

Nuchal translucency measurement: who is right, who is not?

Engin Korkmazer¹, Emine Arslan², Özgür Akkurt³, Muzaffer Temur¹, Tayfur Çift¹

¹Clinics of Obstetrics & Gynecology, Bursa Yüksek İhtisas Training and Research Hospital, Bursa, Turkey ²Department of Obstetrics & Gynecology, Faculty of Medicine, Hitit University, Çorum, Turkey ³Clinics of Perinatology, Bursa Yüksek İhtisas Training and Research Hospital, Bursa, Turkey

Abstract

Objective: Fetal nuchal translucency (NT) measurement at first trimester provides prediction for fetal aneuploidy and cardiac anomalies. Performance of NT as a screening marker has not been consistent in studies. Measurement of NT has high intra- and interobserver variability. Radiologists, obstetricians and perinatology experts measure the NT. There is no consensus about who should perform the NT measurement. In this study we compared the correlation of NT measurement in three groups depend on mean thickness and distribution of NT.

Methods: A total of 929 participants were recruited for this study. 7 radiologists, 8 obstetricians and 1 perinatology expert measured NT. Crown-rump length (CRL), mean NT and NT distribution were calculated for each group.

Results: Perinatology expert's mean NT measurement was significantly higher than that of radiologists and obstetricians (p<0.05). Measurements of the perinatal expert also had significantly different distribution than other groups (p<0.05). There was no significant difference between the groups in terms of CRL values. Interobserver reliability coefficients with 95% confidence intervals for CRL and NT were 0.967 (0.910–0.987, p<0.001) and 0.596 (0.455–0.845, p<0.001), respectively.

Conclusion: There are statistically significant differences for mean NT value and distribution in three groups. Measurements of the perinatal expert has higher mean thickness and distribution. It is obvious that there is a need for standardization in NT measurement and it is necessary to evaluate the perinatal outcomes of these three groups and to approximate the two groups with the most accurate result group.

Keywords: Nuchal translucency, perinatology, obstetrician, radiologist.

Özet: Ense kalınlığı ölçümü: Kim doğru, kim yanlış?

Amaç: İlk trimesterde fetal ense kalınlığı (NT) ölçümü, fetal anöploidi ve kardiyak anomaliler için kestirim olanağı sunmaktadır. Bir tarama belirteci olarak NT ölçümünün yapılması çalışmalarda tutarlı sonuçlar vermemektedir. NT ölçümü, yüksek gözlemciler arası ve gözlemciler içi değişkenliğe sahiptir. Radyologlar, doğum ve perinatoloji uzmanları NT'yi ölçmektedir. NT ölçümünü kimin yapması gerektiği konusunda fikir birliği bulunmamaktadır. Bu çalışmada, ortalama kalınlık ve NT'nin dağılımına bağlı olarak üç grupta NT ölçümünün korelasyonunu karşılaştırdık.

Yöntem: Çalışmaya toplam 929 katılımcı dahil edildi. Yedi radyolog, 8 doğum uzmanı ve 1 perinatoloji uzmanı NT'yi ölçtü. Her grup için tepe-makat uzunluğu (CRL), ortalama NT ve NT dağılımı hesaplandı.

Bulgular: Perinatoloji uzmanının ortalama NT ölçümü, radyologların ve doğum uzmanlarının ölçümünden anlamlı şekilde daha yüksekti (p<0.05). Ayrıca perinatoloji uzmanı ölçümleri diğer gruplardan anlamlı şekilde daha farklı bir dağılıma sahipti (p<0.05). CRL değerleri bakımından gruplar arasında hiçbir anlamlı farklılık yoktu. CRL ve NT için %95 güven aralığıyla gözlemciler arası güvenilirlik katsayıları sırasıyla 0.967 (0.910–0.987, p<0.001) ve 0.596 (0.455– 0.845, p<0.001) idi.

Sonuç: Üç grupta ortalama NT değeri ve dağılımı bakımından istatistiksel olarak anlamlı farklılıklar bulunmaktadır. Perinatoloji uzmanı ölçümleri daha yüksek ortalama kalınlık ve dağılıma sahiptir. NT ölçümünde bir standart oluşturma ihtiyacı olduğu açıktır ve bu üç grubun perinatal sonuçları değerlendirmesi ve iki grubun en doğru sonuç grubuna yaklaşması önemlidir.

Anahtar sözcükler: Ense kalınlığı, perinatoloji, doğum uzmanı, radyolog.

Correspondence: Engin Korkmazer, MD. Clinics of Obstetrics & Gynecology, Bursa Yüksek İhtisas Training and Research Hospital, Bursa, Turkey. e-mail: ekorkmazer@yahoo.com

Received: March 20, 2018; Accepted: May 13, 2018

Please cite this article as: Korkmazer E, Arslan E, Akkurt Ö, Temur M, Çift T. Nuchal translucency measurement: who is right, who is not? Perinatal Journal 2018;26(2):64–68. ©2018 Perinatal Medicine Foundation

Available online at: www.perinataljournal.com/20180262002 doi:10.2399/prn.18.0262002 QR (Quick Response) Code:



deo**med**

Introduction

Nuchal translucency (NT) is a sonoluscent area behind the fetal neck in the first trimester and was first defined by Nicolaides et al.^[1] NT measurement at 11–13+6 weeks of gestation is one of the main points of Down syndrome screening. Increased NT is associated with 70% of Down syndrome alone.^[2] 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%.^[3] Moreover, increased NT is associated with other chromosomal anomalies, genetic syndromes, and structural anomalies.^[4] Performance of NT varies from study to study. Besides the other factors, the difficulty of obtaining proper images changes the success of the test. Small differences in caliper placement not only have the potential to alter an individual patient's calculated risk estimate significantly, but also may decrease the cumulative performance of the screening test by increasing rates of false-positive or false-negative results.^[5] The repeatability of NT measurement was found different in several studies.^[6]

The Fetal Medicine Foundation (FMF) and other foundations have published guidelines to promote a standardized measurement technique for obtaining a proper NT measurement. There is no consensus about who have to measure the NT. In our country radiologists, obstetricians and perinatology experts perform the first trimester aneuploidy scan. In our country, NT measurement is performed in private clinics, hospitals and university hospitals. Although NT measurement is taught in all three groups, it is not controlled in the following period. Aim of this study is to compare the NT measurement performance of these three groups at the same population.

Methods

This is a prospective study including 929 singleton pregnant women with first trimester aneuploidy scan. Inclusion criteria were having a singleton pregnancy and the agreement of the women to participate in this study. Measurements were carried out between June 2017 and September 2017. Each pregnancy was examined for number of fetuses, measurement of crown-rump length (CRL) and NT. Seven radiologists, 8 obstetricians and 1 perinatology expert measured the NT. Only the measurements of radiologists, obstetricians and perinatology experts who performed more than 100 measurements in a period of at least 3 months were included in the analysis. Only the perinatology expert had FMF certification for NT measurement. Radiologists and obstetricians were certified by Ministry of Health. All examiners examined and measured the fetal NT consecutively and independently without knowing the each other's results. Additionally each examiner scanned the CRL.

Demographic characteristics, pregnancy week, CRL and NT values were recorded. NT scans were performed between 11 and 13+6 weeks of gestation with CRL between 45 and 84 mm. The ultrasound machine was GE Voluson S6 (General Electric Healthcare, Chicago, IL, USA) for all groups. Transabdominal ultrasonography was performed on all women using a convex 2-5 MHz probe (4C-RS; General Electric Healthcare, Chicago, IL, USA). CRL and NT were measured as described by Nicolaides at al. three times for each woman.^[1] The NT value was expressed in a decimal of a millimeter. When three images were not obtained within 30 min, transvaginal ultrasonography (using a 5-7.5 MHz probe) was performed and time was added to the previous scans. The average of the three measurements was accepted as the NT value. We compared the CRL, mean NT values, distribution width of NT values and standard deviation of NT values for three groups (radiologists, obstetricians, and perinatology expert).

We used the SPSS version 21.0 software (SPSS Inc., Chicago, IL, USA). Continuous variables were expressed as mean±standard deviation. Differences between the values of NT and CRL in groups were tested by independent variables t test. The measurement results of the three examiners were analyzed for any significant differences and variations. Interobserver comparisons were done by reliability tests (Cronbach's alpha correlation coefficients). Two-way mixed effects model was used in cases where people effects were random and measured effects were mixed (absolute agreement definition) statistical analysis was performed using the Friedman test, the chi-square test and a multivariate analysis for measurement variations. A probability value of <0.05 was considered statistically significant. Study approved by local ethics committee.

Results

Nine hundred and twenty-nine women with singleton pregnancies were included in the study. The mean maternal age was 28.3 ± 5.5 (range: 16 to 47) years and

the mean gestational age at scan was 12 (range: 11 to 13+6) weeks. 14% of the pregnant women included in the study were over 35 years of age and all patients were Caucasian.

There was no significant difference between the groups in terms of CRL values. The mean of all group nonspecific median NT was 1.46±0.55. Interobserver reliability coefficients with 95% confidence intervals for CRL and NT were 0.967 (0.910–0.987, p<0.001) and 0.596 (0.455–0.845, p<0.001), respectively. For all groups median NT values, standard deviation and range are shown at **Table 1**.

Mean NT values for perinatology expert, radiologist and obstetrician were 1.66, 1.41 and 1.25, respectively. Standard deviation values for perinatology expert, radiologist and obstetrician were 0.75, 0.38 and 0.34, respectively. Mean NT values and distribution width were significantly higher in perinatology group (p<0.05) (**Fig. 1**). In the radiology group, the mean NT value and NT range was higher than the obstetrician group (**Table 2**).

Discussion

In this study we compared the inter-operator reliability of CRL and NT measurements by sonographers with different levels of experience. Our study is the largest study comparing three groups in the literature. The CRL and NT measurements are dominant parameters in prenatal screening for Down syndrome. Increased NT value has a major impact on the estimated risk for trisomy 21 and therefore on the patient's decision for or against invasive testing. The difficulties encountered when examining the NT are mainly related to fetal position, an increased maternal body mass index, nuchal cord, maternal abdominal wall thickness, quality of ultrasound machine and the inability of inexperienced sonographers to perform the scan correctly and examine Table 1. Demographic characteristics of the study group.

	n=929
Age	28.3±6
Parity	0.7±1.1
Week	12.2±0.6
CRL	62.5±8.4
NT	1.46±0.55

the NT. Variations in measurement decrease in comparison with experienced examiners.

Unsatisfactory quality of NT measurements can easily lead to overestimation or underestimation of the risk for Down syndrome. Small differences in measurement have the potential to significantly alter an individual's risk estimate for aneuploidy and increase the chance for false-positive or false-negative diagnoses. When the karyotype is normal and the NT is enlarged, the fetus is still at risk for cardiac abnormalities and congenital fetal abnormalities.^[7] However, NT screening displays high-



lue

05 05

Fig. 1. Distrubution of NT values for each group.

	Obstetrician	Radiologist	Perinatology expert	p-va	
CRL (mm)	62.5±8.4	61.7±9.7	62.1±8.7	N	
Mean NT (mm)	1.25	1.41	1.66	<0	
Standard deviation (mm)	0.34	0.38	0.75	<0	
NT range (mm)	0.7–6.1	0.7–5.7	0.8–7		

Table 2. Nuchal translucency measurement results for each group (n=929).

NS: not significant, NT: nuchal translucency.

er variability and significant operator dependence. Kagan et al. showed that an underestimation and overestimation of the CRL has a major impact, resulting in substantial underestimation or overestimation of those risks.^[8] A study regarding the measurement of NT had established that the ability to obtain reproducible results improves with training and good results are achieved after 80 scans.^[9] In our study, only the measurements of the experts who had performed at least 100 NT measurements were included in the study to increase the quality and the repeatability of the test.

Several studies evaluated the performance of NT measurement by providers.^[5,10] The current cross-sectional study analyzes the performance of NT measurement in same population. Mean NT values were statistically different in three groups. Perinatology expert group had the highest mean NT value and standard deviation. If we accept the measurements of perinatology expert as the gold standard, the low mean NT measurement of the radiologist and obstetrician groups will result in a false low calculation of the first trimester ane-uploidy risk. Aksoy et al. compared only obstetricians and radiologists in their study and found low inter-observer reliability of NT measurement.^[11]

NT measurement performance depends on sonographer. Among our results who were previously trained for perinatology, median NT value and NT distribution were higher, suggesting that experience does matter. These variations in the NT measured by the inexperienced sonographer can be explained by the failure and inability to achieve the exact mid-sagittal view.^[12] There is no quality monitoring programme for NT measurement in our country. In our study, none of the obstetricians and radiologists had FMF certificate for NT measurement. In FASTER study Malone et al. showed that 7.4% of NT measurements were measured incorrectly.^[13] Dalton et al. also confirmed that data in their study.^[14] Interobserver reliability coefficients with 95% confidence intervals for CRL and NT were 0.967 (0.910-0.987, p<0.001) and 0.596 (0.455–0.845, p<0.001), respectively. This result shows us that the interobserver reliability and reproducibility were high in the CRL measurement. However, reliability and reproducibility for NT evaluation were relatively low.

In our study there was no statistically significant difference between the groups in terms of CRL measurement. Also in the known literature it is seen that the interobserver variability is low and the reliability is high in terms of CRL measurement. In the largest study on this subject Souka et al. showed that the reliability and reproducibility of CRL measurement was high.^[15-17] Salomon et al. performed a simulation study using a simulation model to evaluate the impact of error in CRL measurements in cases of sequential combined screening for Down syndrome. After more than 3200 simulated cases were analyzed, they reported that Down syndrome screening might be highly sensitive to errors in CRL measurements.^[18]

The fact that there was only one perinatologist in our study and that the perinatal outcomes were not assessed constitutes the weakness of our study. However, in literature our study is the first one in which a large population of 929 patients were evaluated on this subject.

Conclusion

Measuring fetal NT and CRL accurately is essential for optimal combined first-trimester screening performance and prenatal care. Our findings show that precision of NT measurements is still largely dependent on sonographer's personal attitude in terms of endurance and accuracy. Interobserver reliability is high in CRL measurement. Although it was first defined 25 years ago, NT measurement still does not seem to be standardized. There is a need for a system that will standardize NT measurement and will make quality control over the years.

Conflicts of Interest: No conflicts declared.

References

- Nicolaides KH, Azar G, Byrne D, Mansur C, Marks K. Fetal nuchal translucency: ultrasound screening for chromosomal defects in first trimester of pregnancy. BMJ 1992;304:867–9.
- Snijders RJM, Noble P, Sebire N, Souka A, Nicolaides KH. UK multicenter project on assessment of risk of trisomy 21 by maternal age and fetal nuchal-translucency thickness at 10–14 weeks of gestation. Fetal Medicine Foundation First Trimester Screening Group. Lancet 1998;352:343–6.
- Bindra R, Heath V, Liao A, Spencer K, Nicolaides KH. Onestop clinic for assessment of risk for trisomy 21 at 11–14 weeks: a prospective study of 15,030 pregnancies. Ultrasound Obstet Gynecol 2002;20:219–25.
- 4. Jelliffe-Pawlowski LL, Norton ME, Shaw GM, Baer RJ, Flessel MC, Goldman S, et al. Risk of critical congenital heart

defects by nuchal translucency norms. Am J Obstet Gynecol 2015;212:518.e1-10.

- 5. Cuckle H, Platt LD, Thornburg LL, Bromley B, Fuchs K, Abuhamad A, et al.; Nuchal Translucency Quality Review Program of the Perinatal Quality Foundation. Nuchal Translucency Quality Review (NTQR) program: first one and half million results. Ultrasound Obstet Gynecol 2015;45:199-204.
- 6. Pajkrt E, Mol BW, Boer K, Drogtrop AP, Bossuyt PM, Bilardo CM. Intra- and interoperator repeatability of the nuchal translucency measurement. Ultrasound Obstet Gynecol 2000; 15:297-301.
- 7. Bilardo CM, Müller MA, Pajkrt E, Clur SA, van Zalen MM, Bijlsma EK. Increased nuchal translucency thickness and normal karyotype: time for parental reassurance. Ultrasound Obstet Gynecol 2007;30:11-8.
- Kagan KO, Hoopmann M, Baker A, Huebner M, Abele H, 8. Wright D. Impact of bias in crown-rump length measurement at first-trimester screening for trisomy 21. Ultrasound Obstet Gynecol 2012;40:135-9.
- 9 Braithwaite JM, Kadir RA, Pepera TA, Morris RW, Thompson PJ, Economides DL. Nuchal translucency measurement: training of potential examiners. Ultrasound Obstet Gynecol 1996;8:192-5.
- 10. Hermann M, Fries N, Mangione R, Boukobza P, Ville Y, Salomon LJ. Nuchal translucency measurement: are qualitative and quantitative quality control processes related? Prenat Diagn 2013;33:770-4.
- 11. Aksoy H, Ozyurt S, Aksoy U, Mutlu E, Tutus S, Babayigit MA, et al. A prospective study to assess the clinical impact of interobserver reliability of sonographic measurements of fetal

nuchal translucency and crown-rump length on combined first-trimester screening. North Clin Istanb 2015;2:92-100.

- 12. Staboulidou I, Wüstemann M, Vaske B, Scharf A, Hillemanns P, Schmidt P. Interobserver variability of the measurement of fetal nasal bone length between 11+0 and 13+6 gestation weeks among experienced and inexperienced sonographers. Ultraschall Med 2009;30:42-6.
- 13. Malone FD. Nuchal translucency-based Down syndrome screening: barriers to implementation. Semin Perinatol 2005; 29:272-6.
- 14. D'Alton ME, Cleary-Goldman J, Lambert-Messerlian G, Ball RH, Nyberg DA, Comstock CH, et al. Maintaining quality assurance for sonographic nuchal translucency measurement: lessons from the FASTER trial. Ultrasound Obstet Gynecol 2009:33:142-6.
- 15. Pexsters A, Luts J, Van Schoubroeck D, Bottomley C, Van Calster B, Van Huffel S, et al. Clinical implications of intraand interobserver reproducibility of transvaginal sonographic measurement of gestational sac and crown-rump length at 6-9 weeks' gestation. Ultrasound Obstet Gynecol 2011;38:510-5.
- 16. Verburg BO, Mulder PG, Hofman A, Jaddoe VW, Witteman JC, Steegers EA. Intra- and interobserver reproducibility study of early fetal growth parameters. Prenat Diagn 2008;28:323-31.
- 17. Souka AP, Pilalis A, Papastefanou I, Salamalekis G, Kassanos D. Reproducibility study of crown-rump length and biparietal diameter measurements in the first trimester. Prenat Diagn 2012;32:1158-65.
- 18. Salomon LJ, Bernard M, Amarsy R, Bernard JP, Ville Y. The impact of crown-rump length measurement error on combined Down syndrome screening: a simulation study. Ultrasound Obstet Gynecol 2009;33:506-11.