ORIGINAL ARTICLE

The Diagnostic Performance of Nuchal Translucency Alone as a Screening Test for Down Syndrome: A Systematic Review and Meta-analysis

Ma. Sergia Fatima P. Sucaldito, MD, John Jefferson V. Besa, MD and Lia M. Palileo-Villanueva, MD, MSc

Department of Medicine, University of the Philippines College of Medicine

ABSTRACT

Background. Down syndrome or trisomy 21, the most common chromosomal disorder, results from the presence of a third copy of chromosome 21 and manifests as mild to moderate intellectual disability, growth retardation, congenital heart defects, gastrointestinal abnormalities, and characteristic facial features. Several methods have been used to screen for Down syndrome in the prenatal period, such as ultrasound, biomarkers, cell-free DNA testing, and combinations of these tests. A positive result from one or more of these screening tests signals the need for confirmatory karyotyping to clinch the diagnosis. Ultrasound between 11 to 14 weeks of gestation can evaluate nuchal translucency (NT) to screen for Down syndrome. During the second trimester, a triple or quadruple test can also be performed alone or in addition to NT to quantify Down syndrome risk. In limited resource settings however, only the measurement of NT via ultrasound can be performed since biomarker tests are either unavailable or inaccessible. While the diagnostic performance of NT measurement alone has been investigated in several observational studies, there is no consensus on its performance as a sole test to screen for Down syndrome.

Objective. To determine the diagnostic performance of NT during prenatal first-trimester ultrasound as a screening test for Down syndrome.

Methods. We performed a systematic search on the PubMed, ProQuest, and Cochrane Library databases for recent systematic reviews and meta-analyses that addressed the objective. The existing reviews found were then independently appraised by the two reviewers with the AMSTAR-2 checklist. To update the existing reviews, a systematic search was done in the same databases to identify additional primary diagnostic studies, which were appraised using the QUADAS-2 tool. Random-effects univariate meta-analysis and summary receiving operator curve (HSROC) analysis for the outcomes were performed using Review Manager version 5.4 and R version 4.2.2, respectively. Subgroup analysis was performed by stratifying the baseline risk of mothers for fetal anomaly as low- or high-risk. High-risk mothers were defined as women with risk factors such as advanced age, positive serum screen, presence of other ultrasound anomalies, and history of previous fetus with anomaly.

Results. We found 22 cohort studies (n=225,846) of women at low-risk for fetal anomaly. The pooled sensitivity was 67.8% (95% CI: 61.4%-73.6%, I^2=70.4%) and specificity was 96.3% (95% CI: 95.5%-96.9%,

Corresponding author: Ma. Sergia Fatima P. Sucaldito, MD
Department of Medicine
Philippine General Hospital
University of the Philippines Manila
Taft Avenue, Ermita, Manila 1000, Philippines
Email: maserfa11@gmail.com

I^2=96.7%). For low-risk women, the overall certainty of evidence was low, due to different modes of verification and heterogeneity not completely explained by variability in baseline risk or cut-points. Seven studies (n=9,197) were on high-risk women. The pooled sensitivity was 62.2% (95% CI: 54.1%-69.7%, I^2=38.8%) and specificity was 96.5% (95% CI: 93.6%-98.1%, I^2=95.5%). For women at high-risk, the evidence was rated as moderate due to differential verification.

Conclusion. Our analysis showed that NT measured through first-trimester ultrasound is specific for Down syndrome but has low sensitivity. Despite this, it is a useful screening test for Down syndrome in low-resource settings where other strategies may not be available or accessible. Furthermore, interpretation of NT results must take into consideration its limited sensitivity as this may lead to missed cases.

Keywords: nuchal translucency measurement, Down syndrome, sensitivity and specificity

INTRODUCTION

Down syndrome or trisomy 21, the most common chromosomal disorder, results from the presence of a third copy of chromosome 21. The risk of Down syndrome increases with increasing maternal age and manifests as mild to moderate intellectual disability, growth retardation, congenital heart defects, gastrointestinal abnormalities, and characteristic facial features.[1] As many as 40% of infants with Down syndrome present with at least one cardiac defect. The estimated worldwide incidence of Down syndrome is between 1 in 1,000 to 1 in 1,100 live births, which translates to 3,000 to 5,000 cases annually.[2] In the Philippines, 1 out of 800 babies is born with Down syndrome, and it is the top reason for genetic consultation locally.[3] Due to recent medical advancements, there has been a significant increase in the survival of infants and life expectancy of patients with Down syndrome. Between 1979 and 2003, the rate of death among infants with Down syndrome during the first year of life decreased from 8.5% to 5%.[4] Furthermore, their median life expectancy is now 58 years compared to only about 10 years in 1960.[5]

Because it is an important genetic cause of morbidity and mortality, several methods have been used for the prenatal diagnosis of Down syndrome. Ultrasound between 11 to 14 weeks of gestation can evaluate nuchal translucency (NT), a transient subcutaneous collection of fluid behind the fetal neck.[6] Another means is through a combined test during the first trimester which involves measurement of NT and maternal serum beta-human chorionic gonadotropin (beta-hCG) plus pregnancy-associated plasma protein A (PAPP-A). During the second trimester, a triple or quadruple test can be performed through measurement of maternal biochemical markers, specifically alpha-fetoprotein (AFP), unconjugated estriol, beta-hCG, and inhibin-A. If the mother receives a positive result based on one or a combination of these screening tests, further procedures including chorionic villus sampling (CVS) and amniocentesis are offered and the samples obtained are sent for karyotyping, which is the confirmatory test.[5] Once the diagnosis is confirmed, parents will receive counseling and offered the option of pregnancy termination in countries where this is available.

Although the most accurate diagnostic tests are CVS and amniocentesis, these are invasive and known to carry between 0.5 to 1% risk of miscarriage, hence non-invasive tests such as those described above are considered first-line tests for screening.[6] In limited resource settings, only measurement of NT via ultrasound can be performed since biomarker or combined tests are unavailable or inaccessible. The diagnostic accuracy of NT measurement alone has been investigated in several observational studies, but there is no consensus on its use as the sole diagnostic tool to screen for Down syndrome.

OBJECTIVE

This study determined the diagnostic performance of nuchal translucency prenatal ultrasound at 11 to 14 weeks as a screening test for Down syndrome. Diagnostic performance was described in terms of sensitivity and specificity.

METHODS

We performed a search on January 10, 2024 on the PubMed,ProQuest, and the Cochrane Central Register of Controlled Trials (CENTRAL) for systematic reviews and meta-analyses that involved the use of NT alone as measured during prenatal ultrasound done at 11 to 14 weeks as a screening test for Down syndrome among pregnant women. The inclusion criteria for this stage of the literature search were: (1) any systematic review or meta-analysis reporting the diagnostic performance of NT alone as a screening test for Down syndrome, (2) must contain enough data for extrapolation of 2x2 table, (3) must mention karyotyping as the reference standard. Exclusion criteria included use of NT in combination with other tests, lack of numeric data for 2x2 table, not reporting Down syndrome as a single outcome, and not using karyotyping as the reference standard. No language or time of publication restrictions were used. Two investigators screened the results independently. Of the 61 articles retrieved, two studies were included. These studies were then independently appraised by the two reviewers with the AMSTAR-2 checklist.

The search for meta-analyses and systematic reviews yielded one Cochrane review on the utility of NT for Down syndrome screening done by Alldred et al. in 2017, which was assessed as a high-quality review, and another performed by Liu et al. in 2015, which was evaluated as low-quality due to the lack of protocol and funding details.[7,8] The Cochrane

review by Alldred was published in 2017 but contained only studies until August 25, 2011, which numbered 18 studies. We decided to adapt this review and update the search. The study by Liu et al. was also screened for additional relevant non-duplicate articles, which yielded five additional unique studies.

Another search was performed on February 1, 2024 using the above databases for additional primary studies on the outcome of Down syndrome from August 25, 2011 until January 31, 2024. The full search strategy is available in Appendix A. The inclusion criteria for this update were: (1) any observational studies reporting on the diagnostic performance of NT as a screening test for Down syndrome, (2) must contain data in tables, text, or supplementary material allowing for the extrapolation of a 2x2 table, (3) must use karyotyping as the confirmatory reference standard, and (4) published on or after January 1, 2011. Two reviewers independently performed the search, selected studies, and collected the following information from the selected studies: baseline population risk for fetal anomaly, index test, reference test, and the participants' results for the index and reference standard as aggregate values for 2x2 tables. Study characteristics are available in Appendix B. The outcome assessed was diagnostic performance in detecting Down syndrome in terms of sensitivity and specificity. The reviewers also independently assessed the risk of bias in these diagnostic studies using the QUADAS-2 tool. The certainty of evidence was described through the GRADE approach. A third reviewer served to reconcile any disagreement.

Random-effects univariate meta-analysis and hierarchical summary receiving operator curve (HSROC) analysis for the outcomes were performed using Review Manager version 5.4 and R version 4.2.2, respectively. Pooled sensitivity and specificity values were reported with 95% confidence intervals along with I^2 values to reflect heterogeneity. Subgroup analysis was performed by stratifying the baseline risk of mothers for fetal anomaly as low- or high-risk. High-risk mothers were those with any of the following risk factors: advanced age, positive serum screen, presence of other ultrasound anomalies, and history of previous fetus with anomaly. Low-risk mothers are those without any of the risk factors above. Another foreseen source of heterogeneity was the variation in the cut-off points for a positive NT screen, hence an additional subgroup analysis was performed according to cut-off value (Appendix C).

RESULTS

We included 23 relevant studies from previous reviews by Alldred et al. and Liu et al. Of the 1,219 articles yielded by the search strategy adapted from Alldred, we selected a total of six new additional cohort studies (Figure 1). The 1,213 articles were excluded during the screening process due to either combined use of NT with other modalities such as biomarkers or cell-free DNA, unavailability of the numeric values needed for a 2x2 table for diagnostic accuracy calculations, or lack of karyotyping as the confirmatory reference standard. Three out of six of these studies had

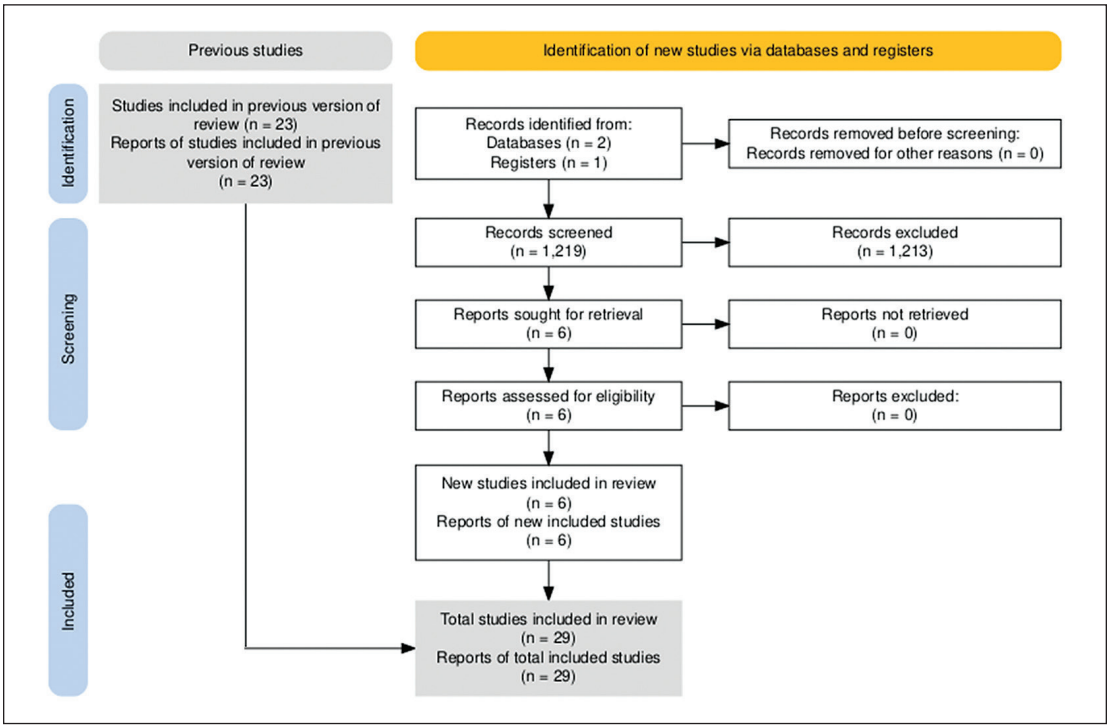

Figure 1. PRISMA diagram for study identification and screening.

a moderate risk of bias in QUADAS-2 due to differential verification of the Down syndrome diagnosis, which could impact the diagnostic characteristics of the test.[9–11] The other three studies were deemed to have low-risk of bias.[12–14] More detailed characteristics of these studies as well as their QUADAS assessments are available in the Appendix B Risk of Bias table. Altogether, we included a total of 29 cohort studies in this evidence review.

Twenty-two cohort studies (n=225,846) recruited participants at low-risk for fetal anomaly.[9–30] Of these studies, 14 studies were prospective and eight studies were etrospective. Participants underwent NT screening at various cut-offs as shown in Figure 2. The procedure for NT measurement was standardized similarly in all the studies through Fetal Medicine Foundation (FMF) certification of the sonographers. The pooled sensitivity was 67.8% (95% CI: 61.4%-73.6%, I^2=70.4%) and specificity was 96.3% (95% CI: 95.5%-96.9%, I^2=96.7%). Additional subgroup analyses based on cut-off points are also available in Appendix C.

Seven cohort studies (n=9,197) recruited participants at high-risk for fetal anomaly.[31–37] Of these studies, five studies were prospective and two studies were retrospective. Participants underwent NT screening at various cut-offs as shown in Figure 3. The pooled sensitivity was 62.2% (95% CI: 54.1%-69.7%, I^2=38.8%) and specificity was 96.5% (95% CI: 93.6%-98.1%, I^2=95.5%). Sensitivity analysis was done by excluding the outlier Acacio et al., but heterogeneity remained substantial with sensitivity values at 61.2% (95% CI: 53.1%-68.9%, I2=44.2%) and specificity at 97.2% (95% CI: 95.7%-98.2%, I2=91.7%). Additional subgroup analysis based on cut-off points are available in Appendix C.

In Figure 4, the HSROCs for various cut-points of NT are shown, which also mirror the sensitivity and specificity values calculated through univariate analysis.

For the outcome of Down syndrome among low-risk pregnant women, the overall certainty of evidence was rated down once for different modes of verification and once for heterogeneity not completely explained by variability in baseline risk or cut-points, hence the certainty of evidence was low. For the outcome of Down syndrome among high-risk pregnant women, the evidence was rated down once for differential verification and was assigned moderate certainty.

DISCUSSION

Our meta-analysis showed that NT measured through first-trimester ultrasound is specific for Down syndrome but has low sensitivity. It is a useful screening test for Down

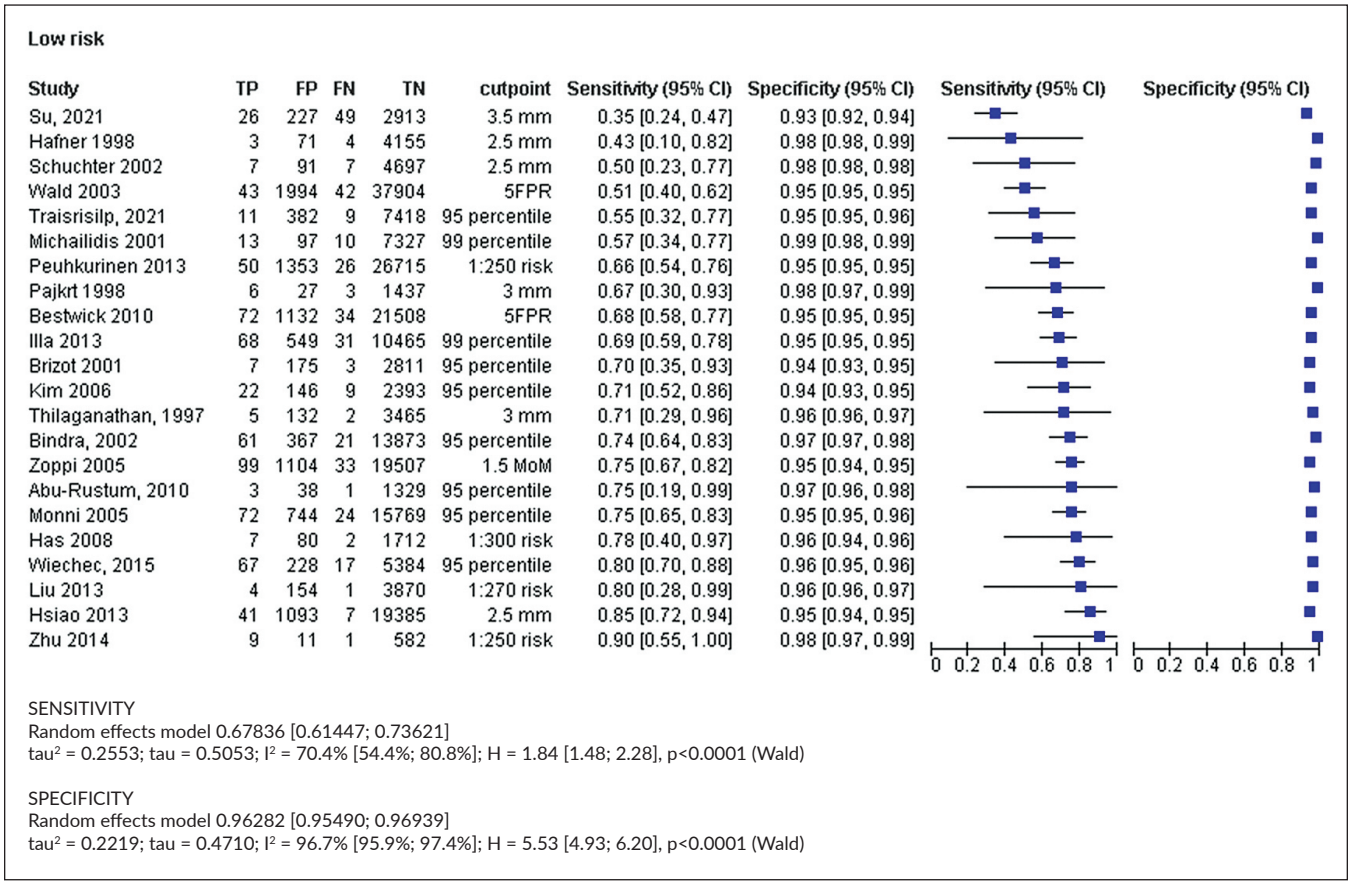

Figure 2. Forest plot of diagnostic characteristics for low-risk pregnant women for the outcome of Down syndrome.

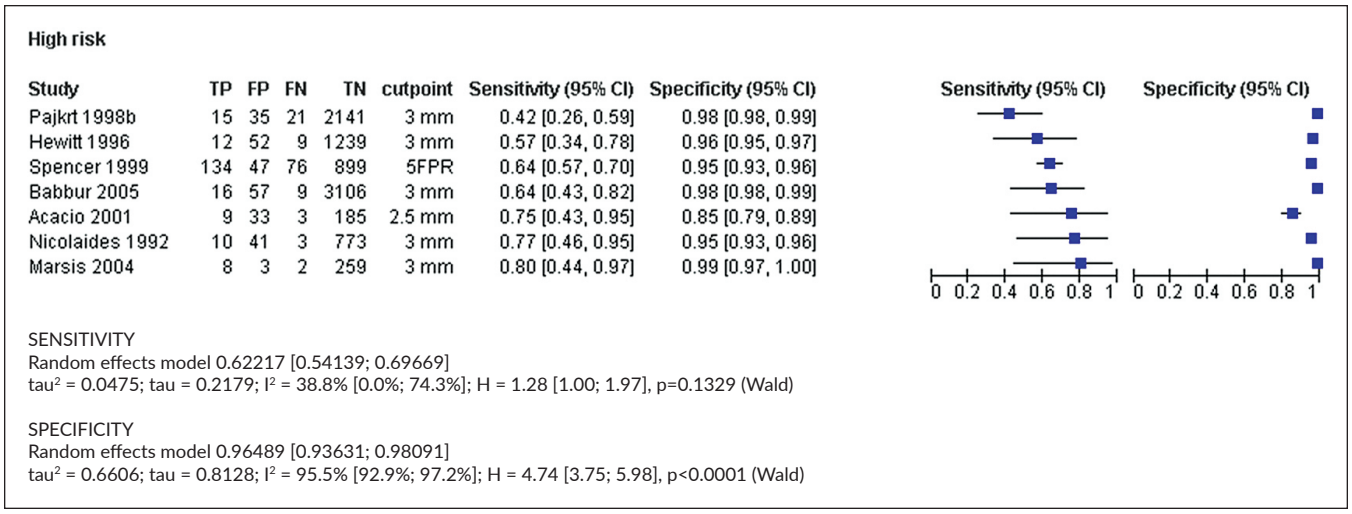

Figure 3. Forest plot of diagnostic characteristics for high-risk pregnant women for the outcome of Down syndrome.

syndrome in low-resource settings, however interpretation of results must take its limited sensitivity into consideration. Another key finding was the variation in certainty of the diagnostic performance of NT depending on the baseline risk of the pregnant population. For mothers at low-risk for fetal anomaly, there was low certainty that NT has high specificity and low sensitivity for Down syndrome. On the other hand, for high-risk women, the evidence for this level of diagnostic performance was of greater certainty (moderate).

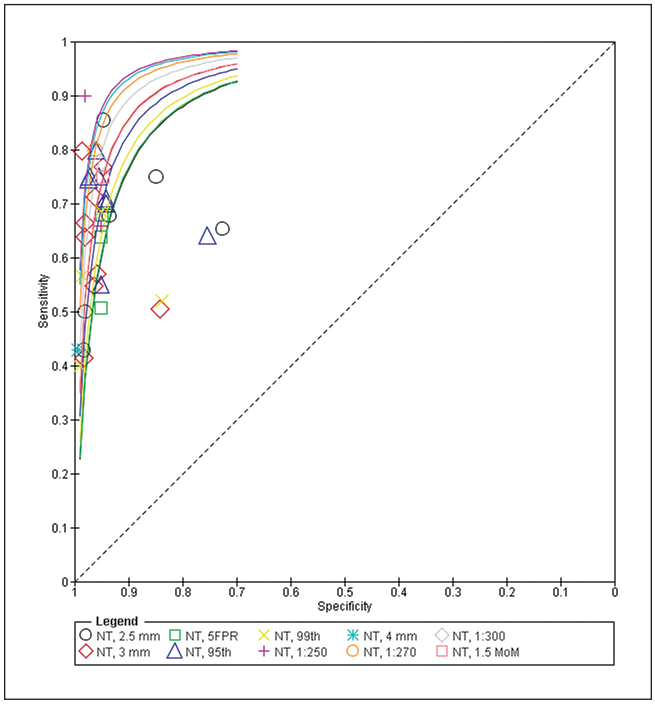

Figure 4. Summary receiver operating characteristics (ROC) curve for different NT cut-off values for Down syndrome.

Aside from NT screening alone, other strategies for Down syndrome screening are also available in urban areas of the Philippines and in high-income countries. In the second trimester, a method called quadruple screening can be performed, which involves blood tests measuring four maternal serum markers (alpha-fetoprotein, beta-human chorionic gonadotropin or beta-hCG, estriol, inhibin A). Patterns of increase and/or decrease among these markers are used to screen for fetal aneuploidies such as Down syndrome, and it has been shown to detect 81% of cases. Combinations of tests have also been formulated in the attempt to enhance the sensitivity of screening. First-trimester combined screening combines NT with blood tests for PAPP-A and beta-hCG, detecting 82-87% of cases. Going further, integrated screening involves doing both first-trimester combined screening and second-trimester quadruple screening, further enhancing the sensitivity to 96% at the expense of more tests with the corresponding financial cost and psychological burden. Another method of screening is stepwise sequential, wherein the women testing positive in the first-trimester combined testing are already offered confirmatory invasive testing, while those who tested negative will continue to receive the second-trimester quadruple testing. This stepwise method detects 95% of cases.[38,39]

Cell-free DNA (cfDNA) testing is a relatively new type of non-invasive screening performed through a maternal blood test drawn as early as 10 weeks age of gestation. CfDNA has been shown to be >99% sensitive and specific for Down syndrome. In high-income countries, it has largely supplanted traditional screening methods such as NT and quadruple serum markers and is recommended for use in all pregnant populations by the American College of Obstetricians and Gynecologists.[40,41]

In low-resource settings however, where the above screening methods may not be available, NT alone is a viable option for Down syndrome screening despite its

limited sensitivity. A positive NT screen will still be helpful for care providers to apprise families of the likelihood of Down syndrome and to provide additional investigation and counseling since pregnancy termination – one option in foreign countries, is illegal in the Philippines.

Another important aspect to consider when screening for Down syndrome is the acceptability and psychological impact of the test on the parents.

A pilot study by Drysdale et al. at St. Mary's Hospital in the UK in 2002 sought to evaluate the acceptability of routine early ultrasound at 12-14 weeks (including NT) among women presenting to a community midwife. In this study, 99% of women accepted the offer of ultrasound at 12-14 weeks age of gestation (AOG). Of these 984 women, 85% agreed to undergo Down syndrome screening by NT. Twenty-seven women were assessed to have high-risk based on NT (3.2%). Of these, 66.7% (n=18) opted to proceed with invasive confirmatory testing, which detected two cases of Down syndrome. Both cases then underwent elective termination, which is notably illegal in the Philippines. After the study, patients answered questionnaires regarding the acceptability of their experience. Majority of the women (83%) answered that they would accept a scan at their next pregnancy and found the scan to be a reassuring experience.[42] Women who have increased nuchal translucency reported significantly greater psychological distress after receiving the full screening report. However, their anxiety scores did not differ significantly from that of women with normal results at 22 weeks of gestation and after delivery. Furthermore, both groups remained supportive of the value of NT screening for their current and future pregnancies.[43] A study done in a district general hospital with multiethnic patients concluded that nuchal translucency screening can be effectively and equitably provided regardless of racial origin. Although African and Asian women consult significantly later during their pregnancy, they still presented early enough for nuchal translucency to be possible.[44] These studies show a predominantly acceptable perception of mothers to NT screening for Down syndrome. However, it must be acknowledged and kept in mind that these studies were performed predominantly in countries where termination is legal and there may be local variations on the acceptability and psychological impact among Filipinos. Individualization of the shared decision-making process is a must when offering any screening test.

In summary, our findings showed that NT as measured by first-trimester ultrasound has an acceptable diagnostic performance in the absence of other screening strategies. Furthermore, supporting evidence presented suggest that it is likely acceptable to mothers and has no evidence of sustained psychological impact. Therefore, NT remains a valuable screening test for the prenatal screening of Down syndrome among all pregnant women in settings where additional tests such as quadruple markers are not widely available.

CONCLUSION

NT as measured by ultrasound at 11 to 14 weeks AOG is specific for Down syndrome but has low sensitivity. It is a viable option for Filipinos in low-resource settings where serum screening tests are not readily available, however interpretation of results must take its limited sensitivity into consideration.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

This study was performed under the Institute of Clinical Epidemiology, National Institutes of Health of the University of the Philippines Manila with funding from the Department of Health, as part of the Philippine Periodic Health Examination (PHEX) guidelines for prenatal disorders.

REFERENCES

1. Akhtar F, Bokhari SRA. Down Syndrome [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan [updated 2023 Aug 8]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK526016/
2. United Nations. World Down Syndrome Day [Internet]. 2023 [cited 2023 Mar 20]. Available from: https://www.un.org/en/observances/down-syndrome-day
3. Makati Medical Center. 7 facts you still don't know about Down syndrome, probably [Internet]. 2020 [cited 2023 Mar 20]. Available from: https://www.makatimed.net.ph/blogs/7-facts-you-still-dont-know-about-down-syndrome-probably/
4. Padilla CD, Cutiongco-de la Paz EM. Genetics and genomic medicine in the Philippines. Mol Genet Genomic Med. 2016 Sep 15; 4(5):494-503. doi: 10.1002/mgg3.247. PMID: 27652277; PMCID: PMC5023934.
5. Centers for Disease Control and Prevention. Data and statistics on Down syndrome [Internet]. 2023 [cited 2023 Mar 20]. Available from: https://www.cdc.gov/ncbddd/birthdefects/downsyndrome/data.html.
6. MacLennan S. Down's syndrome. Innov Educ Inspir Gen Pract. 2020 Jan;13(1):47-52. doi: 10.1177/1755738019886612.
7. Renna MD, Pisani P, Conversano F, Perrone E, Casciaro E, Renzo GC, et al. Sonographic markers for early diagnosis of fetal malformations. World J Radiol. 2013 Oct 28;5(10):356-71. doi: 10.4329/wjr.v5.i10.356. PMID: 24179631; PMCID: PMC3812447.
8. de Castro-Hamoy LG, Tumulak MJR, Cagayan MSFS, Sy PA, Mira NRC, Laurino MY. Attitudes of Filipino parents of children with Down syndrome on noninvasive prenatal testing. J Community Genet. 2022 Aug;13(4):411-425. doi: 10.1007/s12687-022-00597-w. Epub 2022 Jun 2. PMID: 35653002; PMCID: PMC9161766.
9. Wise J. The end of Down's syndrome? BMJ. 2016 Oct 5;355:i5344. doi: 10.1136/bmj.i5344. PMID: 27707709.
10. Alldred SK, Takwoingi Y, Guo B, Pennant M, Deeks JJ, Neilson JP, et al. First trimester ultrasound tests alone or in combination with first trimester serum tests for Down's syndrome screening. Cochrane Database Syst Rev. 2017 Mar 15;3(3):CD012600. doi: 10.1002/14651858.CD012600. PMID: 28295158; PMCID: PMC6464518.
11. Liu Y, Ye X, Zhang N, Zhang B, Guo C, Huang W, et al. Diagnostic value of ultrasonographic combining biochemical markers for Down

syndrome screening in first trimester: a meta-analysis. Prenat Diagn. 2015 Sep;35(9):879-87. doi: 10.1002/pd.4626. Epub 2015 Jun 28. PMID: 26014106.
12. Traisrisilp K, Sirichotiyakul S, Tongprasert F, Srisupundit K, Luewan S, Jatavan P, et al. First trimester genetic sonogram for screening fetal Down syndrome: A population-based study. Taiwan J Obstet Gynecol. 2021 Jul;60(4):706-710. doi: 10.1016/j.tjog.2021.05.021. PMID: 34247811.
13. Wiechec M, Knafel A, Nocun A, Matyszkiewicz A, Juszczak M, Wiercinska E, et al. How Effective Is First-Trimester Screening for Trisomy 21 Based on Ultrasound Only? Fetal Diagn Ther. 2016 Mar;39(2):105-12. doi: 10.1159/000434632. Epub 2015 Aug 19. PMID: 26303666.
14. Thilaganathan B, Slack A, Wathen NC. Effect of first-trimester nuchal translucency on second-trimester maternal serum biochemical screening for Down's syndrome. Ultrasound Obstet Gynecol. 1997 Oct;10(4):261-4. doi: 10.1046/j.1469-0705.1997.10040261.x. PMID: 9383877.
15. Bindra R, Heath V, Nicolaides KH. Screening for chromosomal defects by fetal nuchal translucency at 11 to 14 weeks. Clin Obstet Gynecol. 2002 Sep;45(3):661-70; discussion 730-2. doi: 10.1097/00003081-200209000-00010. PMID: 12370606.
16. Abu-Rustum RS, Daou L, Abu-Rustum SE. Role of first-trimester sonography in the diagnosis of aneuploidy and structural fetal anomalies. J Ultrasound Med. 2010 Oct;29(10):1445-52. doi: 10.7863/jum.2010.29.10.1445. PMID: 20876898.
17. Su L, Wu X, Lin N, Xie X, Cai M, Wang M, et al. Different Cutoff Values for Increased Nuchal Translucency in First-Trimester Screening to Predict Fetal Chromosomal Abnormalities. Int J Gen Med. 2021 Nov 18;14:8437-8443. doi: 10.2147/IJGM.S330960. PMID: 34819751; PMCID: PMC8608408.
18. Hafner E, Schuchter K, Liebhart E, Philipp K. Results of routine fetal nuchal translucency measurement at weeks 10-13 in 4233 unselected pregnant women. Prenat Diagn. 1998 Jan;18(1):29-34. PMID: 9483637.
19. Schuchter K, Hafner E, Stangl G, Metzenbauer M, Höfinger D, Philipp K. The first trimester 'combined test' for the detection of Down syndrome pregnancies in 4939 unselected pregnancies. Prenat Diagn. 2002 Mar;22(3):211-5. doi: 10.1002/pd.288. PMID: 11920896.
20. Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM, et al. First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). Health Technol Assess. 2003 Apr;7(11):1-77. doi: 10.3310/hta7110. PMID: 12709291.
21. Michailidis GD, Spencer K, Economides DL. The use of nuchal translucency measurement and second trimester biochemical markers in screening for Down's syndrome. BJOG. 2001 Oct;108(10):1047-52. doi: 10.1111/j.1471-0528.2001.00241.x. PMID: 11702836.
22. Peuhkurinen S, Laitinen P, Honkasalo T, Ryynanen M, Marttala J. Comparison of combined, biochemical and nuchal translucency screening for Down syndrome in first trimester in Northern Finland. Acta Obstet Gynecol Scand. 2013 Jul;92(7):769-74. doi: 10.1111/aogs.12095. Epub 2013 Apr 9. PMID: 23369035.
23. Pajkrt E, van Lith JM, Mol BW, Bleker OP, Bilardo CM. Screening for Down's syndrome by fetal nuchal translucency measurement in a general obstetric population. Ultrasound Obstet Gynecol. 1998 Sep;12(3):163-9. doi: 10.1046/j.1469-0705.1998.12030163.x. PMID: 9793187.
24. Bestwick JP, Huttly WJ, Wald NJ. Distribution of nuchal translucency in antenatal screening for Down's syndrome. J Med Screen. 2010; 17(1):8-12. doi: 10.1258/jms.2010.009107. Erratum in: J Med Screen. 2010 Jun;17(2):106. PMID: 20356939; PMCID: PMC3104820.
25. Illa M, Mula R, Arigita M, Grande M, Gonce A, Borobio V, et al. Likelihood ratios to apply for nasal bone, ductus venosus and tricuspid flow at the 11-13 weeks' scan in down syndrome screening. Fetal Diagn Ther. 2013 Jun;34(2):116-20. doi: 10.1159/000351854. PMID: 23817147.
26. Brizot ML, Carvalho MH, Liao AW, Reis NS, Armbruster-Moraes E, Zugaib M. First-trimester screening for chromosomal abnormalities by fetal nuchal translucency in a Brazilian population. Ultrasound Obstet Gynecol. 2001 Dec;18(6):652-5. doi: 10.1046/j.0960-7692.2001.00588.x. PMID: 11844209.
27. Kim MH, Park SH, Cho HJ, Choi JS, Kim JO, Ahn HK, et al. Threshold of nuchal translucency for the detection of chromosomal aberration: comparison of different cut-offs. J Korean Med Sci. 2006 Feb;21(1):11-4. doi: 10.3346/jkms.2006.21.1.11. PMID: 16479057; PMCID: PMC2733957.
28. Zoppi MA, Ibba RM, Floris M, Manca F, Axiana C, Monni G. Nuchal translucency measurement at different crown-rump lengths along the 10- to 14-week period for Down syndrome screening. Prenat Diagn. 2005 May;25(5):411-6. doi: 10.1002/pd.1173. PMID: 15906410.
29. Monni G, Zoppi MA, Ibba RM, Floris M, Manca F, Axiana C. Nuchal translucency and nasal bone for trisomy 21 screening: single center experience. Croat Med J. 2005 Oct;46(5):786-91. PMID: 16158472.
30. Has R, Kalelioglu I, Yuksel A, Ibrahimoglu L, Ermis H, Yildirim A. Fetal nasal bone assessment in first trimester down syndrome screening. Fetal Diagn Ther. 2008 May;24(1):61-6. doi: 10.1159/000132409. Epub 2008 May 27. PMID: 18504384.
31. Liu LH, Lu XQ, Liu CH. Application of thickness of fetal nuchal translucency combined with serological test in screening of Downs syndrome. Matern Child Health Care China. 2013; 17:2715-2717.
32. Hsiao CH, Cheng PJ, Shaw SW, Hsu JJ, Chen RC, Tseng YJ, et al. Extended first-trimester screening using multiple sonographic markers and maternal serum biochemistry: a five-year prospective study. Fetal Diagn Ther. 2014 Feb;35(4):296-301. doi:10.1159/000357564. Epub 2014 Feb 6. PMID: 24503519.
33. Zhu YL, Tang L, Yang XH. Measurements of fetal nuchal translucency and nasal bone in first-trimester screening for Down syndrome. Chin J Fam Plann. 2014;22:98-101.
34. Pajkrt E, Mol BW, van Lith JM, Bleker OP, Bilardo CM. Screening for Down's syndrome by fetal nuchal translucency measurement in a high-risk population. Ultrasound Obstet Gynecol. 1998 Sep;12(3):156-62. doi: 10.1046/j.1469-0705.1998.12030156.x. PMID: 9793186.
35. Hewitt BG, de Crespigny L, Sampson AJ, Ngu AC, Shekleton P, Robinson HP. Correlation between nuchal thickness and abnormal karyotype in first trimester fetuses. Med J Aust. 1996 Oct 7;165(7):365-8. doi: 10.5694/j.1326-5377.1996.tb125018.x. PMID: 8890842.
36. Spencer K, Souter V, Tul N, Snijders R, Nicolaides KH. A screening program for trisomy 21 at 10-14 weeks using fetal nuchal translucency, maternal serum free beta-human chorionic gonadotropin and pregnancy-associated plasma protein-A. Ultrasound Obstet Gynecol. 1999 Apr;13(4):231-7. doi: 10.1046/j.1469-0705.1999.13040231.x. PMID: 10341399.
37. Babbur V, Lees CC, Goodburn SF, Morris N, Breeze AC, Hackett GA. Prospective audit of a one-centre combined nuchal translucency and triple test programme for the detection of trisomy 21. Prenat Diagn. 2005 Jun;25(6):465-9. doi: 10.1002/pd.1163. PMID: 15966036.
38. Acácio GL, Barini R, Pinto Júnior W, Ximenes RL, Pettersen H, Faria M. Nuchal translucency: an ultrasound marker for fetal chromosomal abnormalities. Sao Paulo Med J. 2001 Jan 4;119(1):19-23. doi: 10.1590/s1516-31802001000100006. PMID: 11175621.
39. 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 Apr 4;304(6831):867-9. doi: 10.1136/bmj.304.6831.867. PMID: 1392745; PMCID: PMC1882788.
40. Marsis IO. Screening for Down syndrome using nuchal translucency thickness and nasal bone examination at advanced maternal age in Jakarta: a preliminary report. J Med Ultrasound. 2004;12(1):1-6. doi:10.1016/S0929-6441(09)60054-0.
41. Caughey AB, Kuppermann M, Norton ME, Washington AE. Nuchal translucency and first trimester biochemical markers for down syndrome screening: a cost-effectiveness analysis. Am J Obstet Gynecol. 2002 Nov;187(5):1239-45. doi: 10.1067/mob.2002.127144. PMID: 12439512.

42. Odibo AO, Stamilio DM, Nelson DB, Sehdev HM, Macones GA. A cost-effectiveness analysis of prenatal screening strategies for Down syndrome. Obstet Gynecol. 2005 Sep;106(3):562-8. doi: 10.1097/01.AOG.0000174581.24338.6f. Erratum in: Obstet Gynecol. 2006 Jan;107(1):209. PMID: 16135588.
43. Gilbert RE, Augood C, Gupta R, Ades AE, Logan S, Sculpher M, et al. Screening for Down's syndrome: effects, safety, and cost effectiveness of first and second trimester strategies. BMJ. 2001 Aug 25;323(7310):423-5. doi: 10.1136/bmj.323.7310.423. PMID: 11520837; PMCID: PMC37550.
44. Torfs CP, Christianson RE. Socioeconomic effects on the risk of having a recognized pregnancy with Down syndrome. Birth Defects Res A Clin Mol Teratol. 2003 Jul;67(7):522-8. doi: 10.1002/bdra.10071. Erratum in: Birth Defects Res Part A Clin Mol Teratol. 2003 Oct;67(10):893. PMID: 14565624.
45. Hunter JE, Allen EG, Shin M, Bean LJ, Correa A, Druschel C, et al. The association of low socioeconomic status and the risk of having a child with Down syndrome: a report from the National Down Syndrome Project. Genet Med. 2013 Sep;15(9):698-705. doi: 10.1038/gim.2013.34. Epub 2013 Apr 4. PMID: 23558253; PMCID: PMC4122862.
46. Chen Y, Qian X, Zhang J, Li J, Chu A, Schweitzer SO. Preliminary study into the economic burden of Down syndrome in China. Birth Defects Res A Clin Mol Teratol. 2008 Jan;82(1):25-33. doi: 10.1002/bdra.20409. PMID: 18022893.
47. Kageleiry A, Samuelson D, Duh MS, Lefebvre P, Campbell J, Skotko BG. Out-of-pocket medical costs and third-party healthcare costs for children with Down syndrome. Am J Med Genet A. 2017 Mar;173(3):627-637. doi: 10.1002/ajmg.a.38050. Epub 2016 Dec 14. PMID: 27966292.
48. Geelhoed EA, Bebbington A, Bower C, Deshpande A, Leonard H. Direct health care costs of children and adolescents with Down syndrome. J Pediatr. 2011 Oct;159(4):541-5. doi: 10.1016/j.jpeds.2011.06.007. Epub 2011 Jul 23. PMID: 21784457; PMCID: PMC3858577.
49. Drysdale K, Ridley D, Walker K, Higgins B, Dean T. First-trimester pregnancy scanning as a screening tool for high-risk and abnormal pregnancies in a district general hospital setting. J Obstet Gynaecol. 2002 Mar;22(2):159-65. doi: 10.1080/01443610120113300. PMID: 12521697.
50. Chueh HY, Cheng PJ, Shaw SW, Lin CT, Hsu JJ, Hsieh TT. Maternal anxiety about first trimester nuchal translucency screening and impact of positive screening results. Acta Obstet Gynecol Scand. 2007 Sep;86(12):1437-41. doi: 10.1080/00016340701622724. Epub 2007 Sep 4. PMID: 17851819.
51. Thilaganathan B, Khare M, Williams B, Wathen NC. Influence of ethnic origin on nuchal translucency screening for Down's syndrome. Ultrasound Obstet Gynecol. 1998 Aug;12(2):112-4. doi: 10.1046/j.1469-0705.1998.12020112.x. PMID: 9744055.

APPENDICES

Appendix A. Search Strategy

We searched PubMed, Proquest, and CENTRAL using a high-sensitivity method filter for systematic reviews and/or meta-analyses using the following terms:

((nuchal translucency) OR (nuchal translucency measurement[MeSH Terms]) OR (nuchal translucency screening[MeSH Terms])) AND ("MEDLINE"[Text Word] OR "systematic review"[Text Word] OR "meta-analysis"[Publication Type] OR "intervention*"[Title])

Two independent reviewers assessed the retrieved systematic reviews using the AMSTAR-2 checklist. Upon selecting a high-quality 2017 Cochrane review on Down syndrome by Alldred for adaptation, we performed a search update in the same databases to include any new studies from August 25, 2011 onwards using the following terms:

((((("prenatal diagnosis") OR (Antenatal OR prenatal OR trimester OR pregnan OR fetus OR fetal)) AND ("nuchal translucency" AND "ultrasound")) AND (Screen* OR Detect* OR Accura* OR Predict* OR ROC OR "ROC curve" OR AUC OR "False positive" OR "false negative" OR "Likelihood ratio" OR Sensitiv* OR Specific* OR Diagnos*)) AND ("Down syndrome" OR "Trisomy" OR "Aneuploidy" OR "Mosaicism")
Filter: Humans, from 2011-2024

Appendix B. Included Studies

Table 1. Characteristics of Included Studies

Author, year	Location	N	Population	Study Design	Outcome	Index Test	Reference standard	Follow-up	Withdrawal explanation
Abu-Rustum, 2010	Lebanon	1370	Routine screening Pregnant women 11-13 6/7 weeks AOG	Retrospective cohort	Down syndrome 4 cases	NT >95th percentile	Karyotyping, postnatal exam	No details on follow-up methods.	3% of patients were lost to follow-up or had spontaneous losses
Acacio, 2001	Brazil private centers	230	High-risk Pregnant women 21-45 years Singleton 10-14 weeks AOG	Retrospective cohort	Down syndrome 12 cases	NT >2.5 mm	Chorionic villus biopsy, amniocentesis or blood or placenta used for fetal karyotyping	100% karyotyping	No details given
Babbur, 2005	Cambridge Maternity Hospital	3188	Self-request Women with history of fetal aneuploidy Pregnant women 19-46 years Singleton 11-14 weeks AOG	Prospective cohort	Down syndrome 25 cases	NT >3 mm	Invasive testing offered to women with NT >3 mm	Details of follow-up to birth not given	463 patients having wide NT did not go on to have serum testing and were excluded
Bestwick, 2010	London antenatal clinics	22746	Routine screening Pregnant women 11-13 weeks AOG	Retrospective cohort	Down syndrome 106 cases	First trimester NT, PAPP-A and free ßhCG (details not reported) NT at 5% false positive rate	Karyotyping or follow-up to birth	Data obtained from the hospitals, the regional cytogenetic unit and the National Down Syndrome Cytogenetic Register.	No details of withdrawals given
Bindra, 2002	UK hospital	15030	Routine screening Pregnant women 15-49 years 11-14 weeks AOG	Prospective cohort	Down syndrome 82 cases	NT >95th percentile	Invasive testing offered to high-risk patients with risk at least 1:300	Data on pregnancy outcome were obtained from the cytogenetics laboratory, the patients themselves, their general practitioners or the maternity units in which they delivered. Cases were also linked to those recorded in the National Down Syndrome Register.	Excluded from further analysis were 647 cases, because the fetal karyotype was not known and they resulted in spontaneous fetal loss (n = 41), termination of pregnancy (n = 32), or were lost to follow-up (n = 574)
Brizot, 2001	Brazil University Hospital	2996	Routine screening Pregnant women 13-46 years Singleton 10-14 weeks AOG	Prospective cohort	Down syndrome 10 cases	NT at 95th percentile, 99th percentile	Antenatal karyotyping (5.9% of pregnancies: 62% of high-risk, 29% of medium-risk and 3% of the low-risk women) or follow-up to birth (85.3% of women)	85.3% of women were followed up to birth. Of these, 65 were spontaneous miscarriages or intrauterine death with no karyotyping.	No details of withdrawals given
Hafner, 1998	Austria hospital	4233	Routine screening Pregnant women 15-49 years 10-13 weeks AOG	Prospective cohort	Down syndrome 7 cases	NT >2.5 mm	Amniocentesis or CVS in patients with previous Down's pregnancy, >35 years or with a positive biochemical test result. Other women underwent scan at 22 weeks and, if NT >2.5 mm, special examination directed to examination of fetal heart. Follow-up to birth	No details given on methods of follow-up. 138 women lost to follow-up.	No details of withdrawals given
Has, 2008	Turkey	1807	Routine screening Pregnant women 17-45 years Singleton 11-14 weeks AOG	Cohort study	Down syndrome 9 cases	NT cutoff 1:300 risk	Karyotyping or follow-up to birth	Findings recorded in a computer database. Karyotype results obtained directly from the genetics department. Pregnancy outcomes obtained from hospital records or from parents via telephone interview. 110 women (5%) with terminations, miscarriages or malformations and unknown outcome were excluded from the study.	No details of withdrawals given

Table 1. Characteristics of Included Studies *(continued)*

Author, year	Location	N	Population	Study Design	Outcome	Index Test	Reference standard	Follow-up	Withdrawal explanation
Hewitt, 1996	Australia, 2 hospitals and 2 private practices	1317	High-risk referral for invasive testing Pregnant women 21-48 years 10-14 weeks AOG	Prospective cohort	Down syndrome 21 cases	NT >3 mm	Chorionic villus sampling	100% karyotyping	No details of withdrawals given
Hsiao, 2013	Taiwan	20526	Routine screening 11-13 6/7 weeks AOG	Prospective cohort	Down syndrome 48 cases	NT >2.5 mm	Risk of 1/300 or greater was regarded as screen-positive and an invasive diagnostic procedure such as chorionic villus sampling or amniocentesis were offered for chromosomal analysis	Karyotype results, details of the pregnancy outcomes, as well as complications were added into the database as soon as they became available.	Data derived from the 256 (1.2%) pregnancies with twins, as well as 272 (1.3%) with incomplete data due to miscarriage, lost to follow-up, missing data, and incorrect dates for screening were excluded from this study
Illa, 2013	Spain	11261	Routine screening 11-13 weeks AOG	Prospective cohort	Down syndrome 101 cases	NT >99th percentile	CVS offered for high-risk women, postnatal outcome for those without karyotyping	Data on perinatal outcome was sought for the non-karyotyped pregnancies. If the pregnancy was not delivered in our center, details of the pregnancy outcome were obtained by telephone either from the mother or the attending obstetrician.	Pregnancies with: (a) neither perinatal outcome nor karyotype; (b) no combined test results; (c) other chromosomal anomalies than T21; and (d) multiple pregnancies were excluded from the study.
Kim, 2006	Korea, hospitals and women's healthcare center	2570	Routine screening Pregnant women 26-33 years Singleton 10-14 weeks AOG	Retrospective cohort	Down syndrome 31 cases	NT >2.5 mm, 3.0 mm or 95th percentile	Amniocentesis or CVS in 419 patients considered high-risk (NT >2.5, aged >35 years, positive biochemical test result, history of chromosomal abnormality, fetal structural abnormality at ultrasound or other reason). Follow-up to birth	Pregnancy outcomes ascertained from obstetric and neonatal medical records of live or stillborn babies. Only patients with known pregnancy outcome included in the study. 8 patients who terminated their pregnancies because of structural abnormalities on ultrasound with no karyotyping results were excluded. Karyotyping was performed in intrauterine fetal death (n = 4) cases.	No details of withdrawals given
Liu, 2013 (article in Chinese)	China	4029	Routine screening 11-14 weeks AOG	Cohort	Down syndrome 5 cases	NT at 1:270 risk	CVS or postnatal exam	Could not access details	Could not access details
Marsis, 2004	Indonesia, 4 hospitals	262	Screening for patients 35 years and above 35-43 years Singleton 11-13 weeks AOG	Prospective cohort	Down syndrome 8 cases	NT >3 mm	Amniocentesis (unclear in which patients this was conducted) or follow-up to birth	Follow-up to birth in patients with no nasal bone and NT >3 mm. Unclear if screen-negative patients had follow-up to birth.	No details of withdrawals given
Michailidis, 2001	UK, hospital maternity unit	7447	Routine screening Pregnant women 13-50 years 10-14 weeks AOG	Prospective cohort	Down syndrome 23 cases	NT>99th percentile	Karyotyping in women considered at risk due to index test results, age, or family history or those with considerable anxiety (632 women, 8.5%). Follow-up to birth	Outcome at birth assess from hospital database, labour ward records or directly from patients. Follow-up data in 7447 patients (87% of initial patient cohort). Patients without follow-up excluded.	No details of withdrawals given
Monni, 2005	Italy, single center	16654	Routine screening Pregnant women 14-49 years Singleton 10-14 weeks AOG	Prospective cohort	Down syndrome 96 cases	NT >95th percentile	Karyotyping or follow-up to birth	Outcome at birth as recorded in hospital database (provided by outcome sheets or telephone interviews). Of 32,000 cases in the database, 16,654 (52%) patients had NT, nasal bone assessment and follow-up data available. Patients without follow-up data were excluded from the study.	No details of withdrawals given

Table 1. Characteristics of Included Studies *(continued)*

Author, year	Location	N	Population	Study Design	Outcome	Index Test	Reference standard	Follow-up	Withdrawal explanation
Nicolaides, 1992	UK, research center for fetal medicine	827	High-risk referral for invasive testing Pregnant women 22-47 years 10-14 weeks AOG	Prospective cohort	Down syndrome 13 cases	NT >3 mm	Fetal karyotyping by amniocentesis (52%) or CVS (48%)	100% karyotyping	No details of withdrawals given
Pajkrt, 1998a	Netherlands tertiary maternity unit	1473	Routine screening Pregnant women 26-36 years Singleton 10-14 weeks AOG	Prospective cohort	Down syndrome 9 cases	NT >3 mm	Prenatal karyotyping offered to patients considered high-risk or maternal anxiety (conducted in 24%) or follow-up to birth	Follow-up to outcome assessment in the delivery room. 68 women (4.4%) were excluded from the study due to loss to follow-up	No details of withdrawals given
Pajkrt, 1998b	Netherlands, prenatal diagnostic centre	2247	High-risk referral for invasive testing Pregnant women 22-46 years Singleton 10-14 weeks AOG	Consecutive cohort	Down syndrome 36 cases	NT >3 mm	Prenatal karyotyping	100% karyotyping	Patients excluded due to sonographically detected fetal abnormalities at NT measurement, no karyotyping or miscarriages
Peuhkurinen, 2013	Finland	26715	Routine screening 9-13 6/7 weeks AOG	Prospective cohort	Down syndrome 76 cases	NT at risk >1:250	Karyotyping	Data obtained from: 1) Genetics Laboratory of the Department of Clinical Genetics at Oulu University Hospital, which is responsible for chromosomal diagnostics in the Oulu area 2) Finnish National Register of Congenital Malformations, which receives information about all Down's syndrome cases diagnosed in Finland 3) National Research and Development Centre for Welfare and Health, which records the birth of all live and stillborn infants	No details on withdrawals
Schuchter, 2002	Austria, single institution	4802	Routine screening Pregnant women>35 years Singleton 10-12 weeks AOG	Prospective cohort	Down syndrome 14 cases	NT >2.5 mm	CVS and amniocentesis (offered to patients with increased risk (>1:400) at first trimester screening. CVS recommended when NT >3.5 or when women did not want to wait until the 15th week for amniocentesis), or follow-up to birth	Patients without follow-up information (n = 92, 2%) were excluded from the study. 27 women with spontaneous abortions were also excluded from the study.	Women not attending visits were excluded from the study
Spencer, 1999	UK - fetal medicine research centre	1156	Women referred for invasive testing or self-referred for screening Pregnant women 19-46 years 10-14 weeks AOG	Case control	Down syndrome 210 cases	NT at 5% false positive rate	Invasive testing (high-risk women) or follow-up to birth	Stated that pregnancy outcome was ascertained in all women.	No details of withdrawals given
Su, 2021	China	3215	Routine screening Pregnant women Singleton 11-14 weeks AOG	Retrospective cohort	Down syndrome 75 cases	NT at >2.5 mm, >3 mm, >95th percentile, >99th percentile	Karyotyping for all	No follow-up details.	8 participants without increased NT and with failed cell cultures were excluded.
Thilaganathan, 1997	UK	2920	Routine screening 10-14 weeks AOG	Prospective cohort	Down syndrome 7 cases	NT >3 mm, >4 mm	CVS or amniocentesis offered to high-risk women	No follow-up details.	In 452 cases (12.5%), NT was not measured because it was declined by the women (3.1%) or it was not technically possible in the 5 min allocated for the scan (9.4%).

Table 1. Characteristics of Included Studies *(continued)*

Author, year	Location	N	Population	Study Design	Outcome	Index Test	Reference standard	Follow-up	Withdrawal explanation
Traisrisilp, 2021	Thailand hospital	7820	Routine screening 11-13 6/7 weeks AOG	Prospective cohort	Down syndrome 20 cases	NT >95th percentile	Karyotypes derived from chorionic villus sampling, amniocentesis, cordocentesis or neonatal work-up	All women were followed up for pregnancy outcomes, which were assessed by the obstetricians, while neonatal outcomes were assessed by the pediatricians of the research team.	Exclusion criteria were cases with fetal structural anomalies other than trisomy 21 and unavailability of final outcomes.
Wald, 2003	UK and Austria - multicentre trial	39983	Routine screening Pregnant women 9-13 weeks AOG, then 14-20 weeks	Case control	Down syndrome 85 cases	NT at 5% false positive rate	Invasive testing (following second trimester screening) or follow-up to birth	Follow-up by: 1) staff at local hospitals completed a study outcome form at, or just after delivery, 2) study records of CVS, amniocentesis or karyotype at birth linked to information from cytogenic laboratories, 3) study records linked to records of cases of Down's syndrome from the National Down's Syndrome Cytogenetic Register, 4) information obtained from local obstetrical outcome records, 5) forms sent to all women with a request to return details of the outcome of their pregnancy, 6) individual searches in respect of women whose outcomes of pregnancy had not been obtained by any of the previous methods. 4% of total patient cohort did not have a documented outcome of pregnancy. Unclear if any of these were included in the nested case-control study.	No details of withdrawals given
Wiechec, 2015	Poland tertiary care center	6265	Routine screening Singleton 11-13 6/7 weeks AOG	Prospective cohort	Down syndrome 84 cases	NT >95th percentile	Karyotyping, postnatal exam	Follow-up details not mentioned	569 (9.08%) cases were excluded: (a) in 416 (6.6%) patients, it was impossible to establish a fetal karyotype, since they were lost to follow-up, (b) 58 (0.93%) cases had miscarriages unrelated to invasive testing, (c) 28 (0.45%) patients had intrauterine fetal demise without subsequent karyotyping, and (d) in 67 (1%) cases, a chromosomal abnormality other than T21 was found.
Zhu, 2014 (article in Chinese)	China	603	Routine screening	Cohort	Down syndrome 10 cases	NT risk >1:250	CVS or postnatal exam	Could not access details	Could not access details
Zoppi, 2005	Italy	20743	Self-referral or referred by other physicians Pregnant women 10-14 weeks AOG	Prospective cohort	Down syndrome 132 cases	NT >1.5 MoM	Invasive prenatal diagnosis if high-risk	Prenatal karyotype and pregnancy and neonatal outcomes (neonatal assessment or necropsy), as provided by outcome sheets or telephone interviews.	No details on withdrawals

Figure 1. Risk of Bias for Studies on Down syndrome.

Appendix C. Subgroup analysis based on NT cut-off values

***NT cut-off at 2.5 mm* (Figure 2)**

Six cohort studies (n=35,584; 3 prospective, 3 retrospective) investigated the test characteristics of nuchal translucency (NT) measurement by ultrasound at a cut-off of 2.5 millimeters for the diagnosis of Down syndrome. One study by Acacio in 2001 screened pregnant women at high-risk for fetal aneuploidy, while the other five studies were performed among the general population. For all studies, nuchal translucency was measured by ultrasonographers with Fetal Medicine Foundation (FMF) certification for NT measurement. Two of these studies (Acacio et al., 2001; Su et al., 2021) used karyotyping as a uniform reference standard, while four studies used differential verification, with karyotyping done only for those with a positive index test, while those with a negative index test received postnatal examination as reference standard. Among these studies, the mean sensitivity was 64.4% (range: 42.9%-85.4%) with moderate heterogeneity (I^2=52.4%, p=0.0621). The mean specificity was 90.4% (range: 72.6%-98.3%) with substantial heterogeneity (I^2=99.7%, p=0). To explore sources of heterogeneity, subgroup analysis was performed according to population risk; however, heterogeneity within the low-risk population studies (excluding Acacio et al., 2001) remained substantial for sensitivity (I^2=60.9%, p=0.0367) and specificity (I^2=99.8%, p=0). Subgroup analysis by mode of verification and sensitivity analysis by excluding the outlier in the forest plot (Su et al.) were attempted, but substantial heterogeneity remained.

***NT cut-off at 3 mm* (Figure 3)**

Nine cohort studies (n=18,681; 7 prospective, 2 retrospective) tested the diagnostic performance of a cut-off threshold of 3 millimeters for Down syndrome screening. Three of these studies (Hewitt et al., 1996; Nicolaides et al., 1992; Pajkrt et al., 1998b) involved women at high-risk for fetal aneuploidy, while six studies were among pregnant women at low risk. Three studies (Su et al., 2021; Hewitt et al., 1996; Pajkrt et al., 1998b) used invasive testing to uniformly confirm the diagnosis, while six others had differential verification. The pooled sensitivity of the nine studies was 56.8% (95% CI: 47.3%-65.8%, I^2=17.7%, p=0.2855). The mean specificity value was 95.8% (range: 84.4%-98.9%, I^2=98.9%, p<0.0001). This considerable numerical heterogeneity remained even when studies were divided by population risk and mode of verification. Despite this unexplained heterogeneity in specificity, the high magnitude and narrow confidence interval suggest that this inconsistency is not likely to affect clinical decisions.

***NT cut-off at 4 mm* (Figure 4)**

A single prospective cohort study by Thilaganathan et al. in 1997 tested the diagnostic accuracy of NT at 4 millimeters for Down syndrome screening among 3,604 pregnant women. Differential verification was used wherein only patients with a positive index test were offered invasive testing. Sensitivity was pegged at 43.0% (95% CI: 10.0%-82.0%), while specificity was 99.0% (95% CI: 99.0%-100%).

***NT cut-off at 95th percentile* (Figure 5)**

The diagnostic accuracy of NT at 95th percentile for Down syndrome screening was evaluated by FMF-accredited sonographers in eight cohort studies (5 prospective, 3 retrospective) among 54,607 fetuses with mothers having a baseline low risk of aneuploidy. Except for three studies (Bindra et al., 2002; Abu-Rustum et al., 2010; Su et al., 2021), five studies made use of differential verification with invasive confirmatory testing for those with a positive screen. The overall pooled sensitivity for this cutoff was 72.2% (95% CI: 66.9%-76.9%, I^2=15.4%, p=0.3086). Specificity values were substantially heterogeneous (I^2=99.6%, p=0) with a mean of 93.1% (range: 75.4%-97.4%). On visual inspection of the forest plot, the specificity value of the Su et al. study had an outlier value for specificity, hence an analysis was performed excluding this study. Even with this exclusion, heterogeneity remained considerable at I^2=95.9%, p<0.0001. Despite this, the specificity in this analysis remained high at 95.8% (95%

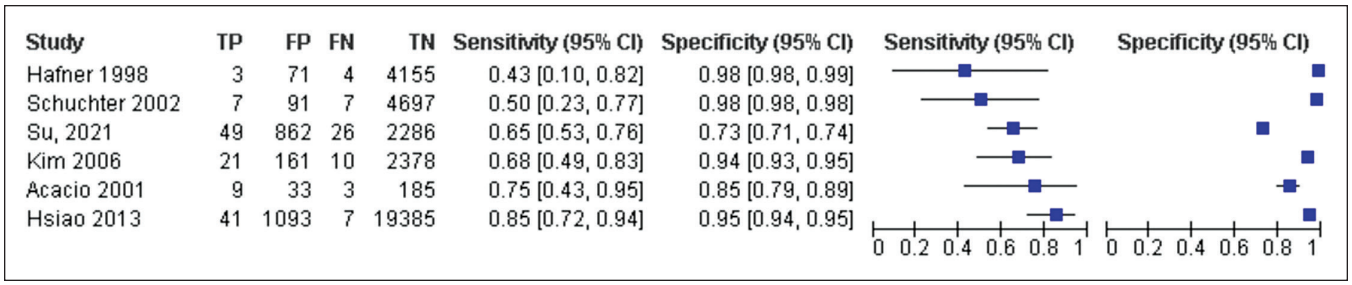

Figure 2. Forest plot of studies with NT cut-off of 2.5 mm.

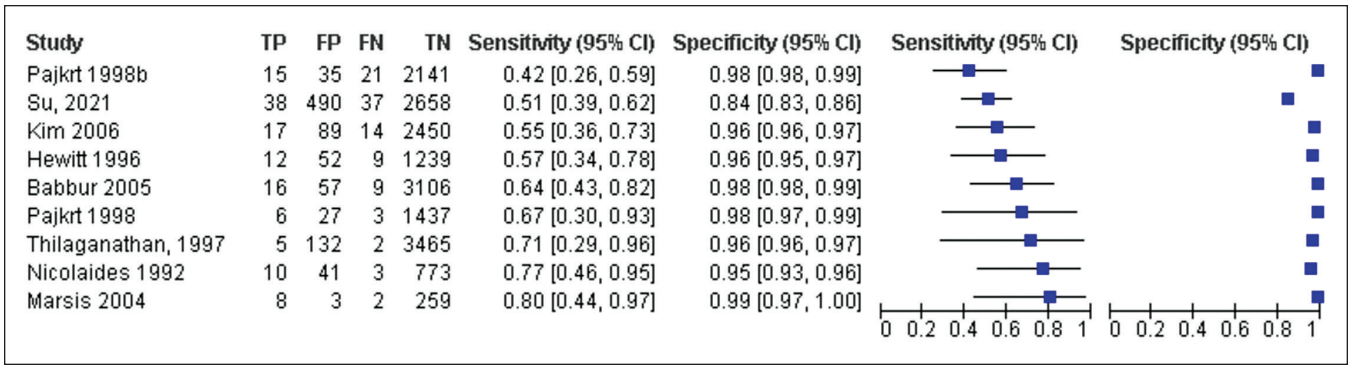

Figure 3. Forest plot of studies with NT cut-off of 3 mm.

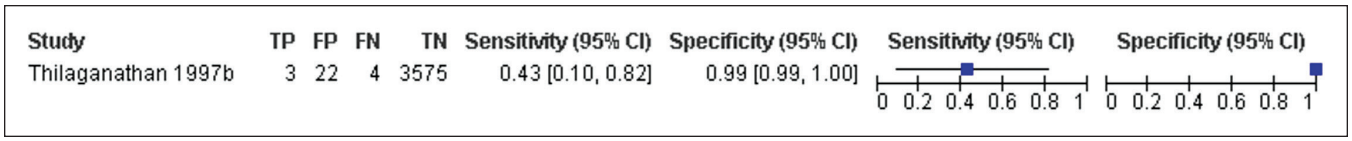

Figure 4. Forest plot of studies with NT cut-off of 4 mm.

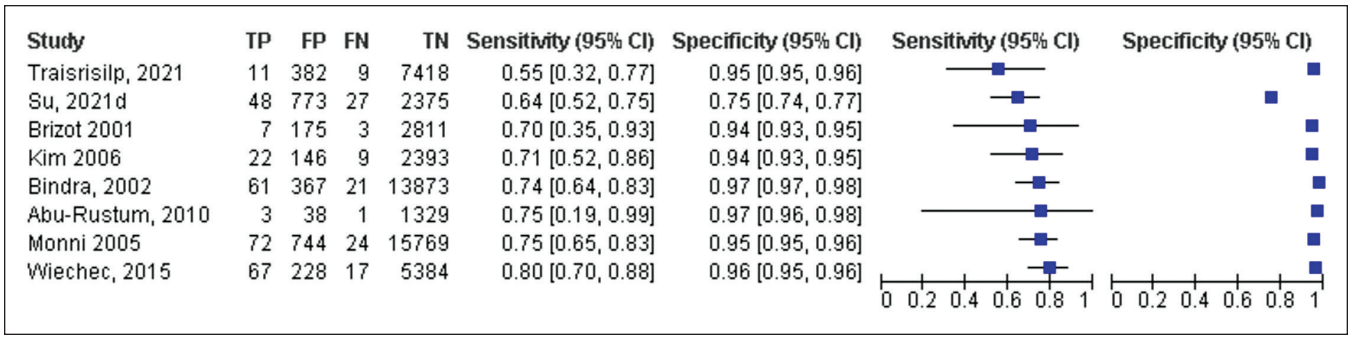

Figure 5. Forest plot of studies with NT cut-off of 95th percentile.

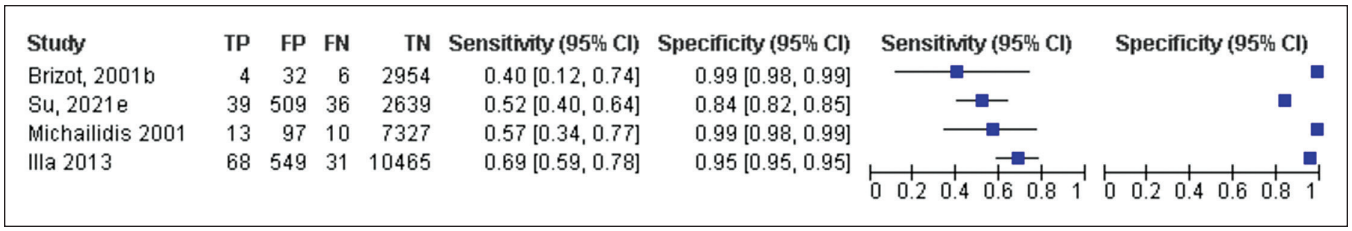

Figure 6. Forest plot of studies with NT cut-off of 99th percentile.

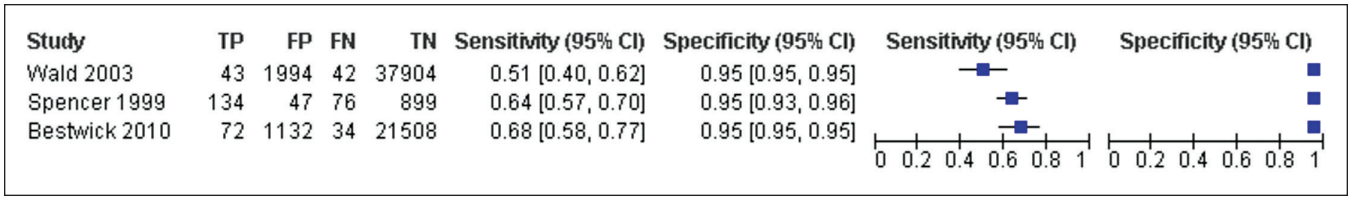

Figure 7. Forest plot of studies with NT cut-off of 5% FPR.

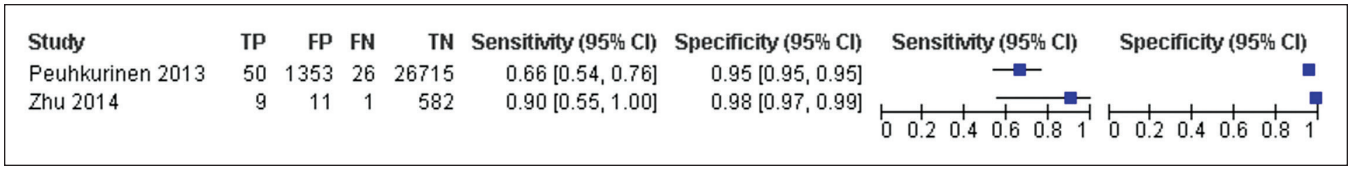

Figure 8. Forest plot of studies with NT cut-off of 1:250 risk.

CI: 94.8%-96.6%), suggesting that the variability may not be clinically important.

***NT cut-off at 99th percentile* (Figure 6)**

Four cohort studies (3 prospective, 1 retrospective) examined the diagnostic accuracy of NT at a cutoff of 99th percentile or greater in screening for Down syndrome among a total of 24,779 fetuses with mothers at low risk of fetal aneuploidy. Three studies used postnatal examination to confirm the diagnosis of Down syndrome among those with a negative index test, while invasive testing was used among those with a positive index test. In contrast, one study (Su et al., 2021) used uniform testing with karyotyping as reference standard. Mean sensitivity among the studies was 54.3% with moderate heterogeneity (range: 40.0%-68.7%, I^2=55.5%, p=0.0807). Mean specificity was 94.1% (range: 83.8%-98.9%, I^2=99.6%, p<0.0001). Visual inspection of the Forest plot showed that sensitivity values generally overlapped, while for specificity values, the study by Su et al. was an outlier. Excluding this study, heterogeneity declined for both parameters but was still substantial for specificity. Pooled sensitivity was 64.4% (95% CI: 55.9%-72.1%, I^2=47.4%, p=0.1497), while pooled specificity was 98.1% (95% CI: 95.8%-99.1%, I^2=99.0%, p<0.0001). The high magnitude of specificity coupled with a narrow CI denote that this residual heterogeneity may not impact clinical decisions.

***NT cut-off at 5% false positive rate* (Figure 7)**

There were three retrospective cohort studies examining the diagnostic characteristics of NT at a cutoff producing a 5% false positive rate depending on the prevalence of Down syndrome in the population. The studies tested 63,885 fetuses with mothers at low-risk for aneuploidy and used differential verification with those with positive index test being offered invasive confirmatory testing. The mean sensitivity was 60.7% (range: 50.6%-67.9%, I^2=69.1%, p=0.0392). Specificity was less variable, with pooled value at 95.0% (95% CI: 94.8%-95.2%, I^2=0).

***NT cut-off at 1:250 risk* (Figure 8)**

Two cohort studies (n=28,747; 1 prospective, 1 retrospective) investigated the diagnostic accuracy of NT cut-off at risk of 1:250 as calculated by FMF method for Down syndrome screening. Both studies tested low-risk populations with differential reference standards, favoring mothers with a positive index test for invasive confirmation. The pooled sensitivity was 68.6% (95% CI: 58.1%-77.5%, I^2=50.9%, p=0.1536) with moderate heterogeneity. Specificity values averaged at 96.7% (range: 95.1%-98.1%), but with considerable heterogeneity (I^2=90.4%, p=0.0013). The specificity value was high and the CI was narrow, hence this variability will likely not affect clinical decisions.

***NT cut-off at 1:270 risk* (Figure 9)**

A single retrospective cohort by Liu et al., 2013 evaluated the diagnostic performance of NT cut-off at 1:270 risk calculated on FMF software. The study included 4,029 fetuses with mothers at low-risk for aneuploidy and the diagnosis was confirmed with invasive testing among mothers with a positive index test. The sensitivity was 80.0% (95% CI: 28.0%-99.0%), and specificity was 96.0% (95% CI: 96.0%-97.0%).

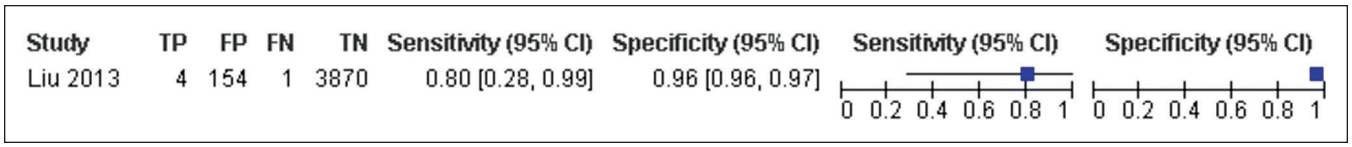

Figure 9. Forest plot of studies with NT cut-off of 1:270 risk.

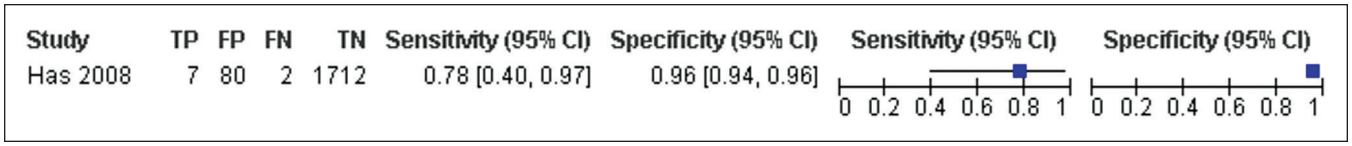

Figure 10. Forest plot of studies with NT cut-off of 1:300 risk.

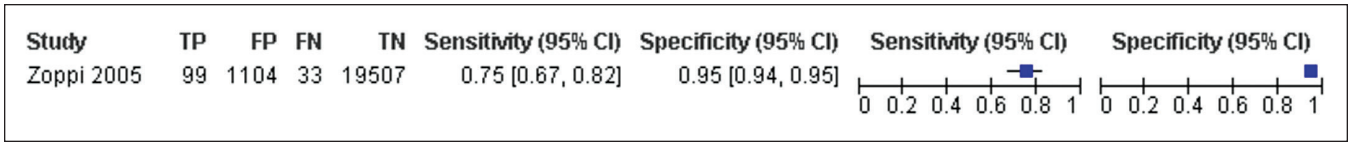

Figure 11. Forest plot of studies with NT cut-off of 1.5 MoM.

***NT cut-off at 1:300 risk* (Figure 10)**
A retrospective cohort study by Has et al. in 2008 investigated the test characteristics of NT cutoff at 1:300 risk among women with low-risk of fetal aneuploidy. This study involved 1,801 fetuses and used differential verification of outcome depending on index test result. The sensitivity of NT in this study was 78.0% (95% CI: 40.0%-97.0%). Specificity was 96.0% (95% CI: 94.0%-96.0%).

***NT cut-off at 1.5 MoM* (Figure 11)**
In 2005, Zoppi et al. performed a diagnostic test accuracy study (prospective cohort) examining the performance of NT at a cutoff of 1.5 MoM (multiples of median) calculated for the sample population. Among 20,743 mothers at low-risk for fetal aneuploidy, NT was measured by FMF-certified sonographers and verification depended on their index test result. Sensitivity was calculated at 75.0% (95% CI: 67.0%-82.0%) and specificity was at 95.0% (95% CI: 94.0%-95.0%).

Overall Certainty of Evidence by Cut-off
Majority of studies (22 of 29 studies) used variable reference standards to confirm the diagnosis of Down syndrome, failed to clarify if all patients received a reference standard, or did not explicitly explain the cause of exclusions, hence evidence for all cut-offs were rated down once for risk of bias. For two of the explored cut-offs (2.5 mm, 5% FPR), there was substantial heterogeneity that could not be explained by subgroups analyses by population risk or mode of verification. The evidence for these two cut-offs were rated down once more for inconsistency, leading to low certainty of evidence. For four cut-offs (3 mm, 4 mm, 1:270, 1:300), confidence intervals were wide and crossed the clinical threshold set at 50% (which would denote similar percentages of detected and missed cases), hence they were rated down once more due to imprecision, for a low certainty of evidence. Four cut-offs (95th percentile, 99th percentile, 1:250, 1.5 MoM) did not have serious inconsistency or imprecision and were assigned moderate certainty of evidence.

Table 2. Summary Table of Diagnostic Performance for each NT Cut-off Point

Cut-off value	No. of studies	Sensitivity	Specificity	Absolute Effects per 1,000 women	
Moderate Certainty of Evidence					
95th percentile	8 studies N=53,607	72.2% (95% CI: 66.9%-76.9%)	93.1% (range: 75.4%-97.4%)	Detected: 5 (4-5) Missed: 1 (1-2)	False positive: 53 (33-85)
1:250 risk	2 studies N=28,747	68.6% (95% CI: 58.1%-77.5%)	96.7% (range: 95.1%-98.1%)	Detected: 4 (4-5) Missed: 2 (1-2)	False positive: 32 (16-63)
99th percentile	4 studies N=24,779	64.4% (95% CI: 55.9%-72.1%)	98.1% (95% CI: 95.8%-99.1%)	Detected: 4 (4-5) Missed: 2 (1-2)	False positive: 19 (9-42)
1.5 MoM	1 study N=20,743	75.0% (95% CI: 67.0%-82.0%)	95.0% (95% CI: 94.0%-95.0%)	Detected: 5 (4-5) Missed: 1 (1-2)	False positive: 50 (50-60)
Low Certainty of Evidence					
2.5 mm	3 studies N=35,584	64.4% (range: 42.9%-85.4%)	90.4% (range: 72.6%-98.3%)	Detected: 3-5 Missed: 1-3	False positive: 17-273
5% FPR	3 studies N=63,885	60.7% (range: 50.6%-67.9%)	95.0% (95% CI: 94.8%-95.2%)	Detected: 3-4 Missed: 2-3	False positive: 48-52
3.0 mm	9 studies N=18,681	56.8% (95% CI: 47.3%-65.8%)	95.8% (range: 84.4%-98.9%)	Detected: 4 (3-4) Missed: 2 (2-3)	False positive: 32 (19-53)
1:270 risk	1 study N=4,029	80.0% (95% CI: 28.0%-99.0%)	96.0% (95% CI: 96.0%-97.0%)	Detected: 5 (2-6) Missed: 1 (0-4)	False positive: 40 (30-40)
1:300 risk	1 study N=1,801	78.0% (95% CI: 40.0%-97.0%)	96.0% (95% CI: 94.0%-96.0%)	Detected: 5 (3-6) Missed: 1 (0-3)	False positive: 40 (40-60)
4.0 mm	1 study N=3,604	43.0% (95% CI: 10.0%-82.0%)	99.0% (95% CI: 99%-100%)	Detected: 3 (1-5) Missed: 3 (1-5)	False positive: 10 (0-10)