

Evaluatie referentiecurve nekplooiingen Nederland

Wendy Koster, Els Lorijn, Rosalinde Sniijders

Inhoudsopgave

| | |
|--|----|
| Achtergrond..... | 2 |
| Materiaal en methode..... | 3 |
| Selectie beeldmateriaal..... | 3 |
| Selectie metingen..... | 3 |
| Beoordeling beeldmateriaal methode 1..... | 3 |
| Beoordeling beeldmateriaal methode 2..... | 3 |
| Resultaten..... | 4 |
| Beeldbeoordeling..... | 4 |
| Selectie van de dataset voor de referentiecurve..... | 5 |
| Referentiecurve..... | 5 |
| Conclusie..... | 9 |
| Beoordeling meetwaarden..... | 9 |
| Beoordeling beeldmateriaal..... | 9 |
| Vervolgadvies..... | 10 |
| Scoringsmethode voor de CRL..... | 11 |
| Vervolgadvies..... | 11 |
| Referenties..... | 12 |

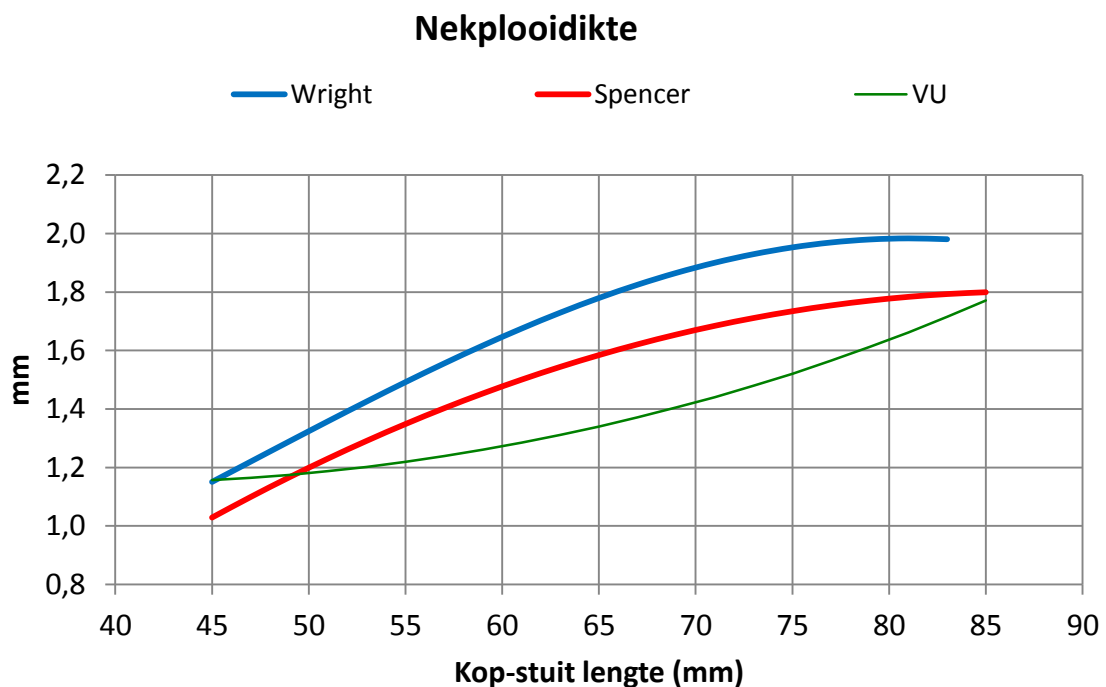
Achtergrond

Bij de introductie van eerste trimester screening begin jaren '90 in Engeland is door de Fetal Medicine Foundation (FMF) een audit programma opgezet voor beoordeling van de kwaliteit van nekplooi (NT) metingen. Bij de jaarlijkse audit worden zowel beeldmateriaal als meetwaarden beoordeeld. Als beeldmateriaal of de verdeling van de metingen niet optimaal zijn krijgt de echoscopist gericht advies om de techniek bij te stellen. Dit heeft ertoe geleid dat centra die deelnemen aan de FMF audit relatief goed presteren (Pandya et al 1995; Snijders et al 2002; Comas et al 2012).

In de loop van de jaren is in verschillende onderzoeken aangetoond dat het bij de beoordeling van beeldmateriaal van belang is om de criteria goed te omschrijven om interpretaties verschillen van beoordelaars te minimaliseren (Herman et al 1998). Bovendien blijkt uit een aantal onderzoeken dat bepaalde items die gebruikt worden om beelden te beoordelen beter te scoren zijn dan andere (Wojdemann et al 2001, Nisbet et al 2011). Het doel van de huidige evaluatie is om de werkzaamheid van twee in Nederland gehanteerde methoden voor beoordeling van beeldmateriaal te vergelijken en om op basis van de bevindingen te komen tot een uniforme reproduceerbare aanpak voor audit.

Voor de evaluatie van de verdeling van NT metingen worden in Nederland momenteel drie verschillende referentiecuren gebruikt. Echocentra die de kansberekening niet door het laboratorium laten doen gebruiken de methode van Wright et al (2008); zes van de zeven laboratoria gebruiken de curve van Spencer et al (2005), en het VUMC gebruikt een eigen curve (Figuur 1).

Figuur 1: Huidige referentiecuren voor foetale nekplooi metingen (NT) in Nederland



Materiaal en methode

Selectie beeldmateriaal

Via de regionale Stichtingen Prenatale Screening (SPS) in Utrecht, Groningen en Amsterdam zijn van 34 echoscopisten drie willekeurig geselecteerde beelden van NT metingen opgevraagd. Deze beelden zijn door een van de onderzoekers (Wendy Koster) geanonimiseerd en in PowerPoint gezet. Deze beelden zijn vervolgens beoordeeld door Els Lorijn en Rosalinde Sniijders volgens zowel de methode van de SPSRU (methode 1) en de SPSNN en SPSAO (methode 2) zonder informatie over de herkomst van de beelden.

Selectie metingen

Uit de beoordelingen zijn 25 echoscopisten geselecteerd waarvoor beide beoordelaars concludeerden dat de metingen goed of uitstekend waren. Metingen van deze echoscopisten zijn gebruikt om te bepalen welke curve het meest geschikt is om metingen van echoscopisten in Nederland te evalueren.

Beoordeling beeldmateriaal methode 1

De eerste methode die gebruikt is om het beeldmateriaal te beoordelen is de originele methode van de Fetal Medicine Foundation in Londen (FMF) aangevuld met het item 'focus'. Deze methode wordt gebruikt in de regio Utrecht door de Stichting Prenatale Screening Regio Utrecht (SPSRU):

Tabel 1: Overzicht van items van beoordelingsmethode 1.

| Item | Score | Omschrijving |
|----------------|-----------|---|
| Doorsnede | 0-1-2 | Is de meting gedaan in een mid-sagittale doorsnede? |
| Vergroting | 0-1-2 | Is het beeld uitvergroot zodat alleen borst en hoofd in beeld zijn? |
| Calipers | 0-1-2 | Zijn de calipers 'on-on' geplaatst? |
| Amnion | 0-1 | Is het amnion in beeld gebracht? |
| Positie | 0-1 | Ligt de foetus in een neutrale positie? |
| Gain | 0-1 | Is de gain optimaal ingesteld? |
| Focus | 0-1 | Staat de focus ingesteld op de nek regio? |
| Maximum | 10 | |

Beoordeling beeldmateriaal methode 2

De tweede methode die gebruikt is om het beeldmateriaal te beoordelen is de nieuwe methode van de FMF die gebruikt wordt in de regio's Noordoost Nederland (SPSNN) en Amsterdam en omstreken (SPSAO). In de aangepaste beoordelingsmethode krijgen alle items een 0-1 score en is in beeld brengen van het amnion vervangen door de vraag of de meting op het widestste punt gedaan wordt.

Tabel 2: Overzicht van items van beoordelingsmethode 2.

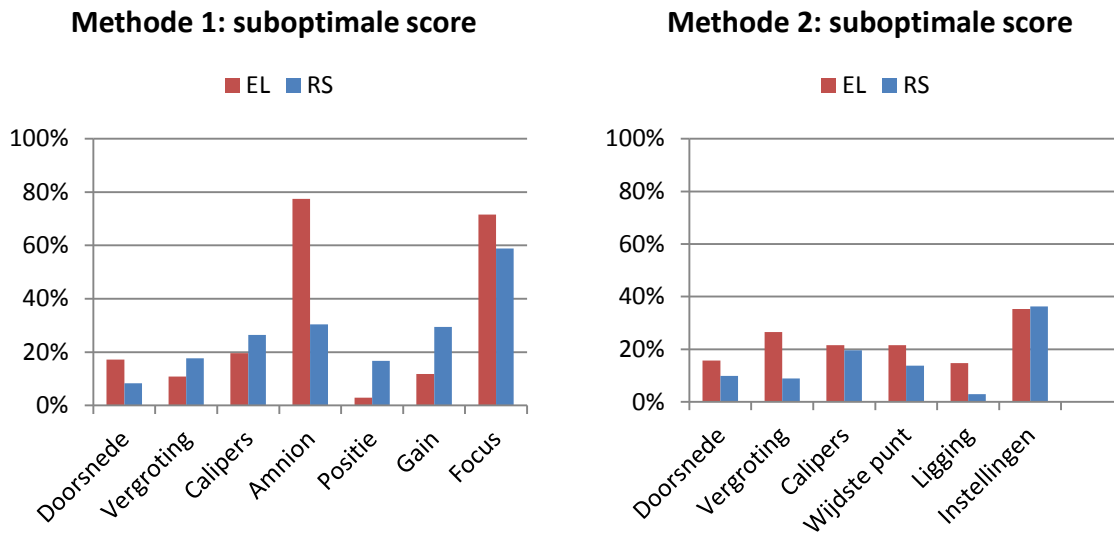
| Item | Score | Omschrijving |
|----------------|----------|---|
| Doorsnede | 0-1 | Is de meting gedaan in een mid-sagittale doorsnede? |
| Vergroting | 0-1 | Is het beeld uitvergroot zodat alleen borst en hoofd in beeld zijn? |
| Calipers | 0-1 | Zijn de calipers 'on-on' geplaatst? |
| Widste punt | 0-1 | Is de meting op het widestste punt gedaan? |
| Ligging | 0-1 | Ligt de foetus in een neutrale positie |
| Instellingen | 0-1 | Zijn optimale instellingen gebruikt? |
| Maximum | 6 | |

Resultaten

Beeldbeoordeling

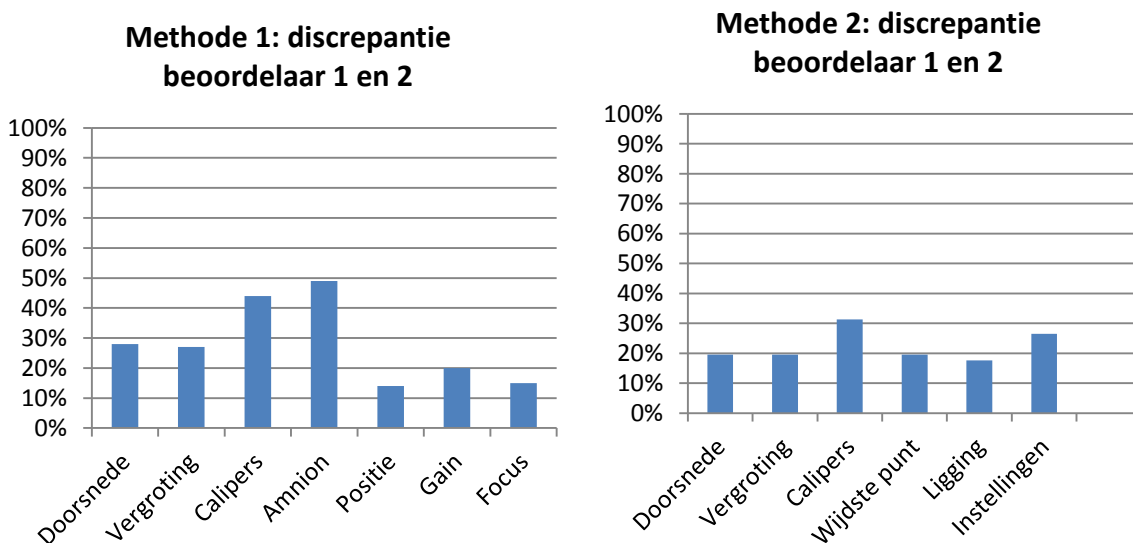
In figuur 2 is per item weergegeven hoe vaak beoordelaars een suboptimale score (0 of 1 bij methode 1; 0 bij methode 2) toekenden bij gebruik van twee in Nederland gehanteerde scoringsmethoden.

Figuur 2: Suboptimale score bij gebruik van respectievelijk beoordelingsmethoden 1 (links) en 2 (rechts) met in rood de score van beoordelaar EL en in blauw de score van RS.



In figuur 3 is weergegeven hoe vaak beoordelaars voor de verschillende items een verschillende score toekenden bij gebruik van de twee scoringsmethoden.

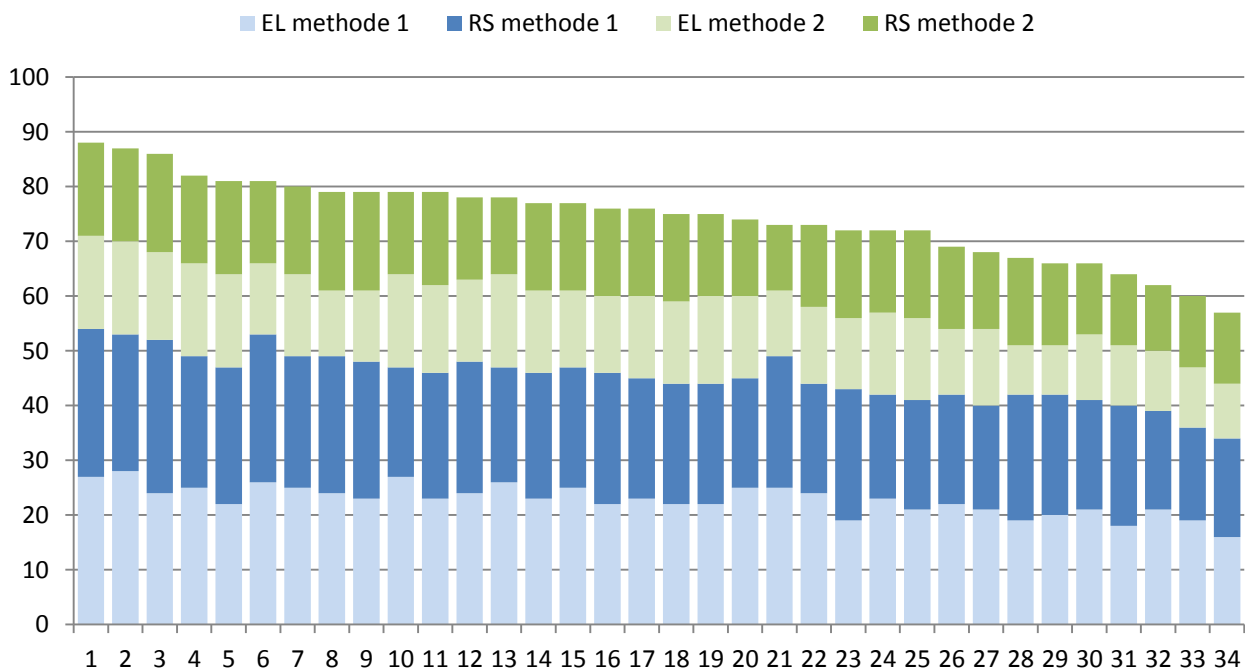
Figuur 3: Discrepantie in de beeldbeoordeling bij gebruik van methode 1 (links) en 2 (rechts).



Selectie van de dataset voor de referentiecurve

Meetwaarden van 25 echoscopisten met een totale score boven de 70 zijn gebruikt om te onderzoeken welke referentie curve het meest geschikt is om de verdeling van de NT metingen van echoscopisten in Nederland te beoordelen. Voor deze selectie zijn de totale scores voor drie beelden van NT metingen per echoscopist berekend. Met beoordelingsmethode 1 was de maximale score 30 (3 x 10 punten) en met beoordelingsmethode 2 was de maximale score 18 (3 x 6). Met de score van twee beoordelaars levert dit maximaal 96 punten (2 x 30 + 2 x 18).

Figuur 4: Totale score voor drie beelden van NT metingen bij beoordeling door twee personen (EL en RS) volgens twee scoringsmethoden.



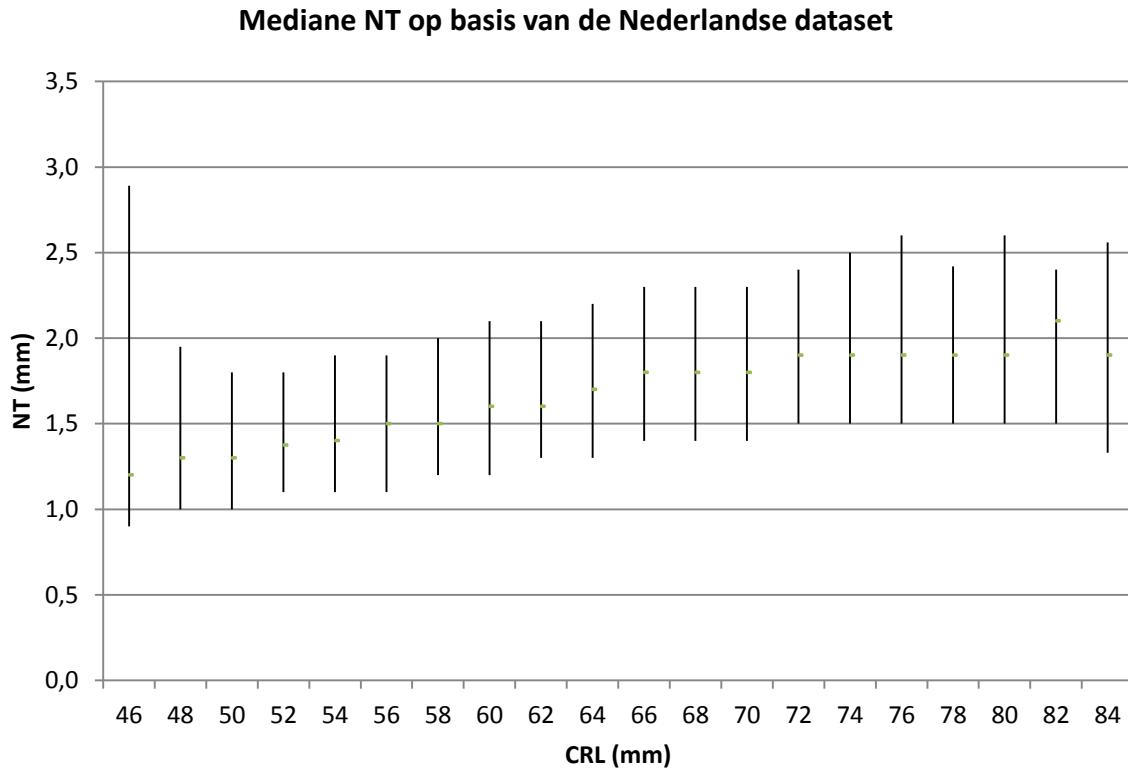
Referentiecurve

Op basis van de beeldbeoordeling zijn gegevens geselecteerd van 25 echoscopisten. Hierbij valt op dat relatief veel echoscopisten (n=11) en meetwaarden (n=4096) afkomstig zijn uit de SPSNN regio. Tabel 3 laat zien dat de mediane CRL en NT in de SPSNN regio iets hoger liggen dan de corresponderende medianen in respectievelijk de SPSAO en SPSU regio's.

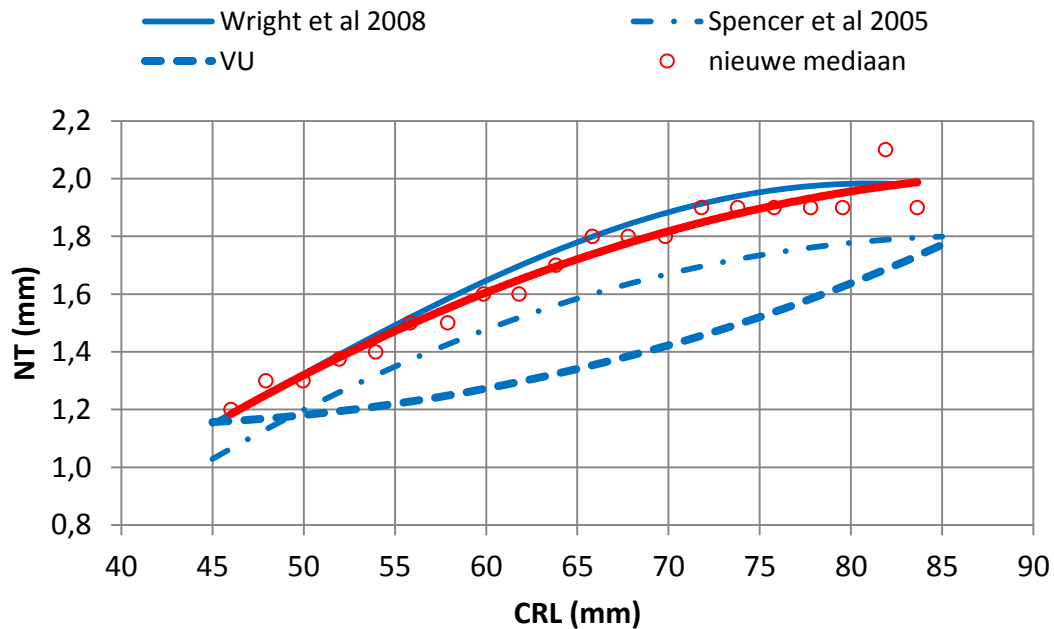
Tabel 3: Mediane kop-stuit lengte (CRL) en nekplooiemeting (NT) per echoscopist met mediane MoM volgens de curve van respectievelijk Spencer en Wright.

| Echoscopist | N | Mediaan CRL | Mediaan NT (mm) | Mediaan NT MoM Spencer | Mediaan NT MoM Wright |
|---------------------|------|-------------|-----------------|------------------------|-----------------------|
| SPSAO totaal | 3535 | 59.9 | 1.6 | 1.08 | 0.97 |
| 1 | 466 | 59.0 | 1.6 | 1.07 | 0.96 |
| 2 | 633 | 59.4 | 1.7 | 1.14 | 1.03 |
| 3 | 1169 | 60.1 | 1.5 | 1.03 | 0.93 |
| 4 | 190 | 59.4 | 1.5 | 1.00 | 0.90 |
| 5 | 393 | 59.8 | 1.7 | 1.13 | 1.01 |
| 6 | 684 | 60.3 | 1.6 | 1.09 | 0.98 |
| SPSNN totaal | 4096 | 61.4 | 1.7 | 1.10 | 0.99 |
| 7 | 250 | 61.3 | 1.7 | 1.15 | 1.04 |
| 8 | 404 | 63.9 | 1.7 | 1.12 | 1.00 |
| 9 | 309 | 60.2 | 1.6 | 1.06 | 0.95 |
| 10 | 907 | 61.5 | 1.7 | 1.11 | 0.99 |
| 11 | 491 | 61.8 | 1.7 | 1.09 | 0.99 |
| 12 | 432 | 62.0 | 1.6 | 1.05 | 0.94 |
| 13 | 265 | 60.3 | 1.7 | 1.16 | 1.04 |
| 14 | 191 | 60.3 | 1.7 | 1.14 | 1.02 |
| 15 | 281 | 61.4 | 1.6 | 1.05 | 0.94 |
| 16 | 315 | 60.3 | 1.7 | 1.13 | 1.01 |
| 17 | 251 | 61.3 | 1.7 | 1.11 | 0.99 |
| SPSRU totaal | 2032 | 60.6 | 1.5 | 1.05 | 0.94 |
| 18 | 190 | 58.3 | 1.4 | 0.96 | 0.87 |
| 19 | 416 | 62.0 | 1.5 | 1.03 | 0.92 |
| 20 | 223 | 59.8 | 1.5 | 1.03 | 0.93 |
| 21 | 268 | 59.0 | 1.6 | 1.08 | 0.97 |
| 22 | 143 | 59.5 | 1.5 | 1.07 | 0.94 |
| 23 | 206 | 62.6 | 1.6 | 1.06 | 0.95 |
| 24 | 324 | 61.9 | 1.6 | 1.07 | 0.96 |
| 25 | 262 | 59.6 | 1.6 | 1.05 | 0.94 |

Figuur 5: Mediane nekplooiemeting (NT) met 10^{de} en 90^{ste} percentiel in relatie met kop-stuit lengte (CRL)



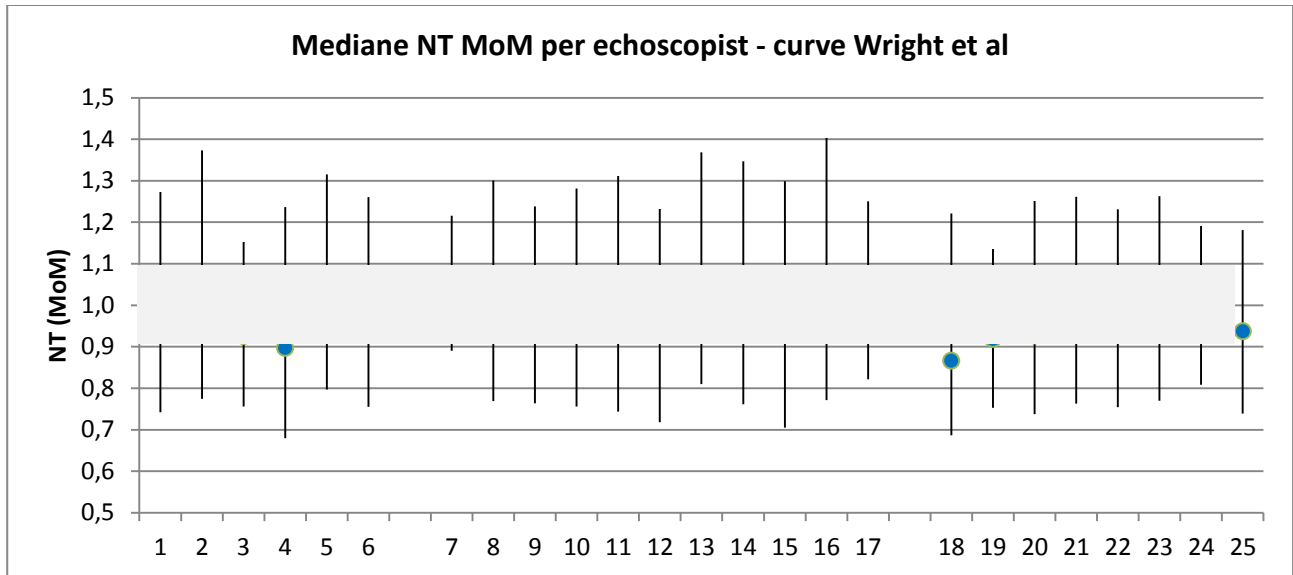
Figuur 6: Mediane NT op basis van de Nederlandse dataset vergeleken met bestaande referentiecuren.



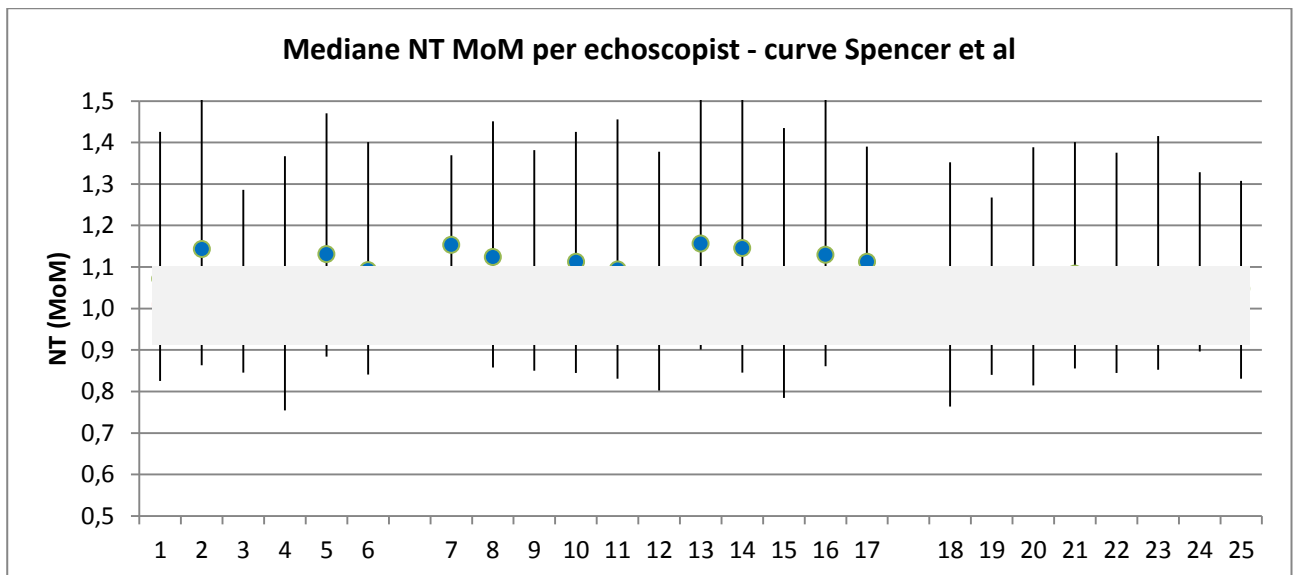
In figure 9 en 10 zijn de mediane NT MoM met 10^{de} en 90^{ste} percentiel per echoscopist weergegeven gebruik makend van respectievelijk de referentie curven van Wright et al en Spencer et al. Met de curve van Wright vallen 23 (92%) van de 25 mediane waarden binnen de door het centraal orgaan gestelde norm van 0.9 tot 1.1 MoM (figuur 9). Met de curve van Spencer vallen 17 (68%) van de mediane NT MoM waarden binnen de norm (figuur 10). De mediane NT MoM waarden van echoscopisten in de regio

Groningen (7-17) liggen gemiddeld iets hoger liggen dan de waarden van echoscopisten in de regio's Amsterdam (1-6) en Utrecht (18-25).

Figuur 7: Mediane nekplooiemeting (NT) in MoM voor kop-stuit lengte met 10^{de} en 90^{ste} percentiel voor 25 echoscopisten uit drie regio's (Amsterdam 1-6, Groningen 7-17 en Utrecht 18-24) gebruik makend van de referentiecurve van Wright *et al.*



Figuur 8: Mediane nekplooiemeting (NT) in MoM voor kop-stuit lengte met 10^{de} en 90^{ste} percentiel voor 25 echoscopisten uit drie regio's (Amsterdam 1-6, Groningen 7-17 en Utrecht 18-24) gebruik makend van de referentiecurve van Spencer *et al.*



Conclusie

Beoordeling meetwaarden

Bij de beoordeling van de meetwaarden blijkt dat de mediane MoM waarde in de regio Utrecht iets lager ligt dan de mediaan in de SPSNN en SPSAO regio's. Dit heeft hoogst waarschijnlijk te maken met het feit dat er in de regio Utrecht een andere referentiecurve (Spencer et al) wordt gebruikt. Het lijkt op basis van deze bevindingen nog niet wenselijk om op een nieuwe landelijke curve over te stappen. Het voorstel is om in alle regio's voorlopig de eigen referentiecurve te behouden en om eerst de nieuwe beeldbeoordeling methode te implementeren zodat de beoordeling en terugkoppeling meer gestandaardiseerd plaatsvinden. Het nu aanpassen van de curves zou kunnen leiden tot over- of onderschatting van combinatierisico's en is daarom op dit moment niet aan te bevelen. Een jaar na bijstelling van de beoordelingsmethode kan nadere evaluatie plaatsvinden om te kijken of de curven dichter bij elkaar liggen. Als er met deze aanpak meer uniformiteit binnen en tussen regio's ontstaat kan op termijn een landelijke curve gehanteerd worden.

Beoordeling beeldmateriaal

Beoordelingsmethode 2 levert minder vaak een suboptimale score op dan methode 1. Een belangrijke reden hiervoor is dat het item *amnion* lastig te beoordelen is op een afdruk. Daarnaast leveren items met een 0-1-2 score uiteraard vaker discrepantie op dan items waarvoor een fout-goed (0-1) score gebruikt wordt. Om te komen tot een uniforme scoringsmethode is het advies om de beoordeling aan te passen op de volgende punten:

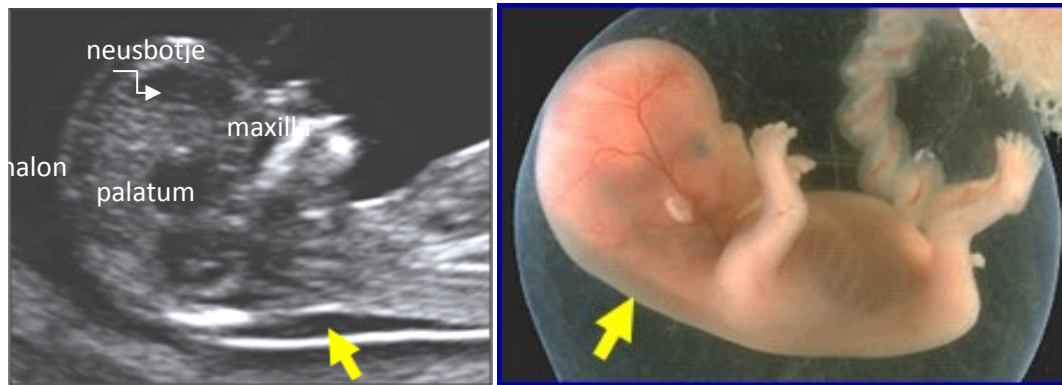
- (1) beoordeling van het amnion in doorsnede voor NT meting vervalt
- (2) beide calipers worden afzonderlijk beoordeeld (1 punt voor elke caliper)
- (3) de locatie waarop de meting wordt verricht wordt meegenomen (wijdste punt)
- (4) er wordt extra aandacht besteed aan instellingen (zoveel mogelijk inzoomen voor freeze, lage gain)
- (5) er wordt een punt toegevoegd voor instelling van de focus op het meetgebied.

Tabel 4: Overzicht van items van de nieuwe beoordelingsmethode.

| Item | Score | Omschrijving |
|------------------|-----------|--|
| Vergroting | 0-2 | Is het beeld uitvergroot zodat alleen borst en hoofd in beeld zijn? |
| Doorsnede | 0-1 | Foetus in midsagittale doorsnede? |
| Ligging Foetus | 0-1 | Foetus in neutrale positie, niet overstrekt, kin niet op de borst? |
| Caliper 1 | 0-1 | Caliper 'on-on' geplaatst? |
| Caliper 2 | 0-1 | Caliper 'on-on' geplaatst? |
| Caliperplaatsing | 0-1 | Staat de 2e caliper op de juiste locatie ten opzichte van de 1e caliper? |
| Wijdste punt | 0-1 | Is de meting verricht op het wijdste punt? |
| Focus | 0-1 | Staat de focus ingesteld op de nekregio? |
| Instellingen | 0-1 | Staat de gain laag, ingezoomd voor freeze? |
| Maximum | 10 | |

NB. Vergroting is een zwaarwegend item. Voldoende vergroot is 2 punten, niet voldoende vergroot is nul punten. Het geven van 1 punt is niet toegestaan bij dit item.

Figuur 11: Doorsnede voor meting van de nekplooidikte



Opties voor logistiek

Er worden 5 casus door regionaal centrum geselecteerd en de echoscopist levert van de beste 3 casus een logboek in;

Regionaal centrum beoordeelt beelden van 5 casus ter plekke en telt daarvan alleen de 3 met de meeste punten.

Vervolgadvies

In totaal kunnen 30 punten behaald worden

26 punten of meer: voldoet aan kwaliteitseisen,

< 26 punten is niet voldoende, advies is dan hands-on training en/of tips en tricks, herbeoordeling van beelden volgens dezelfde methode na een half jaar.

< 22 punten is reden tot overwegen van beëindigen van overeenkomst, afhankelijk van MoM-waarden en vorige beoordelingen.

Evaluatie van het beoordelingssysteem volgt na een jaar.

Scoringmethode voor de CRL

Bij het echoscopisch onderzoek als onderdeel van de combinatietest wordt niet alleen de NT gemeten maar ook de CRL. Deze meting fungeert als uitgangspunt bij de berekening van het risico op trisomie 21, 13 en 18. Bij beeldbeoordeling is het dus van belang ook naar de kwaliteit van de CRL-meting te kijken.

Tabel 5: Overzicht van items van de CRL beoordelingsmethode.

| Item | Score | Omschrijving |
|-------------------|-----------|--|
| Vergroting | 0-1 | Foetus beslaat 2/3 van het beeld |
| Doorsnede* | 0-2-3 | Foetus in midsagittale doorsnede of langste doorsnede? |
| Ligging** | 0-1 | Foetus in neutrale positie, niet overstrekt, kin niet op de borst? |
| Caliper 1 | 0-1 | Caliper op de huid van de kruin? |
| Caliper 2 | 0-1 | Caliper op de huid van de stuit? |
| Langste doorsnede | 0-1 | Is de langste doorsnede gemeten? |
| Focus | 0-1 | Staat de focus ingesteld op de hoogte van het embryo / de foetus? |
| Instellingen | 0-1 | 1e trimester setting, gain functioneel, vruchtwater zwart? |
| Maximum | 10 | |

* midsagitaal: 3 punten; niet midsagitaal, maar wel de langste doorsnede: 2 punten; geen langste doorsnede: 0 punten. **NB.** Doorsnede is een zwaarwegend item, daarom niet de mogelijkheid van het geven van 1 punt.

** niet correct is: embryo in hyperextensie, te sterk geflecteerd hoofd, hoofd te zeer gedraaid ten opzichte van de thorax.

Vervolgadvies

Gelijk aan het advies bij het beoordelen van de NT-metingen.

Bij beoordelen van 3 CRL-metingen kunnen in totaal 30 punten behaald worden

26 punten of meer: voldoet aan kwaliteitseisen, onder 26 punten is niet voldoende, advies hands-on training en/of tips en tricks, herbeoordeling van beelden volgens dezelfde methode na een half jaar.

Referenties

Comas Gabriel C, Echevarria M, Rodriguez I, Serra B. Analysis of quality of nuchal translucency measurements: its role in prenatal diagnosis. *The Scientific World Journal* 2012, Article ID 482832.

Herman A, Maymon R, Dreazen E, Caspi E, Bukovsky I, Weinraub Z, Nuchal translucency audit: a novel imagescoring method. *Ultrasound in Obstetrics and Gynecology* 1998;12-6:398–403.

Nisbet D, McLennan A, Robertson A, Schluter PJ, Hyett J. Reducing inter-rater variability in the assessment of nuchal translucency image quality. *Fetal Diagn Ther* 2011;30:128–134.

Snijders RJM, Thom EA, Zachary JM, Platt LD, Greene N, Jackson LG, Sabbagha RE, K. Filkins K, Silver RK, Hogge WA, Ginsberg NA, Beverly S, Morgan P, Blum K, Chilis P, Hill LM, Hecker J, Wapner RJ, BUN Study Group. First-trimester trisomy screening: nuchal translucency measurement training and quality assurance to correct and unify technique. *Ultrasound Obstet Gynecol* 2002; 19: 353–359.

Spencer K, Cicero S, Atzei A, Otigbah C, Nicolaides KH. The influence of maternal insulin-dependent diabetes on fetal nuchal translucency thickness and first-trimester maternal serum biochemical markers of aneuploidy. *Prenat Diagn* 2005; 25: 927–929.

Wøjdemann KR, Christiansen M, Sundberg K, Larsen SO, Shalmi A, Tabor A. Quality assessment in prospective nuchal translucency screening for down syndrome, *Ultrasound in Obstetrics and Gynecology* 2001;18-6,641–644.

Wright D, Kagan KP, Molina FS, Gazzoni A, Nicolaides KH. A mixture model of nuchal translucency thickness in screening for chromosomal defects. *Ultrasound Obstet Gynecol* 2008; 31: 376–383.