General discussion and perspectives
General discussion

In recent years first-trimester screening for Down syndrome (DS) has become an established method in many Western countries where it has replaced both second trimester screening for DS, and the policy of offering invasive diagnostic testing to women of advanced maternal age (AMA). First-trimester screening offers several advantages compared with the other two strategies. It results in a higher detection rate (DR) than second-trimester screening (triple test) and it reduces the number of invasive diagnostic procedures and thus the number of procedure-related miscarriages. Moreover, affected pregnancies will be diagnosed earlier in gestation, allowing prospective parents to decide earlier either to continue pregnancy or to consider termination of pregnancy. The emotional impact is less when pregnancy termination for fetal abnormalities is performed early in gestation. In the Netherlands, a national screening program, open to all pregnant women, exists from January 2007, including first-trimester screening for DS and second-trimester screening for neural tube defects and other structural anomalies. The research described in this thesis was performed at the VU University medical center and studied data of the first-trimester screening program for DS performed in the province North-Holland in the Netherlands over the years 2004 to 2011. The aim of our research was to demonstrate the effect of the different screening policies on the uptake of prenatal screening (PNS) and prenatal diagnostics (PND) in different maternal age groups and, subsequently, to investigate whether adjustments of the current screening policies improves screening performance.

Uptake of prenatal screening and prenatal diagnostic testing

Offering first-trimester screening for DS should result in a decrease of invasive testing. It has been demonstrated that the total number of PND will decline only if significant numbers of women, especially women of advanced maternal age (AMA), will perform the first-trimester combined test (FCT), and then decide, based on the risk result, to perform PND. In the Dutch situation, however there is a relatively low participation rate in the national screening program, as compared with other European countries. In England the number of PND decreased with 72% over a 9-year period with offering the FCT as the primary screening method. Moreover women are not offered PND on the basis of age alone. The introduction of a combined risk assessment in Denmark resulted in an uptake of 85%, in a sharp decline in the number of PND and halved the number of infants born with DS. In this thesis (chapter 2) we confirm a low overall uptake (30%), being more distinct in women ≤ 30 years of age (20%) in comparison to AMA women (> 40%). A significant increase was seen in the uptake of the FCT in the period 2001-2003 to 2007-2010. The participation of FCT increased over the years with approximately 10%, 20%, 35%, 45% and 30% in women aged ≤ 25 years, 26-30 years, 31-35 years, 36-40 years and 41-45 years respectively. The effect on the uptake of PND for AMA however is minimal, only a 10% decrease is seen. Even yet about half of the diagnostic procedures performed is
General discussion

for the indication AMA. Especially women aged ≥ 39 years opt for PND for AMA; a decline in uptake is seen only in women aged 36 - 38 years. Thus the choice for FCT and PND for AMA is found to be age related and seems dependent on the background risk.

Screening performance

Prenatal screening aims at identifying women with a high risk for carrying a fetus with a certain anomaly. Performance of a screening test will be appropriate if the majority of the affected fetuses will be identified and only a small number of women will receive a false-positive result at a certain cut-off value. The choice of the cut-off value determines the balance between DR and false-positive rate (FPR). In the Dutch screening program the cut-off is set at 1:200 at mid-term. Screening performance of the FCT is reported up to 90% at a FPR of 5%11-15. Most studies reported on screening performance in overall screening populations. However, in case of identical values of NT and serum parameters a younger woman will receive a lower risk result than an older women due to the difference in the a priori maternal age risk. Differentiation in screening performance between women of different maternal ages is relevant. In the counselling of the individual pregnant woman the actual DR and FPR valid for her age should be provided to meet the requirements of informed decision-making. Several studies showed increasing DR and FPR with increasing maternal age16-19. Comparison of screening performance between the different maternal age groups based on the values of DR and FPR alone is difficult. The odds of being affected given a positive result (OAPR) provides a good additional estimate. Our research shows an excellent screening performance of the FCT in different maternal age groups (DR 100% and FPR 2.4% in women aged ≤ 25 years, DR 90% and FPR 3% in women aged 26-30 years, DR 95.1% and FPR 4.8% in women aged 31-35 years, DR 94.9% and FPR 12.2% in women aged 36-40 years and DR 100% and FPR 22.5% in women aged 41-45 years). The balance between FPR and DR is even more favourable in women < 36 years of age compared to AMA women with comparable OAPR. Lowering the cut-off level for increased risk improves screening performance. In women < 36 years of age the best test performance is at a cut-off of 1:150 (DR 94.3% and FPR 3.4%) and in AMA women at a cut-off of 1:100 (DR 94.2% and FPR 6.7%) (chapter 3).

Adjustments of serum marker concentrations

With the introduction of population screening adjustments for serum marker concentrations were applied for maternal weight and smoking behavior. Serum marker concentrations decrease with increasing maternal weight. Without applying a correction, the MoM values in heavier women turn out to be lower and opposite in slim women20. Smoking results in lower median PAPP-A MoM values; and does not seem to affect fβ-hCGMoM values21,22. Ethnicity also influences serum marker distribution, especially Afro-Caribbean and Asian women show higher PAPP-A and fβ-hCG levels23. Adjusting for
ethnicity however is difficult due to lack of a complete and correct classification. In twin pregnancies a correction should be applied for both serum markers and a distinction should be made between mono- and dichorionic twins, since biochemical markers are significantly lower in monochorionic twins. The impact of early vaginal bleeding, diabetes, and fetal gender are either too small or too complicated to justify a correction. In pregnancies conceived with assisted reproduction techniques (ART), like in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI), the distribution of first-trimester serum parameters differ from spontaneously conceived pregnancies. Most studies reported decreased PAPP-A levels in ART pregnancies with unaltered fβ-hCG levels, increased fβ-hCG levels, or decreased fβ-hCG levels. And only few studies showed no changes in maternal serum values in ART pregnancies. None of these studies made a difference in ART pregnancies with or without hormone treatment. Amor and colleagues showed in a large ART population that only in uncomplicated ART pregnancies with hormone treatment significantly lower PAPP-A levels and unaltered levels of fβ-hCG were found with a significantly increased likelihood for a receiving a false-positive result. The prevalence of fetal chromosomal abnormalities in ART pregnancies is not higher than in naturally conceived pregnancies with the possible exception of sex chromosomal anomalies. Applying an adjustment for serum markers in ART pregnancies with hormone treatment is advocated to reduce the overestimate of risk. In chapter 4 we determined the adjustments for fβ-hCG and PAPP-A levels by comparison of ART pregnancies (IVF and ICSI cases) with hormone treatment and spontaneously conceived pregnancies, all with a normal pregnancy outcome to exclude the possible effect of adverse pregnancy outcomes on PAPP-A levels. The computed correction factors for PAPP-A and fβ-hCG were 1.42 and 1.17 for the IVF group and 1.56 and 1.05 for the ICSI group. Before adjustments for maternal serum values can be introduced, it is necessary to study the effect on screening performance in unselected (complicated and non-complicated) IVF and ICSI populations compared to an overall screening population. In chapter 5 we concluded that a correction factor for PAPP-A in ART pregnancies (with hormone treatment) is justified and that correction for fβ-hCG is unnecessary. With adjustment of PAPP-A concentrations the FPR in ART pregnancies decreases to the observed FPR of the overall screening population. These results underline the importance of implementation of a correction for PAPP-A in ART pregnancies with hormone treatment. Furthermore it will improve the reliability of the pre- and post-test counselling for women carrying an ART pregnancy. In 2012 adjustments for ART pregnancies has been introduced in the Dutch screening program.

Serum sampling

Maternal serum is collected between 9+0 and 13+6 weeks of gestation. Studies in DS pregnancies have shown that the earlier the serum is sampled PAPP-A levels are lower and therefore PAPP-A discriminates better early in gestation whereas the opposite applies for fβ-hCG which possibly discriminates better the later it is taken in the first trimester of pregnancy. What is the effect on screening performance when serum is sampled early or late in the first-trimester? Do these
characteristics of PAPP-A and fβ-hCG neutralize one another? Schiøtt et al. showed a high positive predictive value of the FCT in a high risk population when serum was sampled before the NT measurement. Only few studies report on DR and FPR with early versus late serum sampling. Kirkegaard and colleagues demonstrated a significant higher DR when serum was sampled before 10+0 weeks of gestation. A Dutch study showed a significant decrease in FPR with serum sampling before 11+0 weeks. In this thesis (chapter 6) we found that a significantly lower FPR is obtained when serum is taken before 11+0 weeks than when it is taken from 11+0 weeks without differences in DR. Differences in FPR are not explained by differences in maternal age. The effect of serum sampling at different gestational weeks on risk assessment is most pronounced in AMA women with serum sampled in week 13.

Maternal age risk

Although AMA is correlated with the presence of an affected pregnancy, the different screening parameters do not correlate with maternal age. DS fetuses present the same in both younger and older women with increased NT, higher fβ-hCG and lower PAPP-A values. The effect of the background risk by maternal age on the risk assessment is substantial. Due to the higher maternal age risk in older women, more women have a false positive result and the lower maternal age risk in younger women results in an increase in the number of false negative cases. Several studies analyzed screening performance by elimination of maternal age risk from the risk algorithm. Two methods were used, the relative risk method and the absolute risk method. With both methods a lowering of the FPR in older women and a tendency for higher DR in younger women was shown. The studies were, however, limited by relatively small population sizes. In a retrospective analysis (chapter 7) we compared screening performance of the absolute risk method with the current first-trimester screening in a large screening population (more than 35,000 data on the FCT). Exclusion of maternal age from the risk assessment leads to a significant decrease in the overall FPR. Moreover the FPR is the same for women of all ages. Compared to the FCT, the decrease in FPR with the absolute risk method is most pronounced in AMA women. DR is comparable between both methods. However a trend towards detection of more DS cases in women aged < 30 years is recognizable as well as a trend towards less DS cases detected in women aged ≥ 36 years. As suggested earlier, based on the uptake of FCT and PND for AMA, maternal age risk may have a significant influence on the choice for first-trimester screening and invasive diagnostic testing. Introduction of the absolute risk method would simplify pre- and post-test counselling and might deal with the misunderstanding that screening performance of first-trimester screening in younger women would be less and that screening in the more older women would not be advisable due to the high background risk. The absolute risk method might lead to a higher uptake of first-trimester screening resulting in a more efficient screening policy.
Informed decision-making

The effect of the introduction of the national screening program for DS on the participation of the FCT is limited. It is of interest to explore the possible reasons why prospective parents refrain from PNS. The participation rate is low in our country, but the national screening program is not directed at achieving high uptake rates. More important is whether the uptake rate is based on informed decision-making. In our country, the information process on PNS is performed in two stages. In the first stage every pregnant woman is informed about the possibility to participate in PNS and subsequently she is asked whether she would like to receive detailed information on the screening options (the ‘information offer’). All women have the right not wanting to receive any information about PNS: the ‘right not to know’. In that case no detailed information is provided. The second stage of the information process starts when the offer is accepted and only then the woman receives detailed information on PNS (the ‘information provision’).

Relevant for the decision-making process, that is known to be complicated\textsuperscript{66-69}, is a careful, non-directive counselling. However it can not be ruled out that the counsellor’s attitude towards PNS is of influence on the information provided. For example, the assumption that screening performance of the FCT in younger women is less considering their low background risk and opposite that screening in the more older women is useless considering their high background risk, might influence the counselling. One study showed that neither uptake rates, nor attitude towards PNS of pregnant women were significantly predicted by counsellors’ attitudes towards PNS\textsuperscript{70}. The principle of screening on the other hand might not be fully understood by some women or they are unable to assess the outcome of a risk result. A quality assurance study for the Dutch situation demonstrated high levels of informed decision-making and showed that the current information provision on PNS is of good quality. However women often do not have adequate knowledge on DS\textsuperscript{71}. Participation of PNS in the Netherlands has been shown to be lower in specified non-western ethnic minorities\textsuperscript{72}. These ethnic variations might be related to barriers in access to information and barriers in the decision-making process. Also differences in socio-economic status, education level and religion may play a role. An extra reason for women aged < 36 years to refrain from PNS might be the fact that the cost for the test are not reimbursed.

Recommendations

To accomplish a more effective first-trimester screening policy several adjustments of the current screening policies are necessary to improve the uptake of the FCT and improve screening performance. An important goal of PNS should be lowering the number of invasive diagnostic tests and thus the procedure-related miscarriages, constituting an important economic and health benefit. To establish a higher uptake of the FCT, barriers in access to screening and in the decision-making process should be removed. So first of all reimbursement of cost should apply to all women. Furthermore information provision should be related to educational level and adjusted for the different
General discussion

ethnic minorities. It is important that the information is individual oriented information and therefore actual screening performance of the FCT for the individual woman, based on her age, should be communicated to meet the requirements of informed decision-making. Moreover women should know that PND for AMA is an inappropriate screening tool for the detection of chromosomal abnormalities. The FCT should be the first step in prenatal testing because it provides a good risk assessment for women of all ages\textsuperscript{62} and thus PND for AMA should be abolished. Refraining from first-trimester screening, implicates that women do not receive information on DS, but also they will not receive information on other possible fetal structural abnormalities. With the implementation of the FCT, sonographers not only measure the NT, but are trained to check for structural defects.

Improving the information offer to pregnant women requires more education of the counsellors and adapting of the information leaves provided.

Also a timely information offer on PNS is important. In an empirical Dutch study it is proposed to incorporate an ‘information offer’ in a preconceptional care consultation\textsuperscript{73}. In the current screening program the information on PNS is provided in the first trimester of pregnancy and so the time for pregnant women to make a decision about participating in PNS is limited. Time constraints are a key obstacle to informed decision-making\textsuperscript{74,75}. With a preconceptional information offer, the likelihood of informed decision-making could be increased by lengthening time for contemplation and by unchaining the offer from other decisional steps with regard to PNS. This is even more relevant regarding the expected introduction of new PNS programs that can be performed as early as 8 weeks of gestation\textsuperscript{76-85}.

We emphasize the importance of ongoing monitoring of the procedure for providing information about PNS, in order to assess whether the objective of informed decision-making is fulfilled.

Improving screening performance is also relevant in establishing a more effective screening program. We advocate implementing the two-step screening approach with serum sampling before 11 weeks of gestation, despite possible practical issues. The effect of early serum sampling should be brought to the women’s attention. The effect of exclusion of maternal age risk from the risk assessment should be studied in another large screening population. However, it is expected that similar results will be obtained. Then introduction of this absolute risk method should seriously be considered, regarding the low FPR for women of all ages and the higher positive predictive value. Hence it would simplify pre- and post-test counselling. Currently maternal age risk still seems to be an important determinant in the choice for invasive diagnostic testing, although it is already included in the risk assessment. Finally it would be advisable to apply different cut-off values for women < 36 years of age and AMA women, for both the FCT as the absolute risk method.

Future perspectives

Recently enormous progress has been made in non-invasive prenatal testing (NIPT) for the detection of DS and in a lesser extent for trisomy 18 and trisomy 13 by analysis of cell-free fetal DNA in maternal plasma. This technique has now reached diagnostic accuracy with detection rates for DS
reported up to 99.5% at a FPR of less than 0.5% (specificity 99.7%)\textsuperscript{76-85}. Significantly less invasive diagnostic procedures will have to be performed as compared to the FCT. NIPT is now routinely available worldwide, including the USA, China and European countries such as Germany, Switzerland, England, Spain and Belgium. However, before introduction of NIPT in the Netherlands is allowed, a study on the validation of the test characteristics in a real-time daily-practice setting is deemed necessary and will be performed to establish the accuracy, the technical performance, the costs and the suitability for high-throughput analysis. Moreover, the study aims to assess pregnant women’s decision making with regard to NIPT, their attitudes and opinions on NIPT. As compared to the FCT, NIPT may be performed earlier in pregnancy (from week 9), subsequently allowing early decision making on a termination of pregnancy in case of an affected fetus. Also after 14 weeks of gestation NIPT can be performed in contrary to the FCT. NIPT in a low risk population and in women with multiple gestations is, however, not (yet) recommended because it has not been sufficiently validated in these groups. On theoretical grounds the positive predictive value after a positive NIPT result will be lower in low risk pregnancies. However, recent studies showed that in the general population of pregnant women with average risk, the test characteristics of NIPT are at least comparable to the characteristics in the group of pregnant women with an increased risk\textsuperscript{86,87}.

The final goal will be implementing NIPT as part of standard obstetric care, but it will take some time before this goal is reached. In the meantime the most effective screening policy would be to maintain the FCT as the first step in PNS. Women with an increased risk of having a baby with DS, trisomy 18 or trisomy 13 will be offered NIPT. Gradually the cut-off for increased risk can be expanded until NIPT will be available for all women.

With the implementation of NIPT, the FCT will disappear as a screening tool for chromosomal abnormalities. We do know that the NT measurement has an added value for the detection of a lot of structural abnormalities and genetic syndromes\textsuperscript{88-94}. In addition fetal anomaly screening is shifting more towards the first trimester of pregnancy\textsuperscript{95,96}. With the abolition of the FCT, the first-trimester screening with focus on anomalies and pregnancy outcome will probably be introduced.
References


42. Amor DJ, Xu JX, Halliday JL, Francis I, Healy DL, Breheny S, Baker HWG, Jaques AM. Pregnancies conceived using assisted reproductive technologies (ART) have low levels of pregnancy-associated plasma protein-A (PAPP-A) leading to a high rate of false-positive results in first-trimester screening for Down syndrome. *Hum Reprod* 2009;24:1330-1338.


General discussion


