

University of Groningen

Challenges in prenatal screening and diagnosis in the Netherlands

Bakker, Merel

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bakker, M. (2017). *Challenges in prenatal screening and diagnosis in the Netherlands*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.

Copyright


Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Intra-operator and inter-operator
reliability of manual and
semiautomated measurement of
fetal nuchal translucency:
a cross sectional study

M. Bakker¹
P. Mulder¹
E. Birnie^{1,2}
C. M. Bilardo¹

¹ Department of Obstetrics and Gynecology, Fetal Medicine Unit,
University Medical Centre, Groningen, the Netherlands.

² Department of Genetics, University Medical Centre Groningen,
University of Groningen, Groningen, the Netherlands.

Published in *Prenatal Diagnosis* 2013; 33: 1264-1271.

Intra-operator and inter-operator reliability of manual and semiautomated measurement of fetal nuchal translucency: a cross sectional study

M. Bakker¹, P. Mulder¹, E. Birnie^{1,2} and C. M. Bilardo¹

¹ Department of Obstetrics and Gynaecology, Fetal Medicine Unit, University Medical Centre, Groningen, the Netherlands.

² Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands.

Objective:

The goal of this study was to examine the intra-operator and inter-operator differences of the manual and semiautomated nuchal translucency (NT) measurements and to evaluate if these differences alter women's risk status.

Methods:

A cross sectional study was performed. Two operators obtained manual and semiautomated NT measurements of 153 NT images. The maximal acceptable difference in NT measurements within and between operators was 0.15 mm. Intra and inter-operator differences were analyzed by the paired Student's t-test and homogeneity of variances by the Levene's test. Intra-operator and inter-operator agreement were quantified with Bland and Altman's limits of agreement, and changes in women's risk sta-

tus were tested with the binomial test.

Results:

Intra-operator agreement was high for each of the measurement methods. Operator 1 had lower SDS for manual measurements. Conversely, operator 2 had lower SDS of the differences for semiautomated measurements, although the SD never reached the same level as operator 1. Inter-operator agreement was highest for the semiautomated measurements. Changes in risk status occurred between the manual and inner-middle method resulting in different clinical policies in up to 1 out of 20 cases.

Conclusion:

Well-trained operators do not seem to benefit from the use of the semiautomated measurement methods.

Introduction

The combined test (CT) offers women an individual risk assessment for trisomy 21, 13, and 18 in the first trimester of pregnancy.^{1,2} The test is based on the measurement of the fetal nuchal translucency (NT), maternal age and maternal serum markers (β -hCG and PAPP-A). The NT is the most effective marker of trisomy 21 and is able to detect about 75% to 80% of the affected fetuses for a false positive rate of about 3% to 5%.^{1,3} Moreover, an enlarged NT is associated with other chromosomal anomalies, genetic syndromes, and

structural anomalies.⁴ It is therefore important that the NT is measured precisely.

The Fetal Medicine Foundation (FMF) has developed guidelines to promote a standardized measurement technique for obtaining a valid and precise manual NT measurement. Aim of the guideline is to achieve uniformity among different operators and guarantee a reliable risk assessment.⁵ However, the acquirement of the midsagittal plane, the selection of the area containing the maximum NT and the placement of the calipers are still prone to error with the manual technique, which may compromise the performance of screening.⁶⁻⁸

In order to minimize variability in the measurement of the NT, a semiautomated NT measurement has been developed recently (sono-NT, GE Medical Systems). Standardization through semiautomated measurement is thought to lower the standard deviation (SD) of the distribution of NT measurements, increase its precision, and enhance the correct discrimination of normal from trisomic fetuses. Recent studies show that semiautomated NT measurement indeed helps to standardize the NT assessment by lowering the SD, especially in inexperienced sonographers.⁹⁻¹² The aim of this study was to examine the intra-operator and inter-operator agreement of the manual and semiautomated measurements, and to evaluate if differences were not only significant in terms of precision but also in terms of changed individual risk status.

Methods

DESIGN

A cross sectional study on the reliability of the manual and semiautomated measurements of the NT was conducted at the University Medical Centre Groningen. We selected NT images from singleton pregnancies obtained at 11+0 to 13+6 weeks of gestation, stored between 2011 and 2012, and satisfying the FMF guidelines. All images were acquired transabdominally, with harmonics on, using a Voluson E8 equipped with a 4 to 8Hz probe (GE Medical Systems). For each image, two FMF-certified operators measured the NT using the manual first and semiautomated NT measurement technique (SONO-NT[®]: inner-inner and inner-middle method) second. Operator 1 is FMF-certified since May 2010 and operator 2 since October 2009. Each operator performs more than 100 NT measurements annually. Each operator performed the measurements twice, with an interval of 2 to 4 weeks between the first and second measurement to minimize any recall bias. Because we only used images without previous measurements, both operators were blinded for all measurement methods.

NUCHAL TRANSLUCENCY MEASUREMENT USING SONO-NT[®]

After the midsagittal plane is acquired manually, the operator places an adjustable box over a large area of the NT containing the thickest part. The software uses the original image within this box and a corresponding 'edge image', that reflects the changes in brightness rather than the brightness itself, to define the echogenic lines delineating the NT that should be used for measurement. The upper caliper will be automatically placed

on the inner border of the upper line, and the operator then decides whether the lower caliper is placed on the inner border (*inner-inner method*) or on the middle of the lower echogenic line (*inner-middle method*). The software connects every point on one line to all possible points on the other line. For each point on the first line, the shortest distance to the other line is calculated and the maximum NT distance from all these distances is selected. For the inner-middle method, both inner and outer borders of the lower echogenic line are detected, and the midpoint between them is calculated.^{9,10}

STATISTICAL ANALYSIS

Comparisons of NTS between the inner-inner method and the manual method, as well as between the inner-middle method and the manual method were described with conventional statistics. Intra-operator differences of the NTS measured with the manual, inner-inner and inner-middle methods were described as mean difference (SD, 95% CI) and as standard deviation of the difference between the two measurements divided by $\sqrt{2}$. Inter-operator differences of the NTS measured with the manual, inner-inner and inner-middle methods were described as mean difference (SD, 95% CI) of operator 1 to that of operator 2. Differences in SDS (homogeneity of variances) between operators were tested with the Levene's test.

Intra-operator and inter-operator agreement were quantified descriptively with Bland and Altman's limits of agreement (LOAS) and their 95% confidence intervals. The maximal clinically acceptable mean difference in NT measurements within and between operators was considered to be 0.15 mm⁷, and tested with the one-sample Student's

■ Table 1 – Characteristics

N=153	OPERATOR 1			OPERATOR 2		
	Manual	Inner-inner	Inner-middle	Manual	Inner-inner	Inner-middle
Mean [CI 95%]	1.878 [1.781, 1.975]	1.799 [1.696, 1.901]	2.040 [1.940, 2.150]	1.857 [1.759, 1.955]	1.759 [1.655, 1.863]	1.970 [1.870, 2.010]
Standard Deviation	0.606	0.643	0.653	0.6140	0.653	0.644
Range	0.940 – 4.840	0.800 – 4.600	1.000 – 5.000	0.660 – 4.490	0.700 – 4.700	1.000 – 5.000
Median [p25 – p75]	1.730 [1.460 – 2.195]	1.700 [1.300 – 2.150]	1.810 [1.410 – 2.115]	1.600 [1.300 – 2.100]	1.900 [1.500 – 2.300]	1.900 [1.500 – 2.300]

■ Table 2 – Intra-operator differences

N=153	OPERATOR 1			OPERATOR 2		
	SD of measurement error	Mean Δ [SD Δ]	P†	SD of measurement error	Mean Δ [SD Δ]	P**
Manual – manual	0.054	0.010 [0.076]	<0.001	0.156	0.008 [0.220]	<0.001
Inner-inner – inner-inner	0.062	-0.005 [0.088]	<0.001	0.120	0.016 [0.170]	<0.001
Inner-middle – inner-middle	0.086	-0.017 [0.122]	<0.001	0.122	-0.012 [0.172]	<0.001
Manual – inner-inner	0.122	-0.080 [0.173]	<0.001	0.180	-0.098 [0.225]	<0.001
Manual – inner-middle	0.123	0.163 [0.174]	0.348	0.145	0.114 [0.205]	0.032

† One-sample t-test (0.15 mm), applied to the differences between the two observations.

A significant p-value implies that the H0: mean Δ >= 0.15 mm is rejected.

Δ Represents difference between measurements within operators.

SD = standard deviation.

t-test. Agreement of NT measurements within or between operators was expressed with the intraclass correlation coefficient (ICC).

The NT measurements were used to estimate the associated risk of trisomy 21, in Astraia, and to classify women as low (<1 : 200) or high risk ($\geq 1 : 200$), keeping the other parameters of the risk calculation unchanged. In the Netherlands, prenatal invasive testing is offered when the risk estimate after the CT is $\geq 1 : 200$. A difference in NT measurements was considered clinically relevant when a change of woman's risk status occurred (from low to high risk or vice versa). The proportion of changes in women's risk status compared to 0 was tested with the binomial test.

Results

A total of 153 NT images were selected for the study. Median crown-rump length (CRL) was 66.6 mm (IQR 58.5 – 71.4 mm) and median NT was 1.8 mm (IQR 1.5 – 2.2 mm). Table 1 shows the characteristics of the NT measurements for each operator.

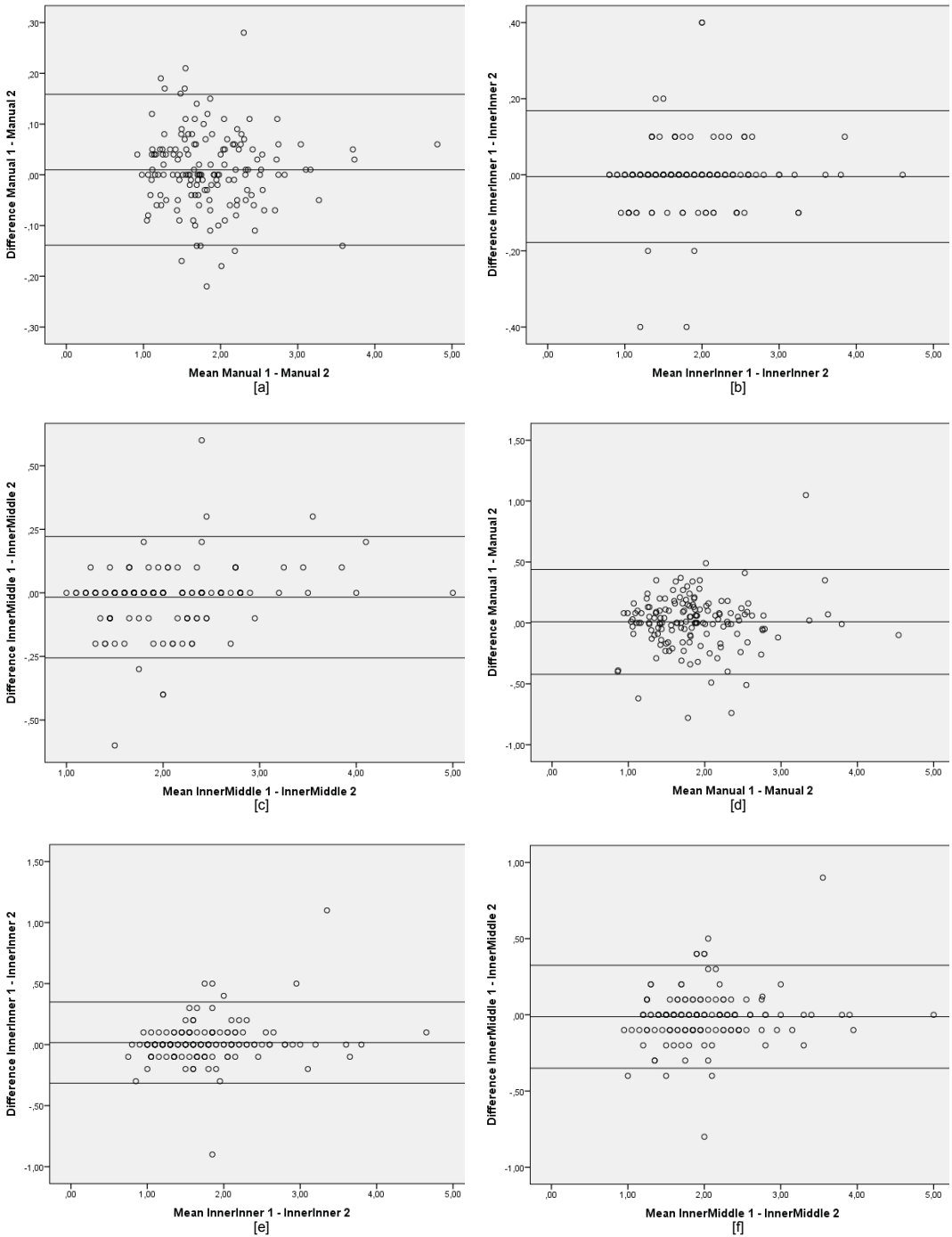
INTRA-OPERATOR AGREEMENT OF MEASUREMENTS

For both operators, the lowest mean NT measurement was obtained by the inner-inner method and the highest by the inner-middle method, whereas the mean of the manual method fell in between. The lowest SD for both operators was obtained when the measurement was performed manually (Table 1).

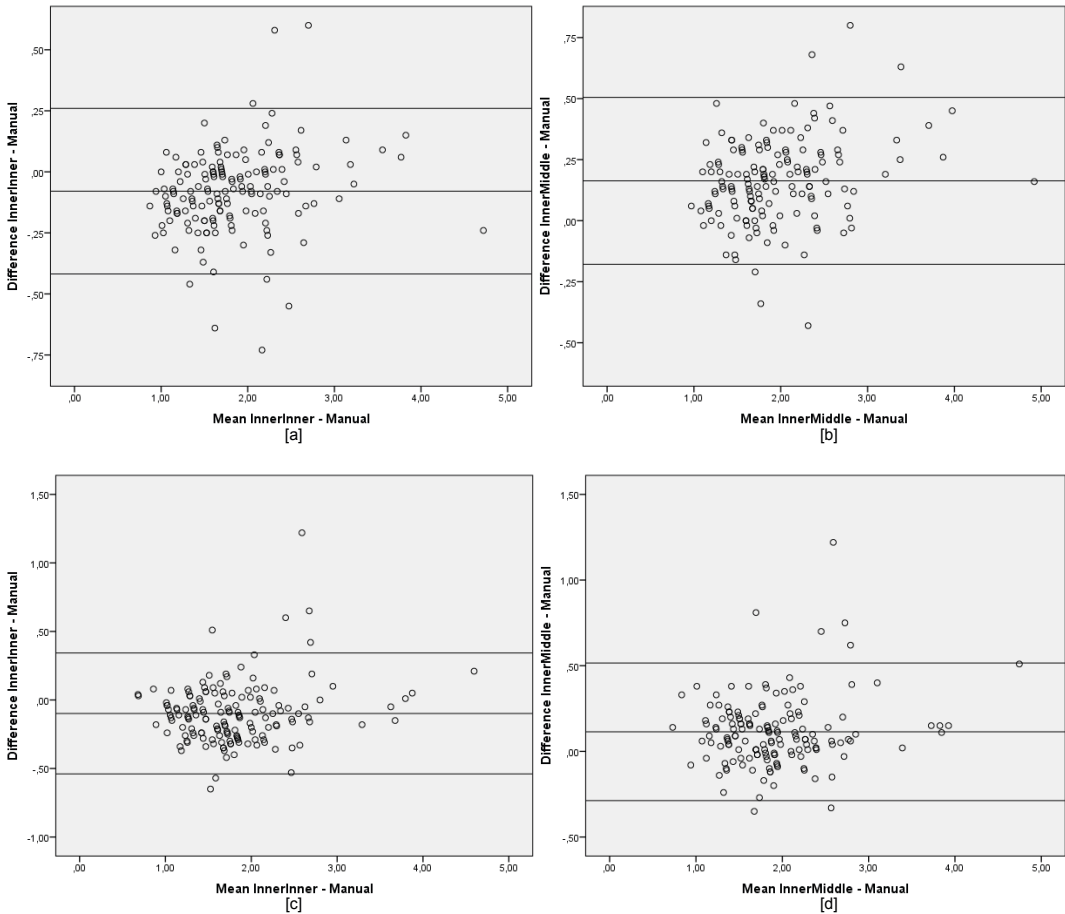
■ Table 3 – Intra-operator limits of agreement

N=153	OPERATOR 1			
	ICC (95% CI)	RC	[LOA]	LOA 95% CI
Manual – manual	0.992 (0.989 – 0.994)	0.149	[-0.139 – 0.159] [-0.178 – 0.169]	(-0.1392 – -0.1388) (0.1588 – 0.1592)
Inner-inner – inner-inner	0.990 (0.987 – 0.993)	0.173	[-0.178 – 0.169]	(0.1783 – -0.1777) (0.1687 – 0.1693)
Inner-middle – inner-middle	0.982 (0.975 – 0.987)	0.239	[-0.256 – 0.222]	(-0.2566 – -0.2554) (0.2214 – 0.2226)
Manual – inner-inner	0.954 (0.917 – 0.972)	0.339	[-0.418 – 0.260]	(-0.4192 – -0.4169) (0.2589 – 0.2612)
Manual – inner-middle	0.931 (0.644 – 0.974)	0.342	[-0.178 – 0.505]	(-0.1792 – -0.1768) (0.5038 – 0.5062)
N=153	OPERATOR 2			
	ICC (95% CI)	RC	[LOA]	LOA
Manual – manual	0.935 (0.911 – 0.952)	0.431	[-0.422 – 0.439]	(-0.4239 – -0.4201) (0.4371 – 0.4409)
Inner-inner – inner-inner	0.964 (0.951 – 0.974)	0.333	[-0.317 – 0.349]	(-0.3181 – -0.3159) (0.3479 – 0.3501)
Inner-middle – inner-middle	0.963 (0.949 – 0.973)	0.338	[-0.350 – 0.325]	(-0.3511 – -0.3489) (0.3239 – 0.3264)
Manual – inner-inner	0.926 (0.873 – 0.954)	0.441	[-0.540 – 0.343]	(-0.5420 – -0.5381) (0.3411 – 0.3450)
Manual – inner-middle	0.932 (0.853 – 0.963)	0.401	[-0.287 – 0.516]	(-0.2886 – -0.2854) (0.5144 – 0.5176)

ICC = intraclass correlation coefficient. RC = repeatability coefficient. LOA = limits of agreement. CI = confidence interval.



■ Figure 1 – Bland Altman plots and their corresponding limits of agreement for differences between repeated measurements within operators.



■ Figure 2 – Bland Altman plots and their corresponding limits of agreement for differences between measurement methods within operators.

Mean intra-operator differences and SDS for the various measurement methods are shown in Table 2. Operator 1 had a lower SD for measurements performed manually than for measurements performed by the inner-inner and innermiddle method. The SD of operator 2, however, was higher using the manual method than for the inner-inner and inner-middle methods, although this difference was not significant.

Figures 1 and 2 and Table 3 show Bland Altman plots and their corresponding limits of agreement (LOA) for differences between repeated measurements within operators and between measurement methods within operators. The LOAS for differences between repeated measurements were $>+0.15$ mm or <-0.15 mm for both operators, except for the upper limit of agreement of the repeat manual measurement of operator 1. However, the mean differences for the repeat measurements were significantly smaller than 0.15 mm for both operators as tested with the one-sample *t*-test, see Table 2.

The LOAS for differences between measurement methods were $>+0.15$ mm or <-0.15

mm. The mean differences between different measurement methods were however significantly smaller than 0.15 mm for both operators, except for the difference between the manual and inner-middle method of operator 1 (Table 2).

For each operator, the agreement between repeat measurements and between different measurement methods was high (Table 3).

Operator 1 had the highest predictability for the repeat manual measurements (ICC 0.992). The ICC was 0.990 for the repeat innerinner measurements and 0.982 for repeat inner-middle measurements. Agreement between the different measurement methods was lower, ICC 0.954 (manual vs. inner-inner) versus ICC 0.931 (manual vs. inner-middle).

For operator 2, agreement of the repeated inner-inner measurements (ICC 0.964) was higher than the repeated manual measurements (ICC 0.935) and inner-middle measurements (ICC 0.963). Predictability between different measurement methods was a lower, ICC 0.926 (manual vs. inner-inner) versus ICC 0.932 (manual vs. inner-middle).

INTRA-OPERATOR DIFFERENCES OF ESTIMATED RISK

Changes in estimated risk of trisomy 21 between repeated measurements were found (Table 4). For operator 1, these differences were only significant for the inner-middle method (0.048) and for the comparison between the manual and semiautomated methods (manual-inner-inner: 0.017, manual-inner-middle: <0.001). For operator 2, these differences were significant for the manual method (0.048), the inner-middle method (0.048), and for the comparison between the manual and semiautomated methods (manualinner-inner: <0.001, manual-inner-middle: <0.001).

■ Table 4 – Intra-operator differences in risk status (risk <1 : 200 or ≥1 : 200)

	OPERATOR 1		OPERATOR 2	
	Discordant pairs (%) [CI 95%]	p*	Discordant pairs (%) [CI 95%]	p*
Manual – manual	0/153 (0.0%) [0 – 0.024]	>0.99	3/153 (2.0%) [0.004 – 0.056]	0.048
Inner-inner – inner-inner	1/153 (0.7%) [0.002 – 0.036]	0.360	2/153 (1.3%) [0.002 – 0.046]	0.130
Inner-middle – inner-middle	3/153 (2.0%) [0.004 – 0.056]	0.048	3/153 (2.0%) [0.004 – 0.056]	0.048
Manual – inner-inner	4/153 (2.6%) [0.007 – 0.066]	0.017	8/153 (5.2%) [0.023 – 0.100]	<0.001
Manual – inner-middle	8/153 (5.2%) [0.023 – 0.100]	<0.001	7/153 (4.6%) [0.019 – 0.092]	<0.001

* Binomial test.

CI = confidence interval.

■ Table 5 – Inter-operator differences

	N	Mean Δ	SD of Δ	Equality of variance p*	p**
Manual (operator 1 – operator 2)	153	0.021	0.247	0.767	<0.001
Inner-inner (operator 1 – operator 2)	153	0.040	0.199	0.591	<0.001
Inner-middle (operator 1 – operator 2)	153	0.070	0.200	0.347	<0.001

* Levene's Test for differences in SD - H0: mean manual (operator 1) = mean manual (operator 2). A non-significant p-value implies that H0 cannot be rejected.

** One-sample t-test (0.15 mm), applied to the differences between the two observations.

A significant p-value implies that the H0: mean Δ >= 0.15 mm is rejected.

Δ Represents difference between operator 1 and operator 2.

SD = standard deviation.

INTER-OPERATOR AGREEMENT OF MEASUREMENTS

Mean differences and SDS of the differences between operators for each method are shown in Table 5. The mean difference between operators was the lowest for the manual method. The SD of the differences of the manual method (0.247) was higher than the SD of the inner-inner (0.199) and inner-middle methods (0.197), but these SDS did not differ significantly (Table 4).

The mean differences between operators for the different measurement methods were significantly smaller than 0.15 mm (Table 5). Agreement was the highest for the semiautomated measurements and the LOAS for the different measurement methods were all $>+0.15$ mm or <-0.15 mm (Table 6).

INTER-OPERATOR DIFFERENCES OF ESTIMATED RISK

In view of the inter-operator differences, women's risk status altered in 2.6% to 3.3% of the cases, depending on the measurement method. All differences in estimated risk of trisomy 21 between the operators were significant (Table 7).

Discussion

In this study, we evaluated differences in intra-operator and inter-operator agreement between the manual and semiautomated measurements, and studied the clinical relevance of lack of agreement in terms of altered CT risk assessment.

The intra-operator agreement was high for each of the measurement methods. Operator 1 had lower SDS for measurements performed manually. Conversely, operator 2 had lower SDS of the differences for semiautomated measurements, although the SD of the

■ Table 6 – Inter-operator differences

	ICC (95% CI)	RC	LOA	LOA (95% CI)
Manual (operator 1) – (operator 2)	0.918 (0.889 – 0.940)	0.484	[-0.463 – 0.505]	(-0.4653 – -0.4607) (0.5027 – 0.5073)
Inner-inner (operator 1) – (operator 2)	0.951 (0.933 – 0.965)	0.390	[-0.350 – 0.430]	(-0.3515 – -0.3485) (0.4285 – 0.4315)
Inner-middle (operator 1) – (operator 2)	0.949 (0.921 – 0.966)	0.386	[-0.316 – 0.455]	(-0.3175 – 0.3145) (0.4535 – 0.4565)

ICC = intraclass correlation coefficient. RC = repeatability coefficient.
LOA = limits of agreement. CI = confidence interval.

■ Table 7 – Inter-operator differences in risk status (risk $<1 : 200$ or $\geq 1 : 200$)

	Discordant pairs (%) [CI 95%]	p*
Manual (operator 1) – manual (operator 2)	5/153 (3.3%) [0.019 – 0.092]	0.006
Inner-inner (operator 1) – inner-inner (operator 2)	5/153 (3.3%) [0.019 – 0.092]	0.006
Inner-middle (operator 1) – inner-middle (operator 2)	4/153 (2.6%) [0.007 – 0.066]	0.017

* Binomial test.

CI = confidence interval.

difference never reached the same level as operator 1. It seemed that operator 2 benefited from the use of the semiautomated measurements.

Agreement between operators was highest for the semiautomated measurements suggesting that the use of semiautomated measurements results in lower variability between operators (standardization), however, differences in variance were not significant. There seems to be no real benefit of the semiautomated measurements over the manual measurement in terms of measurement variability.

Earlier studies have addressed the reproducibility of the manual and semiautomated NT measurement and reported on the benefit of the use of the semiautomated method.^{9,10,12} Moratalla et al. found a between-operator SD of 0.0149 for the inner-middle method, 0.109 for the manual method, and high ICCs for both methods (0.98 for the inner-middle method and 0.85 for the manual method). Their conclusion was that the measurement of the NT is more reliable when a semiautomated approach is used rather than the manual method, because it will reduce variation between operators.⁹ Kagan et al. found an estimated SD of differences between operators of -0.02 for the inner-inner method, -0.03 for the inner-middle method and -0.11 for the manual method.¹² They concluded that especially non-experts will benefit from this method as it helps to standardize caliper placement. Abele et al. found that the inter-operator variability for all measurement methods was similar for experts, and these were lower than those of less experienced sonographers. They also concluded that inexperienced operators will benefit from the semiautomated method, but they added that the most important contributor to inter-operator variability is image acquisition.¹⁰ Our results are consistent with the aforementioned studies; however, we could interpret these studies conversely and state that the semiautomated measurement does not seem to be useful in well-trained operators.

Although the inter-operator and intra-operator differences were larger than the 0.15 mm threshold in this study, the impact of measurement variability within and between operators on women's risk status was small. However, these small differences translated in different clinical policies in up to 1 out of 20 cases, which we clinically cannot neglect.

Both operators in our study are FMF accredited and perform the same number of NT measurements annually, except operator 2 has more years of experience as a sonographer. Most studies evaluating measurement variability between the manual and semiautomated methods conclude that less experienced operators will benefit from the semiautomated measurement as it helps standardize caliper placement.¹⁰⁻¹² Paradoxically, in our study, it was the more experienced operator 2 who seemed to benefit from the use of the semiautomated method. This suggests that more experience does not necessarily imply accurate measurements. Not only experience but also attitude as part of competence seems also prerequisite for obtaining precise NT measurements. This is reflected by the higher SD between measurements of operator 2 in comparison with operator 1.

Precise measurement of the NT not only consists of the selection of the thickest part and accurate placement of calipers, but also requires the appropriate acquisition of the midsagittal plane and accurate measurement of the CRL. Use of the semiautomated measurement theoretically can help reducing measurement variability in the first two steps, however correct image acquisition and measurement of the CRL still remain operator dependent. Kagan et al. showed that an underestimation and overestimation of the CRL has a major impact, resulting in substantial underestimation or overestimation of those

risks.¹³ Also, deviation from the midsagittal plane can have a great impact on women-specific risks, and this impact is higher than measurement variability due to caliper placement.^{14,7} The importance of factors influencing image quality, for example, gain and harmonics, are well known.¹⁵ From a technical point of view, the main contributor to intra-operator and inter-operator differences is thus not caliper placement, but the acquisition of the optimal NT image complying with all the FMF criteria.

This means that precision of NT measurements is still largely dependent on appropriate training, adherence to strict criteria, coincidental circumstances (e.g. fetal position, maternal BMI, and time allocated), and, last but not least, on the operator's personal attitude in terms of endurance and accuracy.

At this moment, there seems to be no real advantage of the semiautomated measurement method over the manual method to warrant its implementation on a large scale. It is unclear which operator will truly benefit from semiautomated measurements because well-trained operators do not appear to benefit from this method, and secondly, how this selection should be performed because attitude besides experiences seems to play a major role. A limitation of this study is that no more than two operators were included. Our results may be reassessed with multiple operators of various levels of experience and attitude in future studies.

At present, normal ranges for NT measurements used in the FMF risk assessment algorithm are based on manual NT measurements. No study has evaluated the impact of the semiautomated methods on the screen positive rate. Advances that may have a larger impact in reducing inter-operator variability are software packages that provide the possibility of selecting the right midsagittal plane while scanning with the aid of a 3D.¹⁶

Conclusion

Well-trained operators do not seem to benefit from the use of the semiautomated measurement methods. Manual measurement of the NT according to the FMF guidelines is mandatory and, in itself, should be sufficient to obtain a reliable risk calculation for prenatal trisomies screening.

References

1. Kagan K. O., Etchegaray A., Zhou Y. et al. Prospective validation of first trimester combined screening for trisomy 21. *Ultrasound Obstet Gynecol* 2009; 34(1): 14-8.
2. Kagan K. O., Wright D., Baker A. et al. Screening for trisomy 21 by maternal age, fetal nuchal translucency thickness, free beta-human chorionic gonadotropin and pregnancy-associated plasma protein-A. *Ultrasound Obstet Gynecol* 2008; 31(6): 618-24.
3. Kagan K. O., Wright D., Valencia C. et al. Screening for trisomies 21, 18 and 13 by maternal age, fetal nuchal translucency, fetal heart rate, free beta hCG and pregnancy-associated plasma protein-A. *Hum Reprod* 2008; 23(9): 1968-75.
4. Bilardo C. M., Pajkrt E., De Graaf I. et al. Outcome of fetuses with enlarged nuchal translucency and normal karyotype. *Ultrasound Obstet Gynecol* 1998; 11(6): 401-6.

5. The fetal medicine foundation, 2010.
6. Pandya P. P., Altman D. G., Brizot M. L. et al. Repeatability of measurement of fetal nuchal translucency thickness. *Ultrasound Obstet Gynecol* 1995; 5(5): 334-7.
7. Kagan K. O., Wright D., Etchegaray A. et al. Effect of deviation of nuchal translucency measurements on the performance of screening for trisomy 21. *Ultrasound Obstet Gynecol* 2009; 33(6): 657-64.
8. Schmidt P., Staboulidou I., Elsasser M. et al. How imprecise may the measurement of fetal nuchal translucency be without worsening first-trimester screening? *Fetal Diagn Ther* 2008; 24(3): 291-5.
9. Moratalla J., Pintoff K., Minekawa R. et al. Semi-automated system for measurement of nuchal translucency thickness. *Ultrasound Obstet Gynecol* 2010; 36(4): 412-6.
10. Abele H., Hoopmann M., Wright D. et al. Intra- and interoperator reliability of manual and semi-automated measurement of fetal nuchal translucency by sonographers with different levels of experience. *Ultrasound Obstet Gynecol* 2010; 36(4): 417-22.
11. Grange G., Althuser M., Fresson J. et al. Semi-automated adjusted measurement of nuchal translucency: feasibility and reproducibility. *Ultrasound Obstet Gynecol* 2011; 37(3): 335-40.
12. Kagan K. O., Abele H., Yazdi B. et al. Intra-operator and interoperator repeatability of manual and semi-automated measurement of increased fetal nuchal translucency according to the operator's experience. *Prenat Diagn* 2011; 31(13): 1229-33.
13. Kagan K. O., Hoopmann M., Baker A. et al. Impact of bias in crown-rump length measurement at first-trimester screening for trisomy 21. *Ultrasound Obstet Gynecol* 2012; 40(2): 135-9.
14. Abele H., Wagner N., Hoopmann M. et al. Effect of deviation from the mid-sagittal plane on the measurement of fetal nuchal translucency. *Ultrasound Obstet Gynecol* 2010; 35(5): 525-9.
15. Pasquini L., Tondi F., Rizzello F. et al. Impact of tissue harmonic imaging on measurement of nuchal translucency thickness. *Ultrasound Obstet Gynecol* 2010; 36(4): 423-6.
16. Cho H. Y., Kwon J. Y., Kim Y. H. et al. Comparison of nuchal translucency measurements obtained using Volume NT(TM) and two- and threedimensional ultrasound. *Ultrasound Obstet Gynecol* 2012; 39(2): 175-80.