Ultrasound Obstet. Gynecol. 5 (1995) 334–337

Repeatability of measurement of fetal nuchal translucency thickness

P. P. Pandya, D. G. Altman, M. L. Brizot, H. Pettersen and K. H. Nicolaides*

Harris Birthright Research Centre for Fetal Medicine, King's College Hospital Medical School; *Medical Statistics Laboratory, Imperial Cancer Research Fund, London, UK

Key words: NUCHAL TRANSLUCENCY THICKNESS, REPEATABILITY, INTRAOBSERVER, INTEROBSERVER, CALIPER PLACEMENT

ABSTRACT

The aim of this prospective study was to assess the repeatability of measurement of fetal nuchal translucency thickness at 10–14 weeks' gestation. The nuchal translucency was measured by two of four operators in 200 pregnant women attending the Harris Birthright Research Centre for Fetal Medicine at 10–14 weeks' gestation. To assess repeatability of different components of variability, six measurements of nuchal translucency were made on each fetus, with a total of 1200 measurements. The data of this study demonstrate that 95% of the time the intraobserver, interobserver and caliper placement repeatability of measuring fetal nuchal translucency were less than 0.54 mm, 0.62 mm and 0.58 mm, respectively. In addition, the repeatability was unrelated to the size of the nuchal translucency. The findings of this study demonstrate that, when the nuchal translucency thickness is measured by well-trained operators, the measurement is highly reproducible.

INTRODUCTION

Abnormal collection of fluid behind the fetal neck (nuchal translucency) at 10–14 weeks of gestation is associated with increased risk for chromosomal defects. We have previously shown that the risk is related to both nuchal translucency thickness and maternal age[1–4]. The nuchal translucency thickness is ≥ 2.5 mm (3 mm when using machines that give measurements to the nearest millimeter) in more than 80% of fetuses with trisomies 21, 18 and 13 and in 4.5% of chromosomally normal fetuses[2]. For translucencies of 3 mm, 4 mm, 5 mm and ≥ 6 mm, there is a corresponding three-fold, 18-fold, 28-fold and 36-fold increase in maternal-age-related risk for trisomy[4]. However, the repeatability of measuring fetal nuchal translucency thickness has not been reported previously.

PATIENTS AND METHODS

This was a prospective study involving 200 pregnant women attending the Harris Birthright Research Centre for Fetal Medicine for measurement of fetal nuchal translucency thickness at 10–14 weeks of gestation. Transabdominal ultrasound examination (Toshiba SSA 250A, Toshiba Medical Systems Limited, Tokyo, Japan; 5-MHz curvilinear probe) was performed to obtain a sagittal section of the fetus for measurement of the maximum thickness of the subcutaneous translucency between the skin and the soft tissue overlying the cervical spine. Care was taken to distinguish between fetal skin and amnion, because at this gestation both structures appear as thin membranes. This was achieved by waiting for spontaneous fetal movement away from the amniotic membrane; alternatively, the fetus was bounced off the amnion by having the mother cough and/or tapping the maternal abdomen.

The nuchal translucency was measured on each of the 200 fetuses by two of four well-trained operators with similar experience. It was intended that each operator would examine the same number of fetuses. The order in which the operators performed the scans was determined by logistic considerations rather than at random. In the event, the number of patients examined by each operator pair varied from 29 to 40, and the number examined by each operator varied from 194 to 216.

To assess repeatability of different components of variability, six measurements of nuchal translucency were made on each fetus. The first operator generated the appropriate image and measured the nuchal translucency in the usual way and then generated a new image and repeated the measurement (intraobserver repeatability). This second image was frozen on the screen, but the calipers were removed and the second operator reset the calipers and made a measurement (caliper placement repeatability). The process was then repeated with the operators reversed (interobserver repeatability). Thus, a

Correspondence: Professor K. H. Nicolaides, Harris Birthright Research Centre for Fetal Medicine, King's College Hospital Medical School, Denmark Hill, London SE5 8RX, UK

Received 3–12–94
Revised 16–1–95
Accepted 21–1–95

sequence of six measurements was obtained: A1, A2, BP, B1, B2 and AP, where A and B represent the two operators, 1 and 2 indicate the two normal measurements, and P indicates the measurement that involved only placement of the calipers. Numeric displays on the screen were covered, so that the operators were blinded to the actual measurement and were unaware of results obtained by the previous operator. Measurements were recorded by an independent observer.

Repeatability of nuchal translucency thickness measurements and interobserver comparisons was assessed by use of the four normal measurements on each fetus (A1, A2, B1 and B2, described above). Intraobserver variation was analyzed by calculation of the standard deviation (SD) of the differences between the 400 pairs of measurements made by the same observer. Interobserver variation was analyzed by calculation of the SD of the differences between the means of pairs of measurements made by two observers on one fetus (n = 200). This value was multiplied by $\sqrt{2}$ to get the corresponding value for single measurements. Caliper placement repeatability was assessed from the SD of the 400 pairs of interobserver differences A2 – BP or B2 – AP. Two-way analysis of variance (ANOVA) for each pair of observers and unbalanced ANOVA for the whole data set led to the same findings, so we do not report these analyses.

RESULTS

The fetal nuchal translucency thickness was successfully measured by both operators in all 200 cases, and a total of 1200 measurements were made.

Distribution of fetal nuchal translucency thickness

The distribution of the nuchal translucency thickness in the 200 fetuses, with the use of the mean of the four 'normal' measurements, is shown in Figure 1. The nuchal translucency was ≥ 2.5 mm in 11 (5.5%) fetuses.

Repeatability and mean nuchal translucency thickness

There was no relationship between mean nuchal translucency thickness and either intraobserver difference (Figure 2) or interobserver difference (Figure 3). Since the repeatability was unrelated to the size of the nuchal translucency, the intra- and interobserver repeatability can be expressed simply by using the SD of the difference.

Repeatability

The standard deviation of differences between repeated readings by the same observer, pooled across the four observers, was 0.27 mm. Likewise, the interobserver standard deviation was 0.31 mm for single measurements and 0.22 mm for the mean of two measurements. Consequently, on 95% of occasions, the difference between two measurements lay within ± 2[5]. For example, 95% of the time two readings by the same observer (intraobserver) would not differ by more than ± 0.54 mm and ± 0.38 mm for the mean of two readings; the corresponding figures for two observers (interobserver) were ± 0.62 mm and ± 0.44 mm, respectively.

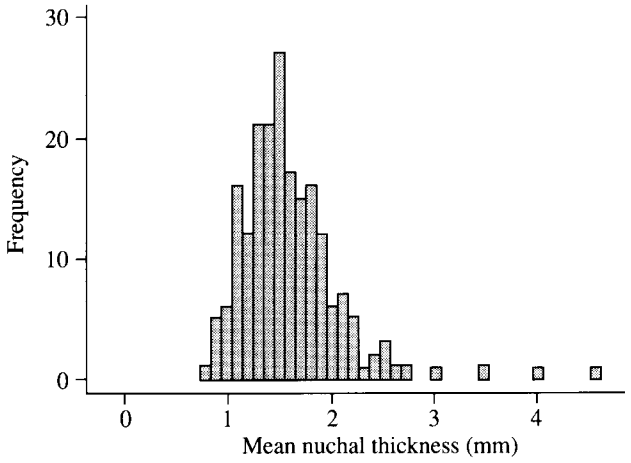

Figure 1 Distribution of the nuchal translucency thickness in the 200 fetuses, with the use of the mean of the four 'normal' measurements

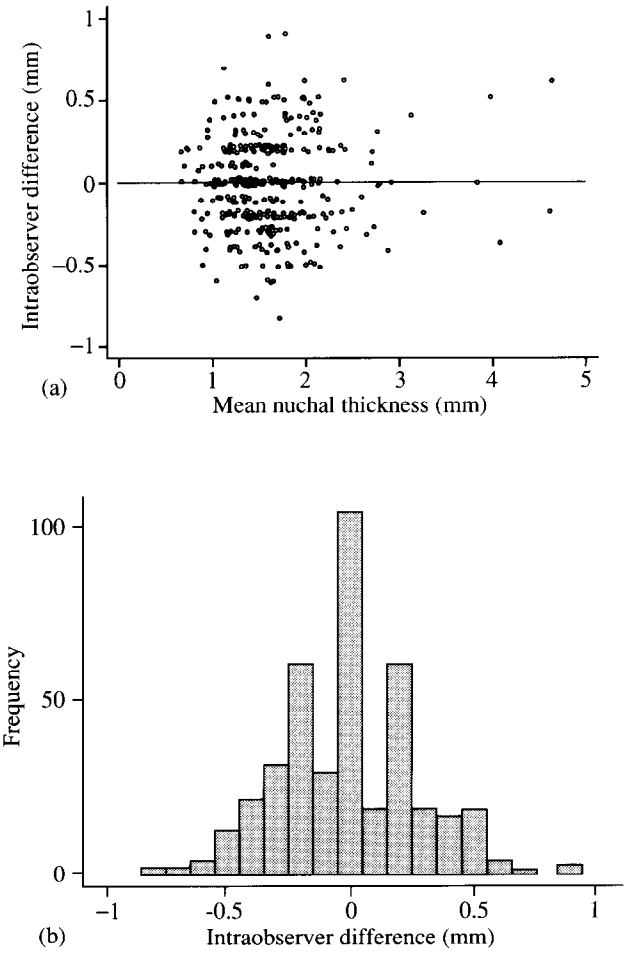

Figure 2 Scatter diagram (a) showing the intraobserver differences plotted against the mean of both readings by the observer and histogram (b) of the intraobserver differences

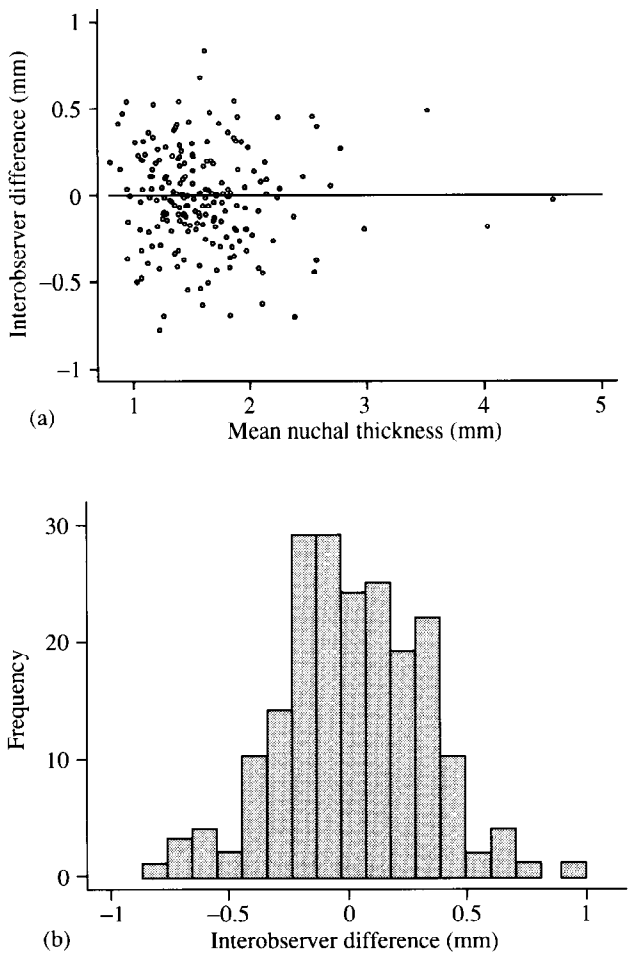

Figure 3 Scatter diagram (a) showing the interobserver differences (based on the mean of two observations) plotted against the mean of all four readings by those observers and histogram (b) of the interobserver differences

Analysis of caliper placement measurements

The interobserver standard deviation when the image was fixed was 0.29 mm (95% of the time, the two readings would not differ by more than ± 0.58 mm). Therefore, nearly all of the interobserver variation was due to the placing of the calipers rather than to the fixing of the image.

Differences between observers

The differences between the four observers were small. The two who agreed least well differed on average by 0.18 mm. However, there was less variation around the mean for these two observers than for most other pairs.

DISCUSSION

The data of this study demonstrate that 95% of the time the intraobserver and interobserver repeatability of measuring fetal nuchal translucency thickness was less than 0.54 mm and 0.62 mm, respectively.

The distribution of nuchal translucency was skewed and only 11 (5.5%) fetuses had a nuchal translucency of ≥ 2.5 mm. However, these findings are consistent with our previous findings of 4.5% and 3.6%[2,6]. The repeatability was unrelated to the size of the nuchal translucency and it is therefore reasonable to infer that the repeatability is the same for large values, even though there were relatively few in this study.

The difference in repeated readings by the same observer and between observers may be accounted for by two main factors: first, the generation of a new image, during which the fetal position may have changed by flexion or extension of the fetal neck and by rotation of the spine to the anterior or posterior; and second, the correct placement of the calipers. We have shown that the caliper placement repeatability would be less than 0.58 mm 95% of the time, which is similar to the intraobserver and interobserver repeatability. These results suggest that a large part of the variation can be accounted for by the placement of the calipers rather than by the generation of the image. Digital image processing and automation of caliper placement should reduce the differences in repeatability. In the meantime, repeatability could be improved by taking the mean of two good measurements rather than one, because the repeatability of two measurements for a single operator is reduced from 0.54 to 0.38 mm and the respective figures for two operators are 0.62 mm and 0.44 mm.

A potential criticism of screening by ultrasound, in contrast to biochemical testing, is that scanning requires highly skilled operators. This is certainly true for many of the subtle markers of chromosomal abnormalities detectable at 16–20 weeks' gestation[7]. However, the skill necessary for measurement of nuchal translucency at 10–14 weeks is no greater than that required to obtain a reliable measurement of the crown–rump length, which is essential for accurate dating of pregnancy and correct interpretation of serum biochemistry results. The findings of this study demonstrate that when the nuchal translucency thickness is measured by well-trained operators, the measurement is highly reproducible.

REFERENCES

1. Nicolaides, K. H., Azar, G., Byrne, D., Mansur, C. and Marks, K. (1992). Fetal nuchal translucency: ultrasound screening for chromosomal defects in first trimester of pregnancy. *Br. Med. J.*, **304**, 867–9
2. Nicolaides, K. H., Brizot, M. L. and Snijders, R. J. M. (1994). Fetal nuchal translucency thickness: ultrasound screening for fetal trisomy in the first trimester of pregnancy. *Br. J. Obstet. Gynaecol.*, **101**, 782–6
3. Pandya, P. P., Brizot, M. L., Kuhn, P., Snijders, R. J. M. and Nicolaides, K. H. (1994). First trimester fetal nuchal translucency thickness and risk for trisomies. *Obstet. Gynecol.*, **84**, 420–3
4. Pandya, P. P., Kondylios, A., Hilbert, L., Snijders, R. J. M. and Nicolaides, K. H. (1995). Chromosomal defects and outcome in 1015 fetuses with increased nuchal translucency. *Ultrasound Obstet. Gynecol.*, **5**, 15–19

5. Bland, J. M. and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, **1**, 307–10
6. Pandya, P. P., Goldberg, H., Walton, B., Riddle, A., Shelley, S., Snijders, R. J. M. and Nicolaides, K. H. (1995). The implementation of first trimester scanning at 10–13 weeks' gestation and the measurement of fetal nuchal translucency thickness in two maternity units. *Ultrasound Obstet. Gynecol.*, **5**, 20–5
7. Nicolaides, K. H., Snijders, R. J. M., Gosden, C. M., Berry, C. and Campbell, S. (1992). Ultrasonographically detectable markers of fetal chromosomal abnormalities. *Lancet*, **340**, 704–7