Intra- and interoperator variability in fetal nasal bone assessment at 11–14 weeks of gestation

M. V. SENAT*, J. P. BERNARD*, M. BOULVAIN† and Y. VILLE*

*Department of Obstetrics and Gynecology, Paris-Ouest University, CHI Poissy-St-Germain, France and †Department of Obstetrics and Gynecology, Hôpitaux Universitaires de Genève, Geneva, Switzerland

KEYWORDS: fetal nasal bones; first trimester; reproducibility; trisomy 21; ultrasound screening

ABSTRACT

Objectives Examination of the fetal nasal bones by ultrasound between 11 and 14 weeks of gestation has been proposed as an additional tool in the detection of trisomy 21 in a high-risk population. However the variability in the identification of fetal nasal bones by ultrasound has not yet been investigated. The aim of this study was to assess the intraoperator and interoperator reproducibility of fetal nasal bone identification by ultrasound between 11 and 14 weeks of gestation.

Methods A total of 1040 consecutive ultrasound examinations were performed at 11–14 weeks of gestation for nuchal translucency (NT) measurement and nasal bone identification by ultrasound. A total of 657 consecutive video-loops were assessed by three experienced operators. Each operator assigned cases to one of three categories, namely present, uncertain or absent nasal bones, and the results were compared between operators. To assess the intraoperator variability each operator reviewed 100 randomly selected videos out of the 657 loops and again used the same classification. Results were compared by pairwise unweighted and weighted Kappa index to evaluate the inter- and intraoperator variability.

Results Among the 1040 fetuses, there were 51 (4.9%) with an NT measurement above the 95th centile. Nasal bones were identified by ultrasound in 948, not seen in eight and impossible to assess in 84 fetuses. Four fetuses had trisomy 21 including three with absent nasal bones and increased NT and one with present nasal bones and normal NT. The Kappa and weighted Kappa values for interoperator variability between the three operators were between 0.26 and 0.37 and 0.33 and 0.44, respectively. The Kappa and weighted Kappa values for intraoperator variability were between 0.35 and 0.48 and 0.43 and 0.53, respectively.

Conclusion The assessment of fetal nasal bones is only fairly reproducible. Although the performance of the test in fetuses at high risk for trisomy 21 has been reported to be good, its implementation as an additional screening technique in the general population must be accompanied by teaching and quality control programs. Copyright © 2003 ISUOG. Published by John Wiley & Sons, Ltd.

INTRODUCTION

Nuchal translucency (NT) measurement is an established screening method for Down syndrome and other chromosomal abnormalities in the first trimester of pregnancy1–4. Risk calculation taking into account maternal age, fetal NT and maternal serum biochemistry at 11–14 weeks of gestation has a sensitivity of up to 85% for a false-positive rate of around 5%4–9.

More recently, ultrasound examination of the fetal nasal bones between 11 and 14 weeks of gestation has been proposed as an additional screening tool with both a reduction in the false-positive rate and an increased detection rate in fetuses at high risk of trisomy 21 by NT measurement10. The prevalence of absent nasal bones by ultrasound and the reproducibility of fetal nasal bone assessment between 11 and 14 weeks of gestation have not been evaluated previously. However this is a critical step should this be applied to the general population as it affects both sensitivity and specificity of the test. Reproducibility of nasal bone assessment was therefore studied at 11–14 weeks of gestation.

METHODS

A total of 1040 consecutive ultrasound examinations were performed at 11–14 weeks’ gestation in an unselected...
Reproducibility of fetal nasal bone assessment

population for NT measurement and nasal bone identification between January 2002 and August 2002. Each pregnancy was examined for number of fetuses and measurement of crown–rump length (CRL), biparietal diameter, femur length and NT. The fetal profile was examined to identify the fetal nasal bones in a mid-sagittal view with the ultrasound beam directed perpendicularly to the nasal bones. The image was magnified for the fetal profile to occupy up to two-thirds of the screen and the transducer was tilted from side to side to ensure that either or both nasal bones could be seen separate from the nasal skin. The examinations of the last 567 consecutive fetuses were recorded as video loops of 30 s duration in order to assess the interobserver reproducibility. The first 473 examinations were recorded as still images and were thus deemed unsuitable for the reproducibility study but they were used for the prevalence calculations. During the screening examination performed by one of the three operators (J.P.B.), fetuses were assessed to classify them in one of the following groups: presence of the nasal bones (Group 1), absence of the nasal bones (Group 2) or the impossibility to confirm either presence or absence (Group 3). There were 530, 7 and 30 fetuses in Groups 1, 2 and 3, respectively. Nine videos were available for the seven fetuses with absent nasal bones and 50 for the 30 fetuses with uncertain assessment. Cases in Group 2 were replicated three times to a total of 27 cases and cases in Group 3 were duplicated to a total of 50 (30 fetuses)

The results of the three operators were compared pairwise and concordance within each of the three pairs was calculated. Inter- and intraoperator variability was calculated with the Kappa index for each pair of operators (i.e. Pair 1, Pair 2, Pair 3). Kappa is defined as the difference between observed and expected agreement expressed as a fraction of the maximum difference. The maximum value of Kappa is 1, which represents perfect agreement and a Kappa value of 0 indicates agreement is only due to chance.

According to Fleiss, values of Kappa exceeding 0.75, values between 0.4 and 0.75 and values of less than 0.4 represent excellent, fair to good and poor agreement, respectively. We considered ‘presence/absence’ to be a worse disagreement than absent/uncertain or present/uncertain. In order to account for this difference in degree of disagreement, we also calculated the weighted Kappa for each pair of operators.

**RESULTS**

Among the 1040 examinations the median (range) maternal age and CRL were 29 (18–44) years and 64 (48–84) mm, respectively. The median NT was 0.1 mm below The Fetal Medicine Foundation’s expected value. There were 51 cases (4.9%) with NT above the 95th centile. In the initial screening examination, nasal bones were identified by ultrasound in 948 (91.1%), not seen in eight (0.8%) and impossible to assess in 84 (8.1%) fetuses, respectively. Four fetuses had trisomy 21 including three with absent nasal bones and increased NT and one with present nasal bones and normal NT.

The distribution of present/uncertain/absent nasal bones among the three operators are shown in Table 2 and Figure 1. Summing up the three operators’ results of 657 cases (n = 1971) nasal bones were present, uncertain and absent in 78.8%, 14.8% and 6.4%, respectively. In 1590/1971 pairs there was good agreement (80.7%), the disagreement was between absent/present in 35 pairs (1.8%), between present/uncertain in 277 pairs (14%) and between absent/uncertain in 69 pairs (3.5%). The Kappa value for nasal bone assessment by one operator compared to each of the other two operators was between

<table>
<thead>
<tr>
<th>Table 2 The distribution of nasal bone assessment by ultrasound among the three observers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operator</strong></td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

*These figures were obtained as a result of the replication three times of the nine (seven fetuses) and duplication of 50 (30 fetuses) video-loops in Groups 2 and 3, respectively.

Copyright © 2003 ISUOG. Published by John Wiley & Sons, Ltd.
0.26 and 0.37 (Table 3). The Kappa value for the intraoperator reproducibility was between 0.35 and 0.48 (Table 4).

**DISCUSSION**

Absence of fetal nasal bones as assessed by ultrasound at 11–14 weeks of gestation may be the consequence of delayed ossification or hypoplasia of nasal bones. It has been suggested that these are more frequent in fetuses with trisomy 21 and that their presence can be assessed by ultrasound both in the first and second trimester of pregnancy. A valuable screening test should be both feasible and reproducible, in addition to demonstrating good sensitivity and specificity, criteria that should be evaluated before its introduction in routine settings. False-positive results in screening for trisomy 21 in the general population have a major impact on the positive predictive value, as the prevalence of the disease is less than 1%.

This study shows that assessment of fetal nasal bones by ultrasound at 11–14 weeks of gestation is only fairly reproducible with good agreement in only 80.7% of the cases and Kappa values around 0.4. Indeed, the impact of a disagreement on risk calculation is maximal when one operator assigns nasal bones as present and another one as absent (likelihood ratio (LR) = 146 vs. 0.27). The difference in LR is less extreme for absent/uncertain disagreement (LR = 1 vs. 146), but still significant. Finally, the difference in LR calculation for present/uncertain disagreement was considered to be moderate (LR = 1 vs. 0.27).

Although NT measurement is widely used and qualifies as a reproducible technique for trained operators, some authors dispute its feasibility and reproducibility in an unselected population. If nasal bone assessment is added to maternal age, NT and maternal serum biochemistry then this is likely to have an impact on risk calculation by increasing it by up to 146 times when nasal bones are not seen and decreasing this risk (LR = 0.27) when they can be identified. In our study we found 1.8%, 3.5% and 14% disagreement between absent/present, absent/uncertain and present/uncertain assignment, respectively, which significantly changes the risk calculation from one operator to another.

Although the performance of the test in fetuses at high risk for trisomy 21 has been reported to be good, its implementation as an additional screening technique in the general population must be accompanied by teaching programs, standardization of the technique and quality control. Both large observational studies in unselected populations and the assessment of the independence of nasal bones and NT as markers for chromosomal abnormalities are needed.

**REFERENCES**

Reproducibility of fetal nasal bone assessment


