First-Trimester Ultrasound Nuchal Translucency Screening for Fetal Aneuploidy

Mitchell D. Feldman, MD, M.Phil.
Professor of Medicine
Division of General Internal Medicine
Department of Medicine
University of CA, San Francisco

California Technology Assessment Forum
February 11, 2004
San Francisco, CA
FIRST-TRIMESTER ULTRASOUND NUCHAL TRANSLUCENCY SCREENING FOR FETALANEUPLOIDY

INTRODUCTION

The California Technology Assessment Forum (CTAF) has received requests to review first-trimester ultrasound nuchal translucency (NT) screening for fetal aneuploidy (abnormal chromosome number), and particularly for Down syndrome.

PRIOR REVIEW

The topic of first trimester screening using NT for fetal aneuploidy and cardiac anomalies was last reviewed by the Blue Shield of California Medical Policy Committee on Quality and Technology in June of 2002. At that time the committee approved the recommendation that NT for fetal aneuploidy and cardiac anomalies did not meet TA criteria. Recently, a major study on first trimester screening that used NT was published in the New England Journal of Medicine (Wapner et al., 2003), prompting this re-evaluation.

BACKGROUND

Down syndrome is often diagnosed at birth on the basis of the typical facial features, hypotonia, and a single palmar crease. In addition to these deformities, mental retardation and several other serious problems may be evident postnataally, including duodenal atresia, congenital heart disease, and leukemia. An Alzheimer-like dementia often develops in the fourth or fifth decade of life and, for those who survive childhood, accounts for a reduced life expectancy (Pyeritz, 2002). Genetic chromosomal analysis (karyotyping) shows that most patients with Down syndrome have simple trisomy (tripling) of chromosome 21 in the G group, or some have unbalanced translocations.

The risk of bearing a child with Down syndrome increases exponentially with the age of the mother at the time of conception and rises markedly after age 35. By age 45, a mother has a 1 in 40 chance of having a child with Down syndrome. The risk of having a child with Down syndrome is 1/1,300 for a 25 year-old women; at age 35 the risk increases to 1/365. Women younger than 35 years give birth to about 70% of infants with Down syndrome (Newberger, 2000). The risk of other conditions associated with trisomy also increases, because of the increased predisposition of older oocytes to nondisjunction during meiosis. The average life expectancy of an individual with Down syndrome has increased to about 50 years of age, mostly due to advances in cardiac surgery (Pyeritz, 2002).
Many fetuses with trisomy 21 can now be detected prenatally by a quantitative assessment of risk followed by invasive testing (chorionic-villus sampling [CVS] or amniocentesis) in the 5% of pregnancies at the highest estimated risk. Selection of the high-risk group is currently accomplished by combining maternal age and results of second or first-trimester maternal serum biochemical tests (Krantz et al, 1996; Taipale et al, 1997; Snijders et al., 1998; Haddow et al, 1998b).

Early in the second trimester (between 15 and 20 weeks of pregnancy), a combination of three biochemical tests, the so-called "triple screen" (human chorionic gonadotrophin (HCG) or free beta-hCG (FBHCG), alpha-fetoprotein (AFP) and unconjugated estriol (uE(3)), yields a detection rate of approximately 67% for a 5% screen-positive rate. In the presence of fetal trisomy 21, the average maternal serum concentrations of human chorionic gonadotropin are higher than normal and those of alpha-fetoprotein and estriol are lower than normal. A fourth marker, inhibin A, increases the detection rate by about 7% for the same false-positive rate. A combination of four biochemical tests, the so-called "quadruple screen" (pregnancy associated plasma protein A (PAPP-A), FBHCG, AFP, and uE(3)), yields a detection rate of approximately 70% for a 5% screen-positive rate. The accuracy of these predictions has been confirmed in more than 20 large prospective intervention studies (Tsukerman et al, 1999; Cuckle, 2000). Integrating the first-trimester and the second-trimester biochemical markers yields a 94% detection rate for a 5% false-positive rate. If the false-positive rate is set at 1%, the detection rate would be 85% (Benn, 2002; Wald et al, 2000).

Prenatal noninvasive screening for Down syndrome is widely used in the U.S., with approximately 2.5 million of the estimated four million births per year receiving such antenatal screening (Malone et al., 2000).

All fetuses have nuchal (nape of the neck) translucency (thickening) detectable by a first trimester ultrasound. Since the early 1990s, first-trimester nuchal translucency thickness has been recognized as a marker of Down syndrome due to trisomy 21 (Mol et al, 1999). The probability that a fetus has trisomy 21 increases strongly with the nuchal translucency thickness. Recently, studies have assessed whether prenatal trisomy detection rates can be further increased, over those obtained by biochemical testing, if first-trimester ultrasound measurement of nuchal translucency thickness is used as an additional marker of risk.
First-Trimester Ultrasound Fetal Nuchal Translucency Screening

The term nuchal translucency refers to the ultrasound measurement of nuchal skin late in the first trimester. Newborns with Down syndrome as well as those with other autosomal trisomies, Turner syndrome, or other genetic syndromes have been observed to have redundant nuchal skin. Nuchal edema occurs in the fetus as well and can be seen with ultrasound. Ultrasound findings range from slight thickening of nuchal skin, to congenital malformations such as cystic hygroma. The size of the NT increases with gestational age and fetal crown-rump length, but disappears after 14 weeks gestational age when the subcutaneous tissue becomes more echogenic (Haak and van Vught, 2003; Chasen and Skupski, 2003). The exact etiology of increased NT is unknown. Possible mechanisms for increased NT include: cardiac failure in association with abnormalities of the heart, venous congestion in the head and neck, altered composition of the extracellular matrix, abnormal or delayed development of the lymphatic system, and others (Bindra et al., 2002; Borruto et al., 2002).

During fetal nuchal translucency screening, ultrasonography is used to assess the degree of thickening (nuchal translucency) or the presence of a fluid collection (cystic hygroma) at the nape of the fetal neck. These abnormalities have been correlated with both genetic disorders and structural anomalies.

Nuchal translucency screening has been proposed for early detection of trisomy 21 (Down syndrome), trisomy 18, trisomy 13, Turner's syndrome (XO), other aneuploidies (Sen, 2001) and for cardiac anomalies (Rizzo et al., 2003; Galindo et al., 2003; Haak et al., 2002; Hyett et al., 1997; Hyett et al., 1999; Zosmer et al., 1999; Mavrides et al., 2001; Hiippala et al., 2001; Ghi et al., 2001). This screening is most accurately performed in the first trimester between 10 and 14 weeks' gestation (Beamer, 2001).

The extent of nuchal translucency is measured as the maximum thickness of the sonolucent zone between the inner aspect of the fetal skin and the outer aspect of the soft tissue overlying the cervical spine or the occipital bone (Taipale et al, 1997). Increased nuchal translucency is defined a priori using a cut-off value (such as >3.0 mm) or percentile (such as > 95th percentile), based on previously collected data suggesting that cut-off as one reliable in differentiating between fetuses at low risk for chromosomal abnormalities and those at high risk (Nicolaides et al., 1992; Taipale et al., 1997).

NT can be measured successfully by transabdominal ultrasound examination in about 95% of cases; the other 5% require transvaginal sonography (Nicolaides et al., 2002). The ability to measure NT and obtain reproducible results improves with training and with the use of appropriate equipment. Good results are generally achieved after 80 and 100 scans for the transabdominal and transvaginal routes respectively (Nicolaides et al., 2002), but ongoing quality assurance is required.
Other sonographic findings are considered "soft" markers for Down syndrome. These markers, often found among normal fetuses, include slightly shortened humerus and femur bones, pyelectasis, echogenic intracardiac focus and hypoplasia of the middle phalanx of the fifth digit. More recently, several studies have noted an association of an absent fetal nasal bone with Down syndrome (Cicero et al, 2003).

First-trimester screening by fetal nuchal translucency, along with serum screening has been proposed as a significant improvement over second-trimester serum screening programs, which is the current standard of care. However, this proposal has been controversial (Mennuti and Driscoll, 2003; Malone et al., 2000; Chasen et al., 2001; Chasen and Skupski, 2003). The potential advantages of nuchal translucency screening include higher detection rates, potentially resulting in fewer women needing to undergo chorionic villus sampling or amniocentesis, and an earlier gestational age at karyotyping and diagnosis, which enables women choosing to terminate a pregnancy, to do so earlier.

**TA Criterion 1:** The technology must have final approval from the appropriate government regulatory bodies.

While the FDA approval process does not speak specifically on the use of ultrasound for nuchal translucency screening, ultrasound is widely used in obstetrics and gynecology. The Medical Device Amendments of 1976 began the review and regulation of all devices including ultrasound. All pre-amendment devices were given clearance at that time. Many abdominal and transvaginal ultrasounds fall into this pre-amendment category.

TA criterion 1 is met.

**TA Criterion 2:** The scientific evidence must permit conclusions concerning the effectiveness of the technology regarding health outcomes.

The major outcome of interest is fetal mortality and neonatal morbidity due to chromosomal abnormalities.

In detection of fetal aneuploidy, published studies have investigated nuchal translucency testing alone and nuchal translucency testing in combination with maternal age, with first-trimester biochemical tests, and with second-trimester biochemical tests. Other studies have investigated nuchal translucency in diagnosis of cardiac anomalies. Observational studies have been published for all five groups. No randomized, controlled trials of nuchal translucency testing have been published.

Outcomes most commonly reported in published studies include test sensitivity and specificity for aneuploidy (usually trisomy 21 and all aneuploidies). Other outcomes reported include positive predictive value and negative predictive value. A high positive predictive value for nuchal translucency would be useful for identifying fetuses likely to have chromosomal abnormalities and who might therefore be selected for further invasive diagnostic tests (chorionic villus...
sampling or amniocentesis). A high negative predictive value for nuchal translucency testing would be useful for identifying fetuses who are unlikely to have chromosomal abnormalities and who therefore do not require invasive diagnostic tests.

Some studies investigating nuchal translucency testing have been performed on mothers who are at elevated risk for fetal trisomy because of advanced maternal age, family history of aneuploidy, or positive maternal biochemical tests, while other studies have investigated the usefulness of increased nuchal translucency testing in unselected women. In most of the prospective screening studies when the fetal nuchal translucency was increased, the parents were counseled about the increased risk of aneuploidy and offered karyotyping by chorionic villus sampling or amniocentesis. Studies have varied in the rate of parental acceptance of the karyotyping. Importantly, studies have also varied in the degree of follow-up (karyotyping and/or pathological examination) of pregnancies terminated, lost through spontaneous abortion, or stillborn, and in the degree of follow-up of apparently normal births.

Several authors have emphasized two important aspects of prenatal screening for fetal abnormalities: first, the need for ascertainment of all affected cases in the population examined, and second, definition of the potential impact of prenatal diagnosis on the prevalence of a given abnormality and live births (Wald et al., 2003; Sebire et al., 1998; Mol et al., 1999; Malone et al., 2000). Unfortunately, many of the published studies do not report data addressing these concerns.

Levels of Evidence: 3, 4, and 5

TA criterion 2 is met.

TA Criterion 3: The technology must improve the net health outcomes.

Patient Benefits

More than 65 studies of nuchal translucency as a marker for fetal aneuploidy have been published since 1992.

Pregnancy Outcome Studies

Fifteen studies do not report detection rates or sensitivities, but report pregnancy outcomes in cases with abnormal nuchal translucency measurements (Hewitt et al., 1993; Pandya et al., 1994; Pandya et al., 1995; Fukuda et al., 1997; Brady et al., 1998; Souka et al., 1998; Bilardo et al., 1998; Van Vugt et al., 1998; Adekunle et al., 1999; Pajkrt et al., 1999; Maymon et al., 2000; Michailidis et al., 2001; Souka et al., 2001; Hiippala et al., 2001; Mangione et al., 2001; Spencer, 2002). These studies generally indicate that measurement of nuchal translucency at 12-14 weeks' gestation is a good marker for chromosome abnormalities (Hewitt et al., 1993; Pandya et al., 1994; Pandya et al., 1995; Bilardo et al., 1998; Adekunle et al., 1999; Maymon et al., 2000; Mangione et al., 2001; Spencer, 2002).
Furthermore, the incidence of chromosome abnormalities increases with increasing nuchal translucency thickness, as well as with increasing maternal age (Hewitt et al., 1993; Pandya et al., 1994; Maymon et al., 2000; Mangione et al., 2001). For example, in one study, 17.8% of fetuses with a nuchal translucency >99th percentile had an adverse pregnancy outcome versus 1.5% for those with a normal measurement (Michailidis et al., 2001).

Enhanced NT is also associated with cardiac defects and other disorders that are not identifiable through cytogenetic testing. When the nuchal translucency is abnormal and the karyotype is found to be normal, the outcome of the pregnancy may sometimes still be unfavorable, with problems including major cardiac defects, terminations for fetal abnormalities, spontaneous abortions, neonatal deaths, diaphragmatic hernias, skeletal defects, renal anomalies, postnatal developmental delays and Noonan's syndrome (Pandya et al., 1995; Brady et al., 1998; Bilardo et al., 1998; Souka et al., 1998; Van Vugt et al., 1998; Adekunle et al., 1999; Pajkrt et al., 1999; Michailidis et al., 2001; Hiippala et al., 2001; Mangione et al., 2001). In this situation, the pregnancy outcome appears to be correlated with the degree of nuchal translucency thickness (Pajkrt et al., 1999). For example, in one review, the chance of a live birth with no congenital defects was 86% in the group with nuchal translucency of 3.5-4.4 mm, 77% for those with translucency of 4.5-5.4 mm, 67% for those with translucency of 5.5-6.4 mm, and only 31% for those with translucency of > 6.5 mm (Souka et al., 2001).

Screening Studies

Table 1 summarizes 17 published studies of nuchal translucency alone in screening for a variety of fetal aneuploidies. Early studies showed wide variation in detection of fetal aneuploidies by nuchal translucency screening. However, more recent studies have shown less variation. Detection rates or sensitivities of nuchal translucency alone for trisomy 21 (Down syndrome) have varied between 43% and 76.5%. Detection (sensitivity) rates have increased over time, with higher rates of approximately 70% to 80% for a 5% false-positive rate in studies published in the last five years (Souter et al, 2001). Specificity of nuchal translucency alone for trisomy 21 has been reported at 95.5-99.9%.

Table 2 summarizes 12 published studies of nuchal translucency combined with maternal age in screening for fetal aneuploidy. Detection rates or sensitivities for nuchal translucency combined with maternal age have varied from between 75% and 91% for trisomy 21, to between 75% and 100% for any aneuploidy,. Specificity rates have not been reported. Positive predictive values of 3.2-4% for trisomy 21 and 5.3-7.8% for any aneuploidy have been reported.

Table 3 summarizes 20 published studies of nuchal translucency combined with first-trimester biochemical testing in screening for fetal aneuploidy. There has been gradual recognition that the most appropriate use of NT for prenatal diagnosis is in conjunction with biochemical testing. Usual first-trimester biochemical tests performed included
PAPP-A, BHCG or FBHCG, and sometimes included AFP. For the combined testing, detection rates have varied between 73% and 100% for trisomy-21, with false positive rates of 4% to 5%.

Wapner et al., (2003) report on the results of a prospective multicenter study (the BUN study) conducted in the U.S., which examined the results of combined first trimester screenings for trisomies 21 and 18. Women of any age with a singleton pregnancy between 74 and 97 days of gestation were offered prenatal screening based on maternal age, maternal levels of free human chorionic gonadotropin and pregnancy-associated plasma protein A, and ultrasonographic measurement of fetal nuchal translucency. They used more than 40 sonographers at 12 centers. Measurements of nuchal translucency were assessed according to the standards of the Fetal Medicine Foundation of London. (A secondary goal of the BUN study was to evaluate methods of training and quality control that best maintain accuracy of NT measurements. The results of this aspect of the study were reported in a recent paper (Snijders et al., 2002). To be consistent with the approach used in second trimester screening, they chose a risk of 1 in 270 pregnancies for trisomy 21 and a risk of 1 in 150 for trisomy 18 as the threshold for a positive result. Fetal chromosomal status was determined by prenatal karyotype analysis or by evaluation of the phenotype at birth. At a cutoff of 1:270, the rate of detection of Down's syndrome was 85.2 percent, with a 9.4 percent rate of positive screening results. At a false positive rate of 5 percent, the risk was 1:129 and the sensitivity was 78.7 percent (95 percent confidence interval, 66.3 to 88.1). Among women 35 years of age or older, screening identified 89.8 percent of fetuses with trisomy 21, with a false positive rate of 15.2 percent, and 100 percent of fetuses with trisomy 18. These results are similar to those of other studies of first trimester screening (Snijders et al., 1998). The authors conclude that first trimester screening that combines maternal age, levels of PAPP-A and free-BhCG and NT is comparable with second trimester screening even after they adjusted for potentially non-viable pregnancies.

Table 4 summarizes five published studies of nuchal translucency combined with second-trimester biochemical testing in screening for fetal aneuploidy, usually trisomy 21. Second-trimester biochemical tests performed usually included AFP, BHCG or FBHCG, and E3. Detection rates or sensitivities for nuchal translucency alone (53.4%-75%) have been lower than those for combined nuchal translucency and biochemical testing (80.6%-95%). Specificity has not been reported. Positive predictive values for trisomy 21 of 2.7% to 12% have been reported for the combined testing. Invasive procedure rates have varied between 4.2% and 8.6%.

Of these studies, only the Serum, Urine and Ultrasound Screening Study (SURUSS) study (Wald et al, 2003) and the ongoing FASTER trial* have the ability to compare the screening performance of first and second trimester tests as well as combine them into a single integrated test in an unbiased way because in these studies intervention was offered after the second trimester screening results were available. Because Down syndrome pregnancies are more likely to miscarry than unaffected pregnancies, bias will arise in estimations of the rate of Down syndrome pregnancies when detected earlier in gestation as compared with Down syndrome pregnancies that are not detected.
The SURUSS was a prospective study that combined first and second trimester strategies to identify the most effective and safest method of antenatal screening for Down Syndrome. Twenty-five maternity units mainly in the UK collected data on over 47,000 unselected singleton pregnancies in both the first and second trimester of pregnancy without planned intervention in the first trimester. There were 101 singleton Down syndrome pregnancies recruited into the study. Measurements included nuchal translucency; biochemical tests including AFP, free beta-hCG, total hCG, unconjugated estriol, PAPP-A and inhibin-A; and several different urine assays. Sonographers in the study received training in NT measurement and they report that a “process of quality control was maintained throughout the study.” Sonographers spent up to 20 minutes to obtain the requested three images (though the mean was around five minutes) and sent a hard copy to the coordinating center for independent review. Overall, an NT measurement could not be obtained in 9% of pregnancies. Factors that predicted the inability to obtain a satisfactory image included sonographer experience and the make and model of the ultrasound machine (independent of sonographer experience). They found that for sonographers who had scanned fewer than 200 pregnancies, images were technically poor 11% of the time compared with 7% for those who had screened 400 or more pregnancies. The main outcome measures of this trial were efficacy (presented as the false-positive rate for an 85% detection rate), safety (as expressed by fetal losses due to amniocentesis or CVS), and cost effectiveness. NT and biochemical measurements were converted to multiples of the median (MoM), which were calculated for NT by dividing the observed nuchal translucency value by the expected median for that gestational age. The most effective screening test was the “integrated test” (NT and PAPP-A at 10 weeks and AFP, uE, free beta-hCG and inhibin-A at 14-20 weeks) with an estimated 85% detection rate for a 1.2% false positive rate. Use of the “combined test,” the same test used in the BUN study (NT, PAPP-A and free beta-hCG), led to an 85% detection rate with a 6% false positive rate. With the integrated test, the rate of procedure related unaffected fetal loss was 9 per 100,000 women screened compared to 45 such losses when either first trimester screening with the combined test or a second trimester quadruple test was performed. The authors concluded that NT alone and the “double test” (AFP and free beta-hCG) were the least safe methods (i.e. led to the largest number of fetal losses per 100,000 women). They conclude that NT is a good screening marker that can be measured satisfactorily in routine practice, but that NT is a poor test when used with maternal age alone. They also concluded that the mean of three NT measures is superior to one measure and that the make and model of the ultrasound machine has an important influence on obtaining satisfactory NT measurement.

As with the BUN study, most of the women enrolled in SURUSS were white (91% in SURUSS, 83% in BUN) potentially limiting the generalizability of their results. Chen et al., (2002) found that there were statistically significant differences in NT measurements between different ethnic groups, though concluded that these were clinically insignificant as reflected by similar screen positive rates.
<table>
<thead>
<tr>
<th>Study Author, Date</th>
<th>Type of Study</th>
<th>No., Type of Population</th>
<th>Nuchal Translucency Abnormal Cut-off</th>
<th>% Sensitivity or Detection Rate</th>
<th>Specificity</th>
<th>False Positive</th>
<th>Positive Predictive Value (PPV)</th>
<th>Aneuploidies Detected</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wayda, 2001</td>
<td>Prospective screening</td>
<td>7,044; singleton or multiple</td>
<td>&gt; 2.5 mm</td>
<td>4.5%</td>
<td>76.47%</td>
<td>97.39%</td>
<td>2.81%</td>
<td>Trisomy 21</td>
<td>NPV: 99.94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.97% any</td>
<td>95.94%</td>
<td>4.05%</td>
<td>10.39%</td>
<td>T18 (8), T13 (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>low-and high-risk</td>
<td>≥ 3.0 mm</td>
<td>2.8%</td>
<td>100%</td>
<td>95.74%</td>
<td>4.26%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>84.35% any</td>
<td>97.61%</td>
<td>2.39%</td>
<td>14.66%</td>
<td>45XO (1), other (4)</td>
<td>NPV: 99.92%</td>
</tr>
<tr>
<td>Tihlagonathan, 1999</td>
<td></td>
<td>11,398; singleton or low-and high-risk</td>
<td>&gt; 1/200</td>
<td>4.90%</td>
<td>76.1%</td>
<td>T21, 97.38%</td>
<td>2.61%</td>
<td>6.81%</td>
<td>Trisomy 21</td>
</tr>
<tr>
<td>Whitlow, 1999</td>
<td>Prospective cross-sectional</td>
<td>6,634; singleton or multiple</td>
<td>&gt; 99th percentile</td>
<td>0.90%</td>
<td>57%</td>
<td>T21, 99.9%</td>
<td>4.26%</td>
<td>5.52%</td>
<td>Triploidy (2)</td>
</tr>
<tr>
<td>Pajkrt, 1998a</td>
<td>Prospective cohort</td>
<td>1,473; singleton or multiple</td>
<td>≥ 3.0 mm</td>
<td>2.2%</td>
<td>67%</td>
<td>T21</td>
<td>3.10%</td>
<td>121, others</td>
<td>NT not measurable in 10%</td>
</tr>
<tr>
<td>Pajkrt, 1998b</td>
<td>Prospective cohort</td>
<td>2,247; singleton or multiple</td>
<td>≥ 3.0 mm</td>
<td>5.4%</td>
<td>69%</td>
<td>T21</td>
<td>4.0%</td>
<td>121, others</td>
<td></td>
</tr>
<tr>
<td>Sebire, 1998</td>
<td>Prospective screening</td>
<td>161,992; singleton percentile</td>
<td>≥ 95th percentile</td>
<td>0.90%</td>
<td>57%</td>
<td>T21, 99.9%</td>
<td>4.26%</td>
<td>5.52%</td>
<td>Triploidy (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>87.3%</td>
<td>45XO</td>
<td>40%</td>
<td>47XXX, 47XY, only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hafner, 1998</td>
<td>Prospective screening</td>
<td>4,233; singleton or multiple</td>
<td>≥ 2.5 mm</td>
<td>1.75%</td>
<td>43%</td>
<td>T21, 98.3%</td>
<td>1.7%</td>
<td>4%</td>
<td>T21, 118, 113</td>
</tr>
<tr>
<td>D’Ottavio, 1997</td>
<td>Prospective screening</td>
<td>3514; singleton or multiple</td>
<td>≥ 4 mm</td>
<td>0.97%</td>
<td>70%</td>
<td>T21, 98.5%</td>
<td>1.5%</td>
<td>14.8%</td>
<td>T21, Transvaginal US.</td>
</tr>
<tr>
<td>Economides, 1998</td>
<td>Prospective screening</td>
<td>2281; singleton or multiple</td>
<td>≥ 3 mm</td>
<td>0.8%</td>
<td>54%</td>
<td>T21, 99.9%</td>
<td>9%</td>
<td>T21, T18, T13</td>
<td>NPV: 100%</td>
</tr>
<tr>
<td>Josephson, 1999</td>
<td>Prospective screening</td>
<td>1,444; singleton or multiple</td>
<td>≥ 4 mm</td>
<td>0.4%</td>
<td>69%</td>
<td>T21, 99.6%</td>
<td>17%</td>
<td>T13</td>
<td></td>
</tr>
<tr>
<td>Taipale, 1997</td>
<td>Prospective screening</td>
<td>10,010; singleton or multiple</td>
<td>≥ 3.0 mm or ≥ 4 mm</td>
<td>0.8%</td>
<td>54%</td>
<td>T21, 99%</td>
<td>9%</td>
<td>T21, T18, T13</td>
<td>NPV: 100%</td>
</tr>
<tr>
<td>Hafner, 1995</td>
<td>Prospective screening</td>
<td>1,972; singleton or multiple</td>
<td>≥ 2.5 mm</td>
<td>50%</td>
<td>T21, 97.1%</td>
<td>1.7%</td>
<td>T21, others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brambati, 1995</td>
<td>Prospective screening</td>
<td>1,819; singleton or multiple</td>
<td>≥ 3.0 mm</td>
<td>2.5%</td>
<td>30%</td>
<td>T21, 96%</td>
<td>11%</td>
<td>45.5%</td>
<td>T21, T18</td>
</tr>
<tr>
<td>Salvesen, 1995</td>
<td>Prospective cohort</td>
<td>97; high-risk</td>
<td>≥ 3.0 mm</td>
<td>11%</td>
<td>45.5%</td>
<td>T21, T18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicolaides, 1994</td>
<td>Prospective screening</td>
<td>1,273; singleton or multiple</td>
<td>≥ 3.0 mm</td>
<td>86%</td>
<td>T21, 95.5%</td>
<td>4.5%</td>
<td>36%</td>
<td>Various</td>
<td>NPV: 99.6%</td>
</tr>
<tr>
<td>Nicolaides, 1992</td>
<td>Prospective screening</td>
<td>827; singleton or multiple</td>
<td>≥ 3.0 mm</td>
<td>6%</td>
<td>35%</td>
<td>T21, 95.5%</td>
<td>4.5%</td>
<td>36%</td>
<td>Various</td>
</tr>
</tbody>
</table>

Footnotes: 1 = Trisomy
<table>
<thead>
<tr>
<th>Study Author, Date</th>
<th>Type of Study</th>
<th>Maternal Age</th>
<th>No. of Patients</th>
<th>Nuchal Translucency</th>
<th>% Estimated Abnormal</th>
<th>Risk</th>
<th>Detection Rate</th>
<th>Positive Predictive Value</th>
<th>Aneuploidies Detected</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brizot, 2001</td>
<td>Prospective</td>
<td>2.996</td>
<td>&gt; 95th percentile</td>
<td>5.8%</td>
<td>&gt; 1/300 in</td>
<td>90% T21</td>
<td>4% T21 (10 cases); screen-positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoppi, 2001</td>
<td>Prospective</td>
<td>10,001</td>
<td>singleton</td>
<td>&gt; 1.5 MoM</td>
<td>&gt; 1/300</td>
<td>90% T21</td>
<td>9% T21(10 cases)</td>
<td>screen-positive patients chose to -12 have invasive testing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gasiorek-Wiens, 2001</td>
<td>Prospective</td>
<td>23,805</td>
<td>singleton</td>
<td>&gt; 95th percentile 9.6%</td>
<td>&gt; 1/300</td>
<td>87.6% T21</td>
<td>13% T21, T18, Transvaginal US in screening self-selected percentile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O'Callaghan, 2000</td>
<td>Prospective</td>
<td>2000</td>
<td>screening</td>
<td>low and high-risk</td>
<td>&gt; 1/300</td>
<td>75% T21</td>
<td>121(8),T18(2),</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zoppi, 2000</td>
<td>Prospective</td>
<td>5,201</td>
<td>singleton</td>
<td>&gt; 1/300</td>
<td>80.8% T21</td>
<td>11.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwarzler, 1999</td>
<td>Prospective</td>
<td>4,523</td>
<td>&gt; 2.5 mm</td>
<td>5.1%</td>
<td>&gt; 1/300</td>
<td>78% all</td>
<td>4.7% T21, T18, Transvaginal US in screening unselected percentile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snijders, 1999</td>
<td>Prospective</td>
<td>100,311</td>
<td>&gt; 95th percentile</td>
<td>5.1%</td>
<td>&gt; 1/300</td>
<td>80.1% T13</td>
<td>T13 only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snijders, 1998</td>
<td>Prospective</td>
<td>96,127</td>
<td>singleton</td>
<td>&gt; 1/300</td>
<td>82.2% T21</td>
<td>8.3% T21, 99.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theodoropoulos, 1998</td>
<td>Prospective</td>
<td>3,550</td>
<td>singleton</td>
<td>&gt; 95th percentile</td>
<td>&gt; 1/300</td>
<td>91% T21</td>
<td>6.2% T21, others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sebire, 1996</td>
<td>Prospective</td>
<td>448 twins</td>
<td>&gt; 95th percentile</td>
<td>7.3%</td>
<td>&gt; 1/300</td>
<td>Twins B88% T21</td>
<td>5.4 - T21 only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pandya, 1995</td>
<td>Prospective</td>
<td>20,804</td>
<td>singleton</td>
<td>&gt; 95th percentile</td>
<td>&gt; 1/300</td>
<td>77% T21</td>
<td>T21, others</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Footnotes:**

MoM = Multiple of Median
NPV = Negative Predictive Value
T = Trisomy
<table>
<thead>
<tr>
<th>Study Author, Type of No., Type of Nuchal First-Trimester Estimated Sensitivity or False Positive Aneuploidies Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
</tr>
<tr>
<td>Date</td>
</tr>
<tr>
<td>Niemimaa, 2001 Prospective 2515 volunteers:</td>
</tr>
<tr>
<td>screening</td>
</tr>
<tr>
<td>Tsai, 2001 Retrospective 1514:</td>
</tr>
<tr>
<td>review</td>
</tr>
<tr>
<td>Spencer, 2000 Retrospective 4,190; singleton</td>
</tr>
<tr>
<td>review</td>
</tr>
<tr>
<td>Krantz, 2000 Prospective 5,809 MoM</td>
</tr>
<tr>
<td>screening</td>
</tr>
<tr>
<td>Age&gt;35, 92% T21 T21, age&gt;35 = 14.3%</td>
</tr>
<tr>
<td>Age&gt;35, 100% T18 T18, age&gt;35 = 0.4%</td>
</tr>
<tr>
<td>Age&gt;35, 100% T18 T18, age&gt;35 = 1.4%</td>
</tr>
<tr>
<td>Spencer, 2000a Retrospective 25 cases triploidy</td>
</tr>
<tr>
<td>Case-control</td>
</tr>
<tr>
<td>947 controls</td>
</tr>
<tr>
<td>Spencer, 2000b Retrospective 42 cases T13</td>
</tr>
<tr>
<td>Case-control</td>
</tr>
<tr>
<td>947 controls</td>
</tr>
<tr>
<td>Spencer, 2000c Prospective 159 twins;</td>
</tr>
<tr>
<td>screening</td>
</tr>
<tr>
<td>947 controls</td>
</tr>
<tr>
<td>Tul, 1999 Retrospective 50 cases T18</td>
</tr>
<tr>
<td>Case-control</td>
</tr>
<tr>
<td>947 controls</td>
</tr>
<tr>
<td>deGraaf, 1999 Retrospective 37 cases T21</td>
</tr>
<tr>
<td>Case-control</td>
</tr>
<tr>
<td>255 controls</td>
</tr>
<tr>
<td>Spencer, 1999 Retrospective 210 singleton T21</td>
</tr>
<tr>
<td>Case-control</td>
</tr>
<tr>
<td>946 controls</td>
</tr>
<tr>
<td>Study Author, Type of No., Type of Nuchal First-Trimester Estimated Sensitivity or False Positive Aneuploidies Comments</td>
</tr>
<tr>
<td>Date</td>
</tr>
<tr>
<td>Date</td>
</tr>
<tr>
<td>DelBiasio, 1999 Prospective 1,467; singleton</td>
</tr>
<tr>
<td>screening</td>
</tr>
<tr>
<td>946 controls</td>
</tr>
</tbody>
</table>

Footnotes:  
AFP = Alpha fetoprotein  
BHCG = Beta human chorionic gonadotropin  
MoM = Multiple of the expected median  
NT = Nuchal translucency measurement  
BT = Biochemical test  
PAPP-A = Pregnancy associated plasma protein A  
FBHC = Free beta human chorionic gonadotropin  
T = Trisomy
<table>
<thead>
<tr>
<th>Study Author, Date</th>
<th>Type of Study</th>
<th>No., Type of Population</th>
<th>Translucency Risk</th>
<th>First-Trimester Biochemical Test (BT)</th>
<th>Estimated Cut-off</th>
<th>Estimated Rate</th>
<th>Sensitivity or Detection Rate</th>
<th>False Positive Detection Rate</th>
<th>Positive Aneuploidies</th>
<th>Predictive Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benattar, 1999</td>
<td>Prospective</td>
<td>1,656; singleton</td>
<td>MoM FBHCG, &gt; 1/250BT</td>
<td>BT - 80% any</td>
<td>5%</td>
<td>T21, T18</td>
<td>Triology</td>
<td>4.5%</td>
<td>T21, T18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biagiotti, 1998</td>
<td>Retrospective</td>
<td>32 cases T21</td>
<td>MoM PAPP-A</td>
<td>BT - 35-49.5% T21</td>
<td>5%</td>
<td>T21</td>
<td>Combined with maternal age</td>
<td>4%</td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimmerman, 1996</td>
<td>Retrospective</td>
<td>1,151; high risk = 3.0mm</td>
<td>MoM PAPP-A</td>
<td>NT - 38.1% any</td>
<td>2%</td>
<td>T21, T18</td>
<td>Triploidy</td>
<td>T13, X0</td>
<td>T21, T18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noble, 1995</td>
<td>Prospective</td>
<td>2,520x &gt; 99th percentile</td>
<td>MoM PAPP-A</td>
<td>NT - 20% T21</td>
<td>5%</td>
<td>T21</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muller, 2003</td>
<td>Prospective</td>
<td>5,664; singleton</td>
<td>MoM PAPP-A</td>
<td>NT - 80%</td>
<td>4.70%</td>
<td>T21</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spencer, 2003</td>
<td>Prospective</td>
<td>12,334; singleton</td>
<td>MoM PAPP-A</td>
<td>NT - 92%</td>
<td>5.20%</td>
<td>T13</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VonKuisenberg, 2002</td>
<td>Prospective</td>
<td>3,864; &gt; 99th percentile</td>
<td>MoM PAPP-A</td>
<td>NT - 84%</td>
<td>5%</td>
<td>T21</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossley, 2002</td>
<td>Prospective</td>
<td>17,229; MoM</td>
<td>MoM PAPP-A</td>
<td>NT - 82%</td>
<td>5%</td>
<td>T21</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wapner, 2003</td>
<td>Prospective</td>
<td>8,514; singleton</td>
<td>MoM PAPP-A</td>
<td>NT - 78.7%</td>
<td>5%</td>
<td>T21</td>
<td></td>
<td></td>
<td>T21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Footnotes:**
- AFP = Alpha fetoprotein
- MoM = Multiple of the expected median
- BHCG = Beta human chorionic gonadotropin
- NT = Nuchal translucency measurement
- BT = Biochemical test
- PAPP-A = Pregnancy associated plasma protein A
<table>
<thead>
<tr>
<th>Study Author, Date</th>
<th>Study Type</th>
<th>Population</th>
<th>Type of Nuchal Translucency</th>
<th>Second-Trimester Biochemical Test</th>
<th>Estimated Cut-off Value</th>
<th>Risk of False Positive Invasive Aneuploidies</th>
<th>Screening Procedure</th>
<th>Invasive Aneuploidies Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michaelidis, 2001</td>
<td>Retrospective</td>
<td>7,447 unselected</td>
<td>&gt; 99th NT percentile</td>
<td>AFP, NT</td>
<td>NT &gt; 99%</td>
<td>BT &gt; 0.5%</td>
<td>6.2%</td>
<td>T21</td>
</tr>
<tr>
<td>Audibert, 2001</td>
<td>Prospective</td>
<td>4,130; singleton, age &lt; 38</td>
<td>&gt; 95th percentile</td>
<td>AFP, NT</td>
<td>NT &gt; 95%</td>
<td>BT &gt; 60%</td>
<td>4.6%</td>
<td>T21</td>
</tr>
<tr>
<td>Schuchter, 2001</td>
<td>Retrospective</td>
<td>9,342; singleton, age &lt; 38</td>
<td>≥ 2.5 mm NT percentile</td>
<td>AFP, NT</td>
<td>NT &gt; 2.5 mm</td>
<td>BT &gt; 60%</td>
<td>8.2%</td>
<td>T21</td>
</tr>
<tr>
<td>Rozenberg, 2002</td>
<td>Prospective</td>
<td>9,444</td>
<td>≥ 3 mm NT percentile</td>
<td>AFP, NT</td>
<td>NT &gt; 3 mm</td>
<td>BT &gt; 60%</td>
<td>8.6%</td>
<td>T21</td>
</tr>
<tr>
<td>Lam, 2002</td>
<td>Prospective</td>
<td>17,590 singleton</td>
<td>≥ 3 mm NT percentile</td>
<td>AFP, NT</td>
<td>NT &gt; 3 mm</td>
<td>BT &gt; 60%</td>
<td>8.6%</td>
<td>T21</td>
</tr>
<tr>
<td>Cicero, 2003</td>
<td>Retrospective</td>
<td>100 cases (121)</td>
<td>MoM</td>
<td>PAPP-A, PAPPA-A</td>
<td>NT + BT &gt; 90.5%</td>
<td>NT + BT &gt; 12%</td>
<td>T21</td>
<td></td>
</tr>
</tbody>
</table>

Footnotes:  
AFP = Alpha fetoprotein  
BHCG = Beta human chorionic gonadotropin  
BT = Biochemical test  
E3 = Unconjugated estriol  
NT = Nuchal Translucency measurement  
T = Trisomy
Studies of Nuchal Translucency Impact on Clinical Decision-Making

There have been no studies published regarding the utility of nuchal translucency testing on clinicians' therapeutic decision-making for pregnant women. Thus, for this diagnostic test, there is not yet evidence that use of the test results has improved medical management in a way that will benefit the patient.

Other Issues

Criticisms about the design of many of the earlier published studies of nuchal translucency measurement have focused on verification bias and selection bias (Malone et al., 1998; Mol et al., 1998; Mol et al., 1999, Malone et al., 2000; Souter et al., 2001). In some studies, fetuses with an increased nuchal translucency thickness were more likely to undergo fetal karyotyping than those with a normal nuchal translucency thickness. This results in verification bias. Down syndrome fetuses with an increased nuchal translucency thickness were almost always detected, whereas the false-negatives (i.e., Down syndrome fetuses with a normal nuchal translucency thickness) had only about a 50% chance of being missed due to fetal loss. This type of bias in the evaluation of a diagnostic test, known as verification bias (or ascertainment bias), occurs when the selection for verification of the diagnosis depends on the results of the test being studied. In the case of nuchal translucency measurement, verification bias is likely to cause an overestimation of the detection rate or test sensitivity. Most studies performed in low-risk populations suffer from this verification bias because fetal karyotyping is restricted to women with an increased fetal translucency thickness (Mol et al., 1999); in high-risk populations, karyotyping is more broadly applied. Mol et al. (1999) examined 25 studies for the presence of verification bias and evaluated its effect on the accuracy of first-trimester nuchal translucency measurement for Down syndrome detection. The authors found 10 studies with verification bias and 15 studies without verification bias. Studies with verification bias reported higher test sensitivities and slightly higher test specificities of nuchal translucency measurement than studies without verification bias. For studies with verification bias, adjusted estimates of the sensitivity were calculated assuming a fetal loss rate for Down syndrome pregnancies of 48%. The sample-size-weighted sensitivity was 55% in studies without verification bias and 77% in those with verification bias, for specificities of 96% and 97%, respectively. After adjustment for verification bias, the sample-size-weighted sensitivity decreased from 77% to 63%.

Several of the published studies have been affected by selection bias (Hewitt et al., 1996; Zimmerman et al., 1996; Orlandi et al., 1997). When selection bias is present, fetuses with a normal nuchal translucency are less likely to be included in the study. Selection bias results in an overestimation of the sensitivity and underestimation of the specificity of the test (Mol et al., 1999).
Advocates of first trimester screening such as Wapner et al. (2003) argue that early diagnosis allows women greater privacy and safer termination of affected pregnancies. However, there is a slightly higher risk of pregnancy loss with chorionic-villus sampling than with second-trimester amniocentesis. In addition, only half the women who elected to terminate the pregnancy in the BUN study did so before 16 weeks of gestation. Thus, it is not clear how often women who undergo first-trimester screening will realize the benefits ascribed to this approach (Mennuti and Driscoll, 2003).

Mennuti and Driscoll (2003) point out in an editorial accompanying the BUN paper, that although the sensitivity of first-trimester screening was higher in this study and in some others that have been reported, the confidence intervals overlap the sensitivity of second-trimester screening involving the measurement of four serum markers. In addition, first-trimester screening is likely to appear to be more sensitive than second-trimester screening in many of the earlier studies because there is a higher prevalence of Down's syndrome in the first trimester. Wapner et al. corrected for this by using published rates of pregnancy loss involving fetuses with Down's syndrome.

**Technical Considerations**

The performance of NT as a screening marker has not been consistent from study to study, presumably because of variability in operator expertise and quality of equipment. Guidelines for the measurement of NT have been suggested to achieve uniformity of results among different operators (Nicolaides et al., 2002). The authors conclude that, "Appropriate training, high motivation and adherence to a standard technique for the measurement of NT are essential prerequisites for good clinical practice" (Nicolaides et al., 2002). Similarly, Chasen and Skupski (2003) assert that although NT should not be considered investigational, it cannot be considered to be the standard of care at this time since: "Without quality testing in experienced centers, the harms of (NT) testing may outweigh the benefits." They assert that NT should not replace second trimester screening until more information comparing the performance of these two approaches is available (Chasen and Skupski, 2003).

Another potential problem with the widespread adoption of NT outside of specialized centers is how to express patient specific risks with NT. An abnormal NT was initially expressed in reference to a fixed cutoff value of 3.0 mm, and many of the earlier studies of NT presented their results in this way. Nicolaiides (1992) and others established that in normal pregnancies NT increases with gestation, so a fixed cut-off provided inaccurate patient-specific risk estimates. Most recent published studies therefore, including Wapner et al., 2003, have converted NT measures into a multiple of the median (MoM) of unaffected pregnancies of the same gestational age, similar to the approach used in screening using maternal serum biochemical markers. However, a recent paper (Spencer et al., 2003), co-authored by Dr. Nicolaides, seems to cast doubt on this approach. In this review of the most appropriate method for calculating accurate patient-
specific risks for trisomy-21 in the first trimester, they concluded that: "the NT MoM approach provides inaccurate individual patient specific risks for trisomy-21 and therefore the use of this method is inappropriate." Instead, they advocate using the delta-NT approach in which patient specific risks for trisomy-21 are calculated by multiplying the a priori maternal age risk with the likelihood ratio of the observed delta-NT. The delta-NT is the difference from the normal median NT at the measured crown-rump length.

In addition, there is some discussion in the literature as to the need to take ethnic origin into account in the interpretation of NT measurements (Snijders and Smith, 2002) and what the effect of image size is on NT measurement. In a recent paper, Edwards et al., (2003) concluded that the measurement of NT decreases significantly with increasing image size and that optimization of NT as a method of screening will, "require agreed standardization of image modification." Finally, the SURUSS trial concluded that the make and model of the ultrasound machine has an important influence on obtaining satisfactory NT measurement.

**Patient Risks**

Transabdominal ultrasonography is considered safe in pregnancy (Taipale et al., 1997). No complications of pregnancy have been noted in the various studies, even those using a vaginal probe (Taipale et al., 1997).

However, there are potential risks to NT screening. NT images must be obtained by properly trained ultrasonographers, using standardized techniques and appropriate ultrasound equipment. Data must be audited periodically to ensure that quality standards are upheld. Failure to meticulously adhere to these guidelines may result in higher false-positive and negative rates that in turn can lead to higher rates of invasive testing and miscarriage or missed diagnoses (Chasen and Skupski, 2003).

Though both CVS and amniocentesis are relatively safe, the risk of CVS is somewhat higher than that of amniocentesis. Between 0.5% and 1% of pregnancies are lost as a complication of first-trimester CVS, whereas less than one in 300 second-trimester amniocenteses results in fetal loss. The risk of "early amniocentesis" performed during gestational weeks 12-14, is similar to that of CVS (Pyeritz, 2002).

**Pending Trials**

The First- and Second-Trimester Evaluation of Risk for Aneuploidy (FASTER) trial is being conducted at 15 clinical centers in the U.S. Data collection ran from 10/99 to 12/02 and the data is now being analyzed and will be presented at the Society of Maternal and Fetal Medicine in early February 2004. A total of 38,000 patients with singleton pregnancies were recruited at 10-13 weeks gestation confirmed by ultrasound. Patients were screened in the first trimester with NT plus PAPP-A and free beta-hCG, and asked to return
between 15 and 18 weeks for the second trimester quadruple marker test (AFP, unconjugated estriol, free beta-hCG and inhibin A). Investigators were blinded to the results of the first trimester screening, so clinical decisions about intervention were not addressed until the second trimester. This design obviates the problem of verification bias discussed above, but does not allow this study to comment on the value of NT in decision-making in first trimester screening. Unfortunately, a different method for measuring the size of the NT was used than has been customary in most other studies so their results may be difficult to compare with this literature.

In sum, there is little doubt that with appropriate training and oversight in the context of a well-designed trial such as the BUN trial (Wapner et al., 2003), NT has been shown to improve the detection rate of fetal aneuploidy such as T-21 without an increase in the false positive rate. SURUSS demonstrates the value of adding NT to biochemical testing across the first and second trimesters. However, the unresolved technical issues and the paucity of data that examines the utility of NT testing on clinician’s therapeutic decision-making make it impossible to conclude that NT has been shown to improve the net health outcomes outside of the specialized settings.

TA Criterion 3 is not met.

**TA Criterion 4: The technology must be as beneficial as any established alternatives.**

Alternatives to fetal nuchal translucency measurement to select high-risk pregnancies for invasive testing (CVS or amniocentesis) are techniques based on maternal age and results of second or first-trimester maternal serum biochemical tests. The sensitivity for screening based on maternal age alone is approximately 20% to 30% and is not considered a reasonable alternative to the use of multiple maternal serum analyses.

The American College of Obstetricians and Gynecologists (ACOG) recommends that all women under the age of 35 be offered serum screening with multiple markers, and that women over the age of 35 be offered an invasive diagnostic procedure at delivery.

Results of marker testing are usually reported as both actual concentration and then converted to the gestation specific multiple of the median (MoM) since these values are constantly changing over the period of gestation in which screening is done. An MoM is defined as the ratio of an individual marker concentration to the median concentration expected for the population of women at the same gestational age. By definition, the central value in unaffected singleton pregnancy is 1.0 MoM. An algorithm that combines maternal age with MoMs can be used to calculate the risk for such chromosomal disorders as
trisomy 21 and 18. In addition, an individual's MoM for a particular marker may be adjusted by certain factors such as race, body weight, diabetes mellitus and multiple gestations that affect their concentrations.

Prospective studies have confirmed the efficacy of a triple-test screening algorithm for Down's syndrome that measures maternal serum AFP, free beta-hCG and uE3 in the second trimester (gestational age 15-22 weeks). The second-trimester "triple screen" when accompanied by ultrasound dating of gestational age yields a detection rate of approximately 70% - 80% with a 5% - 8% false-positive rate (Benn, 2003; Cuckle, 2001). Other studies have found a 69% detection with a 5% false positive rate (Wald et al., 1999).

The best available second trimester serum screening panel is the quadruple marker test. The quadruple test for Down's syndrome calculates the risk of a Down's syndrome term pregnancy from maternal age at term and the concentration of four markers in maternal serum-AFP, unconjugated estriol, free beta-hCG and inhibin A at 15-22 weeks of pregnancy. The detection rate in one recent study was 82% (95% CI 72%-89%) with a false positive rate of 7% (Wald et al., 2003).

First-trimester serum screening is not as effective as standard second-trimester screening. The serum markers used are the free beta-hCG (also a good second trimester marker) and PAPP-A, a complex, high molecular weight glycoprotein of uncertain function. Levels of free beta-hCG are, on average, almost two times higher, while levels of PAPP-A are, on average, almost 2.5 times lower in Down syndrome compared to unaffected pregnancy (Wald et al., 1996). The two serum markers in combination with maternal age constitute a first-trimester maternal serum screening test that achieves 63 percent detection at a five percent false positive rate (Canick and Kellner, 1999).

Current methods of first or second-trimester screening yield a relatively high level of detection with a false positive rate of 5%. The goal of current combined screening tests that include NT are to further reduce the false positives while maintaining a high detection rate. Fewer false positive results will result in fewer instances of fetal loss due to unnecessary invasive screening.

As reported in the BUN study (Wapner et al., 2003) and in SURUSS (Wald et al., 2003), the addition of NT to first and/or second trimester serum markers has the potential to achieve equal or better detection rates to the triple or quadruple tests that are the current standards of care. However, NT measurement is highly operator-dependent. It requires training, external quality control and adequate time to allow accurate measurement; otherwise sub optimal performance will result (Crossley et al., 2002). In addition, the performance of NT as a screening marker has not been consistent from study to study, presumably because of variability in operator expertise and quality of equipment. While few experts question the validity of the association between increased NT and fetal aneuploidy, the published literature reveals a lack of congruence over a standard and reproducible method for measuring NT. In addition, there has been
insufficient attention paid to the impact of NT on individual decision-making by patients. Further study as well as proper training and ongoing quality management are necessary before we can conclude that nuchal translucency screening improves net health outcomes as much as or more than the established alternatives of the second-trimester "triple screen" and the "quadruple screen."

**TA criterion 4 is not met.**

**TA Criterion 5: The improvement must be attainable outside the investigational setting.**

Nuchal translucency testing has been evaluated in multiple centers. Similar results can be anticipated in other obstetric settings, provided uniformly trained operators, with adequate technical experience, using the appropriate equipment do ultrasound studies and with ongoing quality assurance in place.

**TA criterion 5 is met.**
RECOMMENDATIONS OF OTHERS

Blue Cross Blue Shield Association (BCBSA)

The BCBSA Technology Evaluation Center has not reviewed this topic.

American College of Obstetrics and Gynecology (ACOG)

In October 1999, ACOG’s Committee on Genetics developed the following committee opinion: "First-trimester screening for fetal chromosome, cardiac and other abnormalities using the nuchal translucency marker alone or in combination with serum markers appears promising but remains investigational…until further studies confirm the efficacy of first-trimester nuchal translucency screening, with or without serum markers, this modality is not recommended for routine clinical use" (ACOG, 2000).
CONCLUSION

Down syndrome is characterized by moderate to severe learning disabilities and cardiac defects in 40%-50% of those affected. It is the most commonly recognized genetic cause of mental retardation with an estimated prevalence of 9.2 cases per 100,000 live births in the United States (Newberger, 2000). The incidence of fetal trisomies is directly related to maternal age. The risk of having a child with Down syndrome is 1/1,300 for a 25 year-old woman; at age 35 the risk increases to 1/365 (Newberger, 2000). In fact, more than 70% of infants with Down syndrome are born to women less than 35 years old. This is largely because, in most populations, women under 35 years of age bare 85%-90% of babies. Consequently, while women over the age of 35 have a higher individual risk for chromosomal abnormalities in the fetus; the larger number of births in the under-35 age category offsets this. Since the risk from amniocentesis in young women was greater than the risk of trisomy, however, most of these women were not even eligible for amniocentesis.

Definitive prenatal diagnosis of trisomy 21 requires cytogenetic analysis of cells obtained by two main invasive procedures: chorionic villus sampling (CVS) and second trimester amniocentesis. CVS offers the opportunity for first-trimester diagnosis when elective pregnancy termination carries the lowest risk of maternal morbidity and is associated with a risk of fetal loss of 0.5% to 1.5%. Amniocentesis carries a somewhat lower risk of fetal loss of 0.5% to 1.0% (Benn, 2002).

First-trimester screening using a combination of nuchal translucency measurement and maternal biochemical tests is feasible, and holds promise to improve Down syndrome detection compared with currently used second-trimester or first-trimester biochemical testing protocols. Such testing might provide substantial advantages to clinicians and patients. (A potential disadvantage, however, is that earlier screening preferentially identifies those chromosomally abnormal pregnancies that are destined to miscarry (Nicolaides et al., 2002)). Compared with second-trimester screening for chromosomal abnormalities, first-trimester screening offers couples greater privacy, more time for the patients to evaluate their options and to make decisions, and allows for earlier and safer methods of pregnancy termination. Women with positive screening results can undergo chorionic villus sampling (CVS) to confirm the results, rather than wait to undergo amniocentesis later in pregnancy. However, fewer than half of the women in the BUN study who carried fetuses with trisomy 21 and who chose to terminate their pregnancies did so before 15 weeks' gestation. This observation suggests that limited access to both CVS and abortion might reduce the benefits associated with earlier screening.

Current methods of first or second trimester screening yield a relatively high level of detection (75-85%) with a false positive rate of 5%. The goal of combined screening tests that include NT is to further reduce the false positives while maintaining a high detection rate. Fewer false positive results will result in fewer
instances of fetal loss due to unnecessary invasive screening. With an almost 80% detection rate (with 5% false positive), the BUN trial provides evidence that this is an attainable goal.

However, the performance of NT has not been consistent across all studies. Adequate training and ongoing quality management will be key to the more widespread adoption of NT. Until this is assured, the use of NT should be restricted to high volume centers with the appropriate expertise in the technology as well as adequate resources to provide genetic counseling for patients with abnormal results.

In addition, some experts express concerns regarding standard measurement of fetal-nuchal translucency. In selected centers, first-trimester screening can be (and has been) implemented. But for many women, comprehensive prenatal care does not begin until the second trimester and early referral of these women to specialized units would require a significant change from existing clinical management. Second-trimester screening should remain the standard of care, pending results from other studies and revision of practice guidelines.

Results from SURUSS suggest that integrating the results of first- and second-trimester screening would be more sensitive and result in fewer diagnostic tests and fewer procedure-related losses of normal pregnancies than the use of either method alone. The question remains, however, if this approach were widely adopted would it better serve women and couples confronted with decisions about screening? Would women with positive first-trimester screening tests await the results of second-trimester screening or opt for immediate diagnosis? Should we offer women who have negative results on first-trimester screening the option of second-trimester screening (Menutti and Driscoll, 2003)?

Although few experts question the validity of the association between increased NT and fetal aneuploidy, the published literature reveals a lack of congruence over a standard, reproducible method for measuring NT. In addition, there has been insufficient attention paid to the impact of NT on individual decision-making by patients. The ongoing FASTER trial holds promise to address some of these issues, though as it is using a different methodology for measuring and reporting NT than was used in the BUN trial (personal communication, K. Filkins, MD), it may only add to the uncertainty when it is published.

Only with the adoption of uniform methodology and the establishment of international standards for NT measurement is the true potential of this test likely to be realized (Borruto et al., 2002). Guidelines for the measurement of NT have been suggested to achieve uniformity of results among different operators (Nicolaides et al., 2002). In this paper the authors conclude that: "Appropriate training, high motivation and adherence to a standard technique for the measurement of NT are essential prerequisites for good clinical practice." Similarly, Chasen and Skupski (2003) assert that although NT should not be considered investigational, it cannot be considered to be the standard of care at this time since: "Without quality testing
in experienced centers, the harms of (NT) testing may outweigh the benefits." In a recent comprehensive review of first trimester screening for Down Syndrome, Malone and D'Alton (2003) argue that a number of implementation issues need to be addressed before first first-trimester screening can be endorsed for use in clinical practice. These issues include: 1) quality control (initial training, adequate monitoring systems); 2) inconsistent interpretation of NT at different centers; 3) the potential negative impact on second-trimester screening; 4) the lack of widespread availability of CVS; 5) the need for accurate comparative data to evaluate the best combination of tests to implement into practice; and 6) the need for appropriate patient counseling. In conclusion, nuchal translucency with appropriate biochemical markers should not replace traditional second trimester screening until these issues are addressed and more information comparing the performance of these two approaches is available.

TA criteria 3 and 4 are not met.

RECOMMENDATION

It is recommended that:

Ultrasound nuchal translucency screening for fetal aneuploidy does not meet CTAF criteria.

*The California Technology Assessment Forum approved the following alternative recommendation:*

*Ultrasound nuchal translucency with biochemical screening for fetal aneuploidy meets technology assessment criteria and should be restricted to centers meeting California Department of Health Services designation for perinatal diagnostic centers.*

February 11, 2004
REFERENCES


2. ACOG Committee Opinion: Committee on Obstetric Practice: Down syndrome screening: 1994, No. 141 (Replaces No. 76, 1989)


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


53 Eustace DL. Screening for nuchal translucency. Screening provides reliable information on which care can be based. BMJ. 1998 Sep 12;317(7160):748.


REFERENCES (continued)


69 Hampton K. Screening for nuchal translucency. Women are unaware that they can refuse an ultrasound scan. BMJ. 1998 Sep 12;317(7160):748-9.


REFERENCES (continued)


90 Layng J. Screening for nuchal translucency. Counseling should be considered an integral part of screening programmes. BMJ. 1998 Sep 12;317(7160):749.


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


REFERENCES (continued)


