. The significant, but weaker, correlation (P = 0.045) between PT MoM and PFSR suggests that the measurement of d, which represents mid-facial hypoplasia, is independent of PT. Similar to other studies\(^{13,20}\), the NBL MoM and PFSR were the only markers that were not significantly correlated in this large cohort of fetuses with DS. However, the combination of these two independent markers did not yield a better detection rate than did the combination of the PFSR and the PT-NBL ratio. Ideally, an adequate repeatability and reproducibility study should be performed, not only by remeasuring ultrasound markers on stored pictures, but also by reacquiring the desired image. Due to the retrospective nature of this study, the latter was not possible, and our reproducibility figures therefore relate exclusively to reproducibility of the measurement. Reproducibility was good for all markers, with the exception of the PFSR. An explanation may be found in the fact that when using a marker combining multiple measurements performed in the mid-sagittal plane, such as the PFSR, the slight interobserver variation in each measurement is amplified. Proof of this may be the lower interobserver ICC for PFSR of 0.67, compared to an ICC of 0.98 for PT. In comparison with other studies on the PFSR, reproducibility of our measurements is poorer than that reported by Chaveeva et al\(^{20}\) but of the same order as that reported by Sonek et al\(^{13}\) and Yazdi et al\(^{14}\).
A number of images (n = 6) and PFSR measurements (n = 22) were excluded from the study either because they were not in a midsagittal view or because of unclear mandible and/or maxilla. These results suggest that when an ultrasound image is obtained with the intention of measuring the nasal bone and prenasal skin thickness, less care is taken in obtaining good visualization of the bony landmarks of the maxilla and mandible. This may not be the case when measurements are taken prospectively, with special attention given to visualization of the bony landmarks of the profile. This assumption is further substantiated by the fact that all discarded ultrasound records were 2D images. If these images had not been excluded then we may have found lower detection rates. Therefore, no firm statements can be made concerning the true detection and false-positive rates until a prospective study is performed.

Studies on DS screening in the second and third trimester are relatively scarce as in countries with well-established DS screening programs, screening occurs preferably in the first trimester (90% in Denmark and France\textsuperscript{21,22}). However, well-organized and established first-trimester screening programs are not available in all countries and screening uptake can also be low (20% and 32% in certain areas of England and The Netherlands\textsuperscript{23,24}). This means that, whilst non-invasive prenatal testing (NIPT) could potentially replace first-trimester screening, there remains at present a role for the evaluation of DS markers at second-trimester ultrasound examination. A late diagnosis is obviously less desirable when the parents may consider termination of pregnancy. However, in all other cases, a diagnosis may still be important for pregnancy management and for preparation of the birth of an affected child.

In conclusion, according to this large cohort of retrospectively analyzed fetuses with DS, the PT-NBL ratio and PFSR qualify as excellent second-trimester ultrasound markers. The strength of the PT-NBL ratio is that it provides a high detection rate and that it is reproducible. Both markers are easy to use in practice, as no knowledge of gestational age-specific mean values is required.
REFERENCES


