FIRST TRIMESTER SCREENING FOR DOWN SYNDROME

MARKO NIEMIMAA

Department of Obstetrics and Gynaecology, University of Oulu

OULU 2003
Academic Dissertation to be presented with the assent of the Faculty of Medicine, University of Oulu, for public discussion in the Auditorium 4 of the University Hospital of Oulu, on June 27th, 2003, at 12 noon.
The aim of the present study was to evaluate the efficacy of the first trimester screening for Down syndrome (DS) in an unselected low-risk Finnish population. The study involved 4,617 women who attended screening between the 8th and 14th weeks of pregnancy in 1998-2000. They gave a blood sample for the measurement of pregnancy associated plasma protein A (PAPP-A) and free beta human chorionic gonadotrophin (β-hCG). Of these women, 3,178 also had an ultrasound examination for the measurement of fetal nuchal translucency (NT). The risk figure for every screened woman was calculated using a computerized risk figure program. The risk 1 in 250 was used as a cut-off. The subgroup of screen positives comprised 5.8% of the study group.

There were 16 DS cases. The combined method (maternal age, NT and the biochemical markers) detected 77% of the affected pregnancies. NT combined with maternal age gave a detection rate of 69%. Serum markers without NT combined with maternal age found 75% of the Down's.

In 49 consecutive singleton in-vitro-fertilization pregnancies, the β-hCG value was more often elevated compared to spontaneous pregnancies, increasing the false positive rate. In 67 twin pregnancies, the serum marker levels were approximately double those in singletons. Smoking reduced PAPP-A by 20% making the smokers more likely to get a positive screening result.

To determine the impact of the screening on the live born incidence of DS, two historical populations were compared. The first group was screened by second trimester serum samples (β-hCG and AFP) and the second group by first trimester ultrasound examination. When detection rates were at the same level, the second trimester screening reduced the number of live born Down's children more effectively.

In conclusion, the first trimester combined method (maternal age, NT, β-hCG and PAPP-A) for Down syndrome screening is efficient in an unselected low risk population. The biochemical screening is not recommended in IVF-pregnancies.

Keywords: beta-human chorionic gonadotropin, Down syndrome, fertilization in vitro, first trimester of pregnancy, nuchal translucency, pregnancy, pregnancy-associated alpha-plasma protein, prenatal diagnosis, smoking, twin pregnancy
“Aloille aatteheni mun
vei tuntemattomille,
 uus elo syttyi sieluhun,
aavistamaton sille;
kuin siivin lensi aikani,
oi, kuink’ on lyhyt kirjani!”

J.L.Runeberg: Vänrikki Stool

To my family
Acknowledgements

This study was carried out at the Department of Obstetrics and Gynaecology, University of Oulu, during the years 2000-2003.

I owe a great debt of gratitude to Professor Pentti Jouppila, M.D., Head of the Department of Obstetrics and Gynaecology, for his kind encouragement and for providing a good opportunity to carry out research in the clinic.

This book never would have been started or completed without the unwavering support of Professor Markku Ryynänen, M.D., my supervisor, co-writer and friend, who introduced me to the field of prenatal screening. He was the one to start to collect the data. I am deeply indebted to him for asking me to write my thesis on the subject. I want to express my most sincere gratitude to him for his expert guidance and warm support during this long-term project. I will always remember the moments of scientific and philosophical discussions we shared together.

I am most grateful to my second supervisor Professor Seppo Heinonen, M.D., for his valuable contribution to this thesis. His enthusiasm, diligence and efficiency have helped me many times to continue this work.

I warmly thank my co-workers Maija Seppälä, Ph.D., and Mikko Suonpää, Ph.D., for their invaluable contributions to the original publications in spite of the long geographical distance between us.

I am also deeply grateful to Professor Pertti Kirkinen, M.D., and Docent Anna Alanen, M.D., for their constructive criticism and kind advice during the preparation of the final manuscript.

I express my gratitude to Mrs. Heather Kannasmaa, for carefully revising the English language. I also want to thank warmly Risto Bloigu, Ph.D., and Pentti Nieminen, Ph.D., for their advice in statistics. I am thankful to Mrs. Liisa Kärki of the photographic laboratory for providing illustrations for articles. I am grateful to all the midwives around Finland who recruited subjects for this project. Warmest thanks go to all the women taking part in the study.

I want to thank all my friends and colleagues in Oulaskangas Hospital and Oulu University Hospital for their encouragement during this project.

I wish to thank my dear parents Elli and Toivo Niemimaa for their everlasting love and kindness. They have shown me by example how to live a good life.
Finally, my most loving thanks go to my wife Mervi and our wonderful children Noora, Nelli and Ville, for their love and patience especially during the last few months of this project.

This research has been supported by grants provided by the University of Oulu, the Finnish Gynaecological Association and the Alma and K.A. Snellman Foundation.
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>alphafetoprotein</td>
</tr>
<tr>
<td>β-hCG</td>
<td>free beta human chorionic gonadotropin</td>
</tr>
<tr>
<td>CHD</td>
<td>congenital heart disease</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CRL</td>
<td>crown rump length</td>
</tr>
<tr>
<td>DR</td>
<td>detection rate</td>
</tr>
<tr>
<td>DS</td>
<td>Down syndrome</td>
</tr>
<tr>
<td>DV</td>
<td>ductus venous</td>
</tr>
<tr>
<td>FMF</td>
<td>Fetal Medicine Foundation</td>
</tr>
<tr>
<td>FPR</td>
<td>false positive rate</td>
</tr>
<tr>
<td>HhCG</td>
<td>hyperglycosylated hCG</td>
</tr>
<tr>
<td>IVF</td>
<td>in vitro fertilization</td>
</tr>
<tr>
<td>MoM</td>
<td>multiple of median</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NT</td>
<td>nuchal translucency</td>
</tr>
<tr>
<td>NTD</td>
<td>neural tube defect</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>pregnancy associated plasma protein-A</td>
</tr>
<tr>
<td>PPV</td>
<td>positive predictive value</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>uE3</td>
<td>unconjugated estriol</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>mU/L</td>
<td>milliunits per litre</td>
</tr>
<tr>
<td>ng/mL</td>
<td>nanograms per millilitre</td>
</tr>
</tbody>
</table>
List of original publications

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:


Contents

Abstract
Acknowledgements
Abbreviations
List of original publications
Contents
1 Introduction ...................................................................................................................15
2 Review of the literature .................................................................................................17
  2.1 Screening for Down syndrome in the first trimester of pregnancy .................17
    2.1.1 Maternal age ....................................................................................................17
    2.1.2 Measurement of fetal nuchal translucency ......................................................17
    2.1.3 Other possible first trimester sonographic markers of Down syndrome......21
    2.1.4 Measurement of maternal serum samples ......................................................23
    2.1.5 Maternal serum and combined first trimester screening studies ..............25
    2.1.6 Maternal urine samples in pregnancies with Down syndrome ..............26
    2.1.7 Fetal DNA in maternal blood .................................................................27
    2.1.8 Integrating first and second trimester screening .......................................27
    2.1.9 Cost effectiveness of screening in the first trimester ......................................28
3 Purpose of the present study ...................................................................................30
4 Subjects and methods ..................................................................................................31
  4.1 Subjects ................................................................................................................31
  4.2 Methods .................................................................................................................32
5 Results .........................................................................................................................34
  5.1 Population medians and the performance of the screening tests .................34
  5.2 Serum concentrations in IVF pregnancies ..........................................................38
  5.3 Serum concentrations in twin pregnancies .........................................................39
  5.4 Impact of screening on the live born incidence of Down syndrome ..........40
  5.5 The influence of smoking on the first trimester screening parameters ............41
6 Discussion ...................................................................................................................43
  6.1 Efficacy of the screening tests .................................................................................43
  6.2 Serum concentrations in IVF-pregnancies .........................................................45
  6.3 Serum concentrations in twin pregnancies .........................................................46
6.4 Impact of fetal loss on the screening ................................................................. 47
6.5 Impact of maternal smoking on the first trimester screening parameters ........ 49
7 Conclusions ......................................................................................................... 51
References
1 Introduction

Down syndrome is the most common chromosomal disease among live born infants, with an incidence of 1 in 600. The syndrome is caused by trisomy of chromosome 21 in the vast majority of cases (95%), the rare reasons for this condition being unbalanced chromosome translocation and chromosomal mosaicism. Trisomy itself is a consequence of meiotic nondysjunction of chromosome alleles, and in 80% it is maternal in origin. DS is associated with mental handicap, cardiac and gastrointestinal anomalies, and vulnerability to infections and leukemia and later to Alzheimer-like dementia. (Simola 1998).

The syndrome was named after Dr. Langdon Down, a physician at London Hospital, who recognized in 1866 that the condition was congenital, dating from intrauterine life (Down 1866). Antenatal diagnosis became possible after the establishment of the method of cultivating fetal cells from amnion fluid for chromosome analysis (Steele & Breg 1966). DS prevalence was known to increase strongly with advancing maternal age, and amniocentesis was therefore offered to women over 35 years of age. This group made up 5% of the pregnant population at that time and 5% was thus established as a suitable screen positive ratio, so that the needs for further investigations could be met. Screening on the basis of maternal age suffered from low sensitivity, however; only about 30% of affected pregnancies were identified. DS screening then took its next step forward in 1984, when alphafetoprotein was found to be reduced in the maternal serum in Down’s pregnancies (Merkatz et al. 1984). Serum screening was now made available to all women regardless of age and it was implemented almost without extra cost, because AFP was already used to screen for neural tube defects (Wald et al. 1974). Human chorionic gonadotrophin was reported to be elevated in DS in 1987 (Bogart et al. 1987), and later unconjugated estriol was accepted as a third valid serum marker (Canick et al. 1988).

Midtrimester maternal serum screening was later established as an efficient method (Haddow et al. 1992). The detection rate is in the order of 60%, with a 5% false positive rate. The invasive diagnostic test, amniocentesis, carries a risk of about 1% of fetal loss as a result of a complication in the procedure. Serum sample collection occurs in the 15th week of pregnancy or later, because AFP is not effective for NTD-screening in early pregnancy (Sebire et al. 1997a). Women assessed as screen positive, i.e. at increased risk of having an affected fetus, have to wait a couple of weeks to get the final results after
amniocentesis. Those who receive bad news have to make a difficult decision about the fate of the fetus quite late, and in most cases after sharing the joy of the pregnancy with other people.

The development of ultrasound technology enabled the introduction of a new screening method in the early 1990s. Fetuses with DS were observed to have an increased nuchal translucency thickness more often than normal fetuses. Screening by certain serum markers was also shown to be effective in early pregnancy. The time had come to start to shift the screening from the second to the first trimester, which is exactly what women wish (de Graaf et al. 2002). Earlier screening brings certain benefits. Screening will possibly cause less anxiety when the results are obtained quickly, and the vast majority will be reassured by normal findings. Those women with positive screening results will get the chromosome analysis sooner after CVS than after AC. This allows them more time to consider the fate of the pregnancy. Termination in the first trimester is also safer than a procedure close to the 20th week. The termination of pregnancy is always problematic ethically, although Finnish society accepts termination as a legal medical procedure. Screening is based on free will and the women decide whether to participate or not.

In Finland, municipalities may choose independently which kind of screening they offer. Most commonly, screening is based on maternal serum samples (AFP, ß-hCG) in the 15th to 16th weeks of pregnancy. Many women also attend an ultrasound screening for structural defects in the 18th to 19th weeks. NT measurement is getting more common and it is being combined with simultaneous serum samples in many centers.

The screening is moving from the 2nd to the 1st trimester, which may even pose some problems for the Finnish maternity care system. Counseling has to be extensive, because screening is one of the most urgent matters to be discussed when a pregnant woman visits a maternity clinic for the first time. Counseling should be neutral and non-directive, and the fact that participation is voluntary needs to be stressed. The woman has to be prepared for the decision making process if she chooses to attend screening. This is the only way to achieve in reality the most important concept in the screening, i.e. the informed choice of a woman about her pregnancy. The aim is to provide people with appropriate information, so that a possible poor pregnancy outcome will not come as a surprise.
2 Review of the literature

2.1 Screening for Down syndrome in the first trimester of pregnancy

2.1.1 Maternal age

Down syndrome prevalence increases strongly with advancing maternal age. For example, the at-term risk for a 20-year-old woman is 1 in 1500, but for a 40-year-old it is 1 in 100. The risk is even higher at the time of screening in the 12th week of pregnancy, because about 30% of affected pregnancies will abort before term (Snijders et al. 1999). Thus, in many communities in Finland, CVS or AC is offered to women older than 37-38 years.

2.1.2 Measurement of fetal nuchal translucency

In the 1980s, a thickened nuchal fold of the fetus in the second trimester was found to be associated with chromosomal disorders (Benacerraf et al. 1985). The respective finding in the first trimester was reported by Szabo (Szabo & Gellen 1990). The term “nuchal fold” in the second trimester was replaced by “nuchal translucency” in the first trimester and it denotes a sonolucent region in the posterior aspect of the fetal neck.

In the early 1990s, several studies in high-risk pregnancies demonstrated a possible association between increased NT and chromosomal defects in the first trimester (Nicolaides et al. 1992, Johnson et al. 1993, Nadel et al. 1993, Savoldelli et al. 1993, Pandya et al. 1995b). Subsequently, a series of observational studies in high-risk pregnancies were carried out: these involved measurement of NT immediately before fetal karyotyping, mainly for advanced maternal age. The studies reported different DRs of aneuploidy (30-80%) (Brambati et al. 1995, Pandya et al. 1995a).
The next step was to study the implementation of NT screening in routine practice. The studies are summarized in table 1.

**Table 1. Studies examining the implementation of fetal NT screening. Modified from (Nicolaides et al. 1999).**

<table>
<thead>
<tr>
<th>Author</th>
<th>Gestation (weeks)</th>
<th>n</th>
<th>Successful measurement (%)</th>
<th>NT cut-off (≥ mm)</th>
<th>FPR (%)</th>
<th>DR trisomy 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Pandya et al. 1995a)</td>
<td>10-14</td>
<td>1,763</td>
<td>100</td>
<td>2.5</td>
<td>3.6</td>
<td>3/4 (75%)</td>
</tr>
<tr>
<td>(Szabó et al. 1995)</td>
<td>9-12</td>
<td>3,380</td>
<td>100</td>
<td>3.0</td>
<td>1.6</td>
<td>28/31 (90%)</td>
</tr>
<tr>
<td>(Bewley et al. 1995)</td>
<td>8-13</td>
<td>1,704</td>
<td>66</td>
<td>3.0</td>
<td>6.0</td>
<td>1/3 (33%)</td>
</tr>
<tr>
<td>(Kornman et al. 1996)</td>
<td>8-13</td>
<td>923</td>
<td>58</td>
<td>3.0</td>
<td>6.3</td>
<td>2/4 (50%)</td>
</tr>
<tr>
<td>(Zimmermann et al. 1996)</td>
<td>10-13</td>
<td>1,131</td>
<td>100</td>
<td>3.0</td>
<td>1.9</td>
<td>2/3 (67%)</td>
</tr>
<tr>
<td>(Taipale et al. 1997)</td>
<td>10-16</td>
<td>10,010</td>
<td>99</td>
<td>3.0</td>
<td>0.8</td>
<td>7/13 (54%)</td>
</tr>
<tr>
<td>(Hafner et al. 1998)</td>
<td>10-14</td>
<td>4,371</td>
<td>100</td>
<td>2.5</td>
<td>1.7</td>
<td>4/7 (57%)</td>
</tr>
<tr>
<td>(Pajkrt et al. 1998)</td>
<td>10-14</td>
<td>1,547</td>
<td>96</td>
<td>3.0</td>
<td>2.2</td>
<td>6/9 (67%)</td>
</tr>
</tbody>
</table>

The prevalence of fetal trisomy at 9-14 weeks of gestation in different maternal age groups was reported in 1999 (Snijders et al. 1999). NT thickness over 3 mm seems to increase the maternal age related risk of abnormal karyotype up to 25-30-fold (Nicolaides et al. 1992, Nicolaides et al. 1994, Pandya et al. 1995b). The observed number of trisomies with NT under 3 mm is approximately one fifth of the number expected on the basis of maternal age (Nicolaides et al. 1994).

In the early studies a fixed cut off for NT was used; later, however, it was learned that NT increases with CRL (Braithwaite et al. 1996). Thus it is essential to take gestational age into account when determining whether a given NT is increased. The fixed millimeter cut-off can be replaced by a certain NT percentile, for example the 95th or the 99th. Furthermore, the use of MoMs of the unaffected population and the distributions of the NT in normal and affected populations, allows a combined risk estimate by maternal age and NT measurement. Table 2 summarizes the studies by a combination of maternal age and nuchal translucency.
Table 2. First trimester studies on screening by a combination of maternal age and NT.

<table>
<thead>
<tr>
<th>Author</th>
<th>Screened women (n)</th>
<th>Median maternal age</th>
<th>Down syndrome (n)</th>
<th>FPR (%)</th>
<th>DR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Snijders et al. 1998)</td>
<td>96,127</td>
<td>31</td>
<td>326</td>
<td>8</td>
<td>82</td>
</tr>
<tr>
<td>(Theodoropoulos et al. 1998)</td>
<td>3,550</td>
<td>29</td>
<td>11</td>
<td>5</td>
<td>91</td>
</tr>
<tr>
<td>(Biagiotti et al. 1997)</td>
<td>3,241</td>
<td>38</td>
<td>32</td>
<td>5</td>
<td>59</td>
</tr>
<tr>
<td>(Thilaganathan et al. 1999)</td>
<td>11,398</td>
<td>29</td>
<td>21</td>
<td>5</td>
<td>76</td>
</tr>
<tr>
<td>(Schwärzler et al. 1999)</td>
<td>4,523</td>
<td>29</td>
<td>12</td>
<td>5</td>
<td>77</td>
</tr>
<tr>
<td>(Zoppi et al. 2001)</td>
<td>12,495</td>
<td>33</td>
<td>64</td>
<td>9</td>
<td>90</td>
</tr>
<tr>
<td>(Brizot et al. 2001)</td>
<td>2,996</td>
<td>28</td>
<td>10</td>
<td>7</td>
<td>90</td>
</tr>
<tr>
<td>(Gasiorek-Wiens et al. 2001)</td>
<td>21,959</td>
<td>33</td>
<td>210</td>
<td>13</td>
<td>88</td>
</tr>
</tbody>
</table>

With a 5% false positive rate, the largest study showed a 77% detection rate for DS in the UK. The risk of trisomy 21 was calculated from the maternal age and gestational-age related prevalence, multiplied by a likelihood ratio depending on the deviation from normal in NT thickness for crown-rump length. (Snijders et al. 1998).

However, the UK multicenter study was criticized for being an interventional study, where the decision regarding diagnostic procedure (CVS) was based on the results of the test (NT) under study. This kind of study arrangement increases the likelihood of screen positive affected fetuses being included and screen negative affected fetuses being excluded, thus causing a verification bias. A proportion of the undetected screen negative trisomic fetuses abort and are lost from the follow-up, causing an overestimation of the detection rate. Therefore, the real DR of Down syndrome in the UK study is more likely to be in the order of 60% than 77% (Haddow 1998). Another source of criticism was the lack of information on the number of patients who were screened but whose NT measurement proved inadequate as a result of technical ultrasonographic difficulties (Malone et al. 2000).

NT measurement seems to help certain groups of pregnant women who previously could not be effectively served by serum screening. The DR of nuchal translucency is not affected by assisted conception (Liao et al. 2001) or twin pregnancy (Sebire et al. 1996). Increased and discordant NT in monochorionic twin pregnancies may also indicate an increased risk of subsequent development of twin-to-twin transfusion syndrome (Sebire et al. 2000).

The reasons for increased NT both in aneuploid and euploid fetuses are various. First, cardiac anomalies, mostly ventricular or atrioventricular septal defects, coincide with DS in about 50% of cases (Hyett et al. 1997, Paladini et al. 2000). Narrowing of the aortic isthmus occurs frequently, and there is an association between the degree of narrowing and translucency thickness. Narrowing of the isthmus can result in overperfusion of the tissues of the head and neck, leading to subcutaneous edema. With advancing gestation the diameter of the isthmus increases and the hemodynamic consequences of the narrowing of the isthmus may be overcome. Thus, relative narrowing of the aortic isthmus may partly explain the spontaneous resolution of NT later during pregnancy. (Hyett et al. 1997). However, no specific type of CHD is associated with increased NT, and heart failure seems not to explain the association between congenital heart defects.
and increased NT (Simpson & Sharland 2000). It has to be noted that most fetuses with NT have no cardiac defects; thus abnormal or delayed fetal heart development might explain nuchal edema in many cases.

A second theory suggests that fluid collects in the neck due to the impaired or delayed development of lymphatic drainage. In a study in which the fetuses presented increased NT at 10-14 weeks' gestation, seven fetuses were terminated because of an abnormal karyotype. The pathological specimens, compared by morphometric analysis with normal fetuses of the same gestational age, showed edema and dilatation of lymphatic capillary vessels. No particular relationship was found with any structural abnormality. (Greco et al. 1996).

One further reason for pathologic NT in fetuses with diaphragmatic hernia might be venous congestion in the head and neck due to mediastinal compression (Sebire et al. 1997b). Finally, alteration in the extra cellular matrix of fetal skin due to over-expression of certain collagen genes in trisomic fetuses has been suggested as a cause of increased NT (von Kaisenberg et al. 1998).

Furthermore, NT is associated with cardiac defects also when the fetus is chromosomally normal. A large study showed that the prevalence of major defects of the heart and great arteries increases with increasing NT. The prevalence was almost 20% where NT was over 5.4 mm. In that study, the use of the 99th percentile of NT as a screening cut-off would have detected 40% of cardiac anomalies. (Hyett et al. 1999). However, other studies have reported lower detection rates: 11% (Mavrides et al. 2001a), 11% (Schwärzler et al. 1999) and 27% (Michailidis & Economides 2001). Nevertheless, increased NT is definitely an indication for a detailed ultrasound examination by a specialist due to a high prevalence of cardiac defects (Brady et al. 1998, Hyett et al. 1999, Zosmer et al. 1999, Ghi et al. 2001, Orvos & Wayda 2002), although the low sensitivity of NT for major CHD in the general population indicates that NT cannot be relied on as the sole or major screening tool for this condition (Mavrides et al. 2001a, Bilardo et al. 2001b). In Finland, among children aged 2-7 years with a previous measurement of NT over 3 mm during pregnancy, a CHD prevalence of 12 % was reported (Hiippala et al. 2001).

Besides fetal aneuploidies and heart defects, increased NT is also associated with a wide range of other defects, such as diaphragmatic hernia (Sebire et al. 1997b), exomphalos, body stalk anomaly, fetal akinesia deformation sequence and possibly with rare skeletal dysplasias and genetic syndromes (Souka et al. 1998).

Increased NT in chromosomally normal fetuses predicts an adverse pregnancy outcome. The risk of spontaneous abortion or intrauterine death was 5.2%, and for neonatal and infant deaths the risk was 1.4%. Of the survivors, 5.6% had abnormalities requiring medical or surgical treatment or leading to mental handicap. Again, the prognosis got worse with increasing NT. The chance of a live birth with no defects was only 31% if NT exceeded 6.4 mm. (Souka et al. 2001). There are few long-term follow-up studies on healthy children who presented with increased NT in the first trimester of pregnancy. In general the parents can be reassured that, in the majority of cases, postnatal developmental is normal. (Maymon et al. 2000, Hiippala et al. 2001). The prevalence of neurodevelopmental delay was reported at 10% (Adekunle et al. 1999) and 5.6% (Van Vugt et al. 1998). A recent study showed that among the normal neonates, 11% were
considered to have a significant neurological handicap or orthopedic problems at 12 to 72 months of age (Senat et al. 2002).

Besides NT screening, the first trimester ultrasound examination has several other benefits. It is non-invasive, it can be done early and the results are obtained quickly. It helps to accurately define gestational age, the number of fetuses and their viability and structure, chorionicity in the case of a twin pregnancy, and the location of the placenta.

2.1.3 Other possible first trimester sonographic markers of Down syndrome

The NT is at the present time the most established screening test which can be measured during an ultrasound examination in the first trimester. In addition to NT, other ultrasonic markers for trisomy 21 have been investigated. Studies of fetal heart pattern screening have yielded conflicting results; ductus venosus flow measurement is technically demanding; and there have so far been only preliminary reports on some others like diameter of umbilical cord and placental volume. The absence of nasal bone seems a promising marker, which could be connected to NT measurement. However, some words of caution are necessary here: “It is of paramount importance that new prenatal tests are scrutinized and their efficacy is assessed before they are introduced into clinical practice, in order to avoid too early and over enthusiastic application of a test, which may cause more harm than benefit. There are two important considerations: first, what are the variability and reproducibility of a test, and, second, has its performance been tested in a low-risk or a high-risk population?” (Hecher 2001).

Fetal heart rate screening needs more evaluation before it can be used clinically. As regards trisomy 21, some studies have reported tachycardia among affected fetuses (Jauniaux et al. 1996, Hyett et al. 1996a, Liao et al. 2000) while others have reported bradycardia (Martinez et al. 1996) or normal heart rate (Van Lith et al. 1992). The sensitivity of fetal heart pattern to detect Down syndrome has been 10% (Liao et al. 2000) and 21% (Hyett et al. 1996a). Thus, inclusion of fetal heart rate in a first trimester screening program for trisomy 21 by a combination of maternal age and NT is unlikely to provide improvement in sensitivity (Liao et al. 2000).

The ductus venosus is a shunt vein delivering well-oxygenated blood from the umbilical vein directly to the fetal heart. In the first trimester the entire length of the DV measures 2-3 mm only. Increased pulsatility index and absent or reversed flow during atrial contraction are the possible pathologic findings. A few studies have reported an association between abnormal DV flow and aneuploidy in high-risk pregnancies, often with increased NT measurement. These studies are summarized in table 3.
Table 3. Ductus venosus Doppler studies to detect fetal aneuploidy among high-risk women.

<table>
<thead>
<tr>
<th>Study</th>
<th>Patients (n)</th>
<th>Successful measurement (%)</th>
<th>DR</th>
<th>FPR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Matias et al. 1998)</td>
<td>486</td>
<td>57/63 (90%)</td>
<td>57/63 (90%)</td>
<td>3.1</td>
</tr>
<tr>
<td>(Borrell et al. 1998)</td>
<td>414</td>
<td>82</td>
<td>8/11 (73%)*</td>
<td>5</td>
</tr>
<tr>
<td>(Bilardo et al. 2001a)</td>
<td>186</td>
<td>86</td>
<td>30/46 (65%)</td>
<td>21</td>
</tr>
<tr>
<td>(Antolin et al. 2001)</td>
<td>1,371</td>
<td>13/20 (65%)</td>
<td>13/20 (65%)</td>
<td>4.3</td>
</tr>
<tr>
<td>(Matias &amp; Montenegro 2001)</td>
<td>515</td>
<td>55/69 (80%)</td>
<td>55/69 (80%)</td>
<td>1</td>
</tr>
<tr>
<td>(Zoppi 2002)</td>
<td>156</td>
<td>97</td>
<td>23/33 (70%)</td>
<td></td>
</tr>
<tr>
<td>(Mavrides et al. 2002)</td>
<td>256</td>
<td>98.5</td>
<td>27/46 (59%)</td>
<td>4.5</td>
</tr>
</tbody>
</table>

*Down syndrome only.

Some authors have suggested that the evaluation of ductal flow between the 11th and 14th weeks of gestation should be adopted as a second level screening test to reduce the invasive test rate derived from the measurement of nuchal translucency (Matias et al. 1998, Matias & Montenegro 2001, Antolin et al. 2001). However, others have stated that as the DV flow pattern is correlated with NT measurement it cannot be used as an independent variable to reduce the indication of fetal karyotyping (Bilardo et al. 2001a). Instead, in the group of increased NT and normal chromosomes, abnormal DV flow predicts major cardiac defects (Matias et al. 1999) and adverse pregnancy outcome (Bilardo et al. 2001a). Therefore, the real clinical value of DV Doppler seems to be the possibility to evaluate the further prognosis in terms of pregnancy outcome if NT is increased and the karyotype is normal (Hecher 2001). Moreover, the reproducibility (Prefumo et al. 2001) and variability studies (Mavrides et al. 2001b) have to be confirmed with a higher prevalence of abnormal waveforms.

New first trimester markers are being sought as the resolution of ultrasound technology improves. Single reports about thickened umbilical cord (Ghezzi et al. 2002) and diminished placental volume (Metzenbauer et al. 2002) with aneuploidy were published recently. In an observational study, the nasal bone in fetuses with DS at 11-14 weeks was absent in 43 of 59 (73%), but in only three of 603 (0.5%) chromosomally normal fetuses. The likelihood ratio for trisomy 21 was 146 for absent and 0.27 for present nasal bone. In screening for trisomy 21, a combination of maternal age, NT, and fetal profile for the presence or absence of nasal bone might achieve a DR of 85% and decrease the FPR to about 1% (Cicero et al. 2001). A retrospective case-control study comprising 100 trisomy 21 pregnancies estimated a 97% DR for the combination of maternal age, NT, PAPP-A, ß-hCH and absent fetal nasal bone. For a FPR of 0.5%, the detection rate was 90% (Cicero et al. 2003).
2.1.4 Measurement of maternal serum samples

2.1.4.1 Pregnancy-associated plasma protein A

Pregnancy-associated plasma protein A is a large glycoprotein (MW 720,000 daltons). PAPP-A was first described in 1974. (Lin et al. 1974). Its biological function is mostly unknown, although recent research has demonstrated granulosa cells as a source of PAPP-A in ovaries and suggested that PAPP-A is a marker of ovarian follicle selection and corpus luteum formation (Conover et al. 2001). During pregnancy PAPP-A levels rise all the way to term (Bischof et al. 1982). It has been shown that PAPP-A is released into the medium by cultured early and late pregnancy deciduas as well as by endometrial stromal cells (Bischof 1984, Bischof & Tseng 1986). Reports in the early 1990s suggested that PAPP-A is reduced in pregnancies with trisomic fetuses. The deviation from normality decreases with advancing gestation (Brambati et al. 1993, Bersinger et al. 1994). The latter finding is compatible with reports that in the second trimester there is no significant difference in maternal serum PAPP-A between pregnancies with trisomy 21 and controls (Aitken et al. 1994). A meta-analysis stated that median maternal serum PAPP-A level in DS pregnancies is 0.35 MoM, 0.40 MoM and 0.62 MoM at gestational weeks 6-8, 9-11 and 12-14 respectively, and 0.94 MoM thereafter. The estimated Down syndrome detection rate for a 5% FPR was 52% for PAPP-A alone. (Cuckle & Van Lith 1999).

Brizot et al studied the possible causes for the decrease of PAPP-A in trisomic pregnancies. They investigated the relationship between placental messenger-RNA expression and the concentration of PAPP-A in both placental tissue and maternal serum in normal and trisomic pregnancies. The maternal serum concentration of PAPP-A in the trisomic group of pregnancies was significantly lower than in the normal controls. However there were no significant differences in PAPP-A mRNA expression or PAPP-A protein concentration in the placental tissues. There was no significant association between the level of placental mRNA and maternal serum PAPP-A concentrations in the normal or trisomic pregnancies. These findings suggest that the decrease in maternal serum PAPP-A in trisomic pregnancies is due to alterations in post-translational events such as protein stability, alterations in the release mechanism of the protein, impaired protein transport across the placenta or modified serum stability of PAPP-A. (Brizot et al. 1996).

Another attempt to explain decreased PAPP-A values in trisomic pregnancies is called the placental compensation hypothesis. This theory suggests that, very early in a DS pregnancy, the maternal serum concentration of all fetoplacental markers might be reduced. With advancing gestation, the placental, but not fetal, markers gradually increase and finally reach or even exceed the normal range. The theory of passive release of placental proteins into the maternal circulation suggests that the speed of this process is inversely related to the molecular weight of the marker protein. Thus, the crossing point between the normal and the affected pregnancy curves would be early for low MW placental proteins (such as \(\beta\)-hCG), and later for high MW markers (such as PAPP-A). (Bersinger et al. 1995).
2.1.4.2 Human chorionic gonadotrophin

Human chorionic gonadotrophin is a glycoprotein with a small MW (39,500 daltons). It was purified from the urine of pregnant women in 1927 (Ascheim & Zondek 1927). Human chorionic gonadotrophin is produced by the placenta, reaching its peak value in the maternal circulation at 8 to 10 weeks of pregnancy. Thereafter, hCG levels decrease rapidly, continuing to fall up to 20 weeks of pregnancy, when a plateau is reached (Braunstein et al. 1976).

Human chorionic gonadotrophin and its beta subunit are well-established markers for DS in the second trimester. Thus, it was natural to investigate their effectiveness in the first trimester. Data from numerous studies concluded that the use of the intact molecule of hCG in DS screening is not productive during the first trimester (Cuckle et al. 1988, Macintosh et al. 1994). Fortunately, early studies suggested that $\beta$-hCG is markedly elevated in DS pregnancies (Aitken et al. 1993, Macintosh et al. 1994). This finding was confirmed by others (Krantz et al. 1996, Forest et al. 1997). The latest meta-analysis gathered data from 17 series and 579 DS cases expressing a mean MoM value of 1.98 for $\beta$-hCG. Statistical modeling gave a detection rate of 42% for a 5% FPR for $\beta$-hCG alone. (Cuckle & Van Lith 1999).

2.1.4.3 Other studied serum markers

Alpha-fetoprotein and unconjugated estriol are well-established markers of DS in the second trimester of pregnancy. AFP is also lowered in DS in the first trimester. A meta-analysis of 542 cases gave a mean MoM of 0.79 for AFP (Cuckle & Van Lith 1999). However, it has been shown that the standard deviation for AFP is increased by 20% compared to that reported in the second trimester. Therefore its contribution to the detection rate at a given false-positive rate will be lowered by the increase in overlap of the affected and unaffected distributions (Berry et al. 1995). Modeling suggests that adding AFP to the combination of PAPP-A and $\beta$-hCG increases the DR only by 2.0% (from 64.6% to 66.6%) (Cuckle & Van Lith 1999).

Unconjugated estriol is also lowered in DS in the first trimester. A meta-analysis of 226 DS cases gave a mean MoM of 0.74. However, adding this marker to the combination of PAPP-A and $\beta$-hCG might increase the detection rate by only 4% (from 64.6% to 68.6%). (Cuckle & Van Lith 1999).

Inhibin-A is reported to be increased in DS in the second trimester (Aitken et al. 1996, Cuckle et al. 1996) but its benefits as an additional serum marker are not accepted by all (Reynolds 2000). However, a recent report showed that quadruple screening yields a DR of 70% and performs better than triple or double screening (Wald et al. 2003). In the first trimester, the results are controversial. Some have reported a difference between affected and unaffected pregnancies (Wallace et al. 1995, Noble et al. 1997b) while others have not (Spencer et al. 2001). Inhibin-A seems to correlate strongly with $\beta$-hCG. Thus the sensitivity for trisomy 21 achieved through the combination of maternal serum inhibin-A...
and ß-hCG is not significantly different from that achieved with ß-hCG alone. (Noble et al. 1997b).

### 2.1.5 Maternal serum and combined first trimester screening studies

From among the many studied serum markers, the combination of PAPP-A and ß-hCG has been found the most useful for screening in the first trimester. Results of the studies are summarized in table 4.

**Table 4. Estimated DRs for trisomy 21 by a combination of maternal age, PAPP-A and ß-hCG with a 5% FPR**

<table>
<thead>
<tr>
<th>Study</th>
<th>Down cases (n)</th>
<th>DR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Wald et al. 1996)</td>
<td>77</td>
<td>63</td>
</tr>
<tr>
<td>(Krantz et al. 1996)</td>
<td>22</td>
<td>63</td>
</tr>
<tr>
<td>(Berry et al. 1997)</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>(Forest et al. 1997)</td>
<td>18</td>
<td>56</td>
</tr>
<tr>
<td>(Haddow et al. 1998)</td>
<td>48</td>
<td>60</td>
</tr>
</tbody>
</table>

Because the serum markers and NT do not correlate with each other either in chromosomally normal or abnormal fetuses (Brizot et al. 1994, Brizot et al. 1995), combining their information gives better results than using either of them alone. The concept of the combined screening is based on NT being the strongest marker. The serum samples are added because they can possibly increase the detection rate for a given false positive rate. The risk calculation software stresses the importance of maternal age for the background risk. To calculate the likelihood ratios, the software program uses the Gaussian distributions of NT and serum values of normal and affected cases. These are described by their means of log10 MoMs, standard deviations and correlation coefficients between the markers. The screening test performs well if the Gaussian distributions of the markers in the normal and affected population are separated. Conversely, the screening test is inefficient if the distributions overlap widely. The degree of the overlap is influenced by the median MoMs and SDs in the populations.

With the advent of rapid immunoassays, it has become possible to provide pretest counseling, biochemical testing of the mother, ultrasound examination of the fetus and post-test counseling on a combined risk estimate, all within a one-hour visit to one-stop clinic for assessment of risk (OSCAR) for fetal anomalies (Bindra et al. 2002). However, this combined method is claimed to have been inadequately studied (Reynolds 2000, Malone et al. 2000). The results of retrospective combined studies are given in table 5. Prospective studies are presented in table 15 in the discussion section.
Table 5. Retrospective combined studies on trisomy 21 at a 5% FPR rate in the first trimester.

<table>
<thead>
<tr>
<th>Study</th>
<th>Down’s cases (n)</th>
<th>Detection rate</th>
<th>Serum screening (%)</th>
<th>Nuchal translucency (%)</th>
<th>Combined screening (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Wald &amp; Hackshaw 1997)</td>
<td>86 + 77*</td>
<td></td>
<td>62</td>
<td>63</td>
<td>80</td>
</tr>
<tr>
<td>(Biagiotti et al. 1998)</td>
<td>32</td>
<td></td>
<td>59</td>
<td>68</td>
<td>76</td>
</tr>
<tr>
<td>(de Graaf et al. 1999)</td>
<td>37</td>
<td></td>
<td>55</td>
<td>68</td>
<td>85</td>
</tr>
<tr>
<td>(Spencer et al. 1999)</td>
<td>210</td>
<td></td>
<td>67</td>
<td>73</td>
<td>89</td>
</tr>
</tbody>
</table>

*Two different datasets on DS cases.

2.1.6 Maternal urine samples in pregnancies with Down syndrome

Initial studies in the second trimester indicated that the maternal urine beta-core fragment of hCG was an outstanding marker, detecting over 80% of DS cases (Cuckle et al. 1994). Since these reports, widely varying results have been published, indicating between 20% and 66% detection of cases at a 5% FPR. This poor screening performance is possibly due to aggregation of the beta-core fragment molecules upon storage in the freezer (Cole et al. 1999a). Another possible urinary marker, hyperglycosylated hCG is a form of hCG with more complex oligosaccharide side chains. In the second trimester, in studies among high risk patients, HhCG has given promising results (Cole et al. 1998, Cuckle et al. 1999, Cole et al. 1999b, Bahado-Singh et al. 2000). Furthermore, HhCG might not suffer from a stability problem like beta-core fragment (Cole et al. 1999b). Other studied urine markers are β-hCG and total estriol. However, due to a lack of prospective studies among low-risk population, urinary screening has not been established in clinical practice.

Studies of urinary markers in the first trimester are sparse. In a study of 8 cases of DS, the median MoM of HhCG of the affected pregnancies was 3.6 MoM (Weinans et al. 2000). Another report of 23 DS cases indicated an 80% detection rate for HhCG (Cole et al. 1999b). In contradiction to this, a study of 5 affected cases suggested that beta-core fragment of hCG is not a promising marker for DS screening (Kornman et al. 1997). Finally, a study of 22 DS cases concluded that any of the studied urine markers (β-hCG, beta core hCG, total estriol) is unlikely to be of value in first trimester screening, because their additional detection rate to the NT measurement is marginal (Spencer et al. 1997).
2.1.7 Fetal DNA in maternal blood

The recent discovery of high concentrations of fetal DNA in maternal plasma represents a promising noninvasive approach. Compared with the analysis of the cellular fraction of maternal blood, the analysis of fetal DNA extracted from maternal plasma has the advantage of being rapid, robust, and easy to perform. The fetal DNA detected is limited to the current pregnancy. Fetal DNA has been found to be increased in maternal blood when the fetus has trisomy 21, possibly due to accelerated apoptosis of fetal cells, although there is a considerable degree of overlap with euploid fetuses. The relatively low sensitivity and specificity implies that a combination of circulating fetal DNA with other markers for fetal trisomy 21 is needed before the measurement of fetal DNA is useful as a screening test for Down syndrome. In addition, DNA markers that would identify female fetuses with DS are needed, as the current basis of detection uses gene sequences from the Y chromosome. (Pertl & Bianchi 2001).

2.1.8 Integrating first and second trimester screening

The concept of integrated screening, in which the same patient undergoes both first and second trimester screening, was introduced in 1999. The integrated method, including maternal age, NT measurement and PAPP-A in the first and AFP, hCG, uE3 and inhibin A in the second trimester, yielded very good results; detection rates of 94% and 85% with FPRs of 5% and 1% respectively. (Wald et al. 1999a). Another study also using statistical modeling gave very similar figures (Cuckle 2001). A recent non-interventional study gave a DR of 86% at a FPR of 5% for a combination of NT, hCG, AFP and maternal age. This was better than NT or serum samples alone but the confidence intervals of detection rates overlapped due to the small sample size (35 DS cases). (Lam et al. 2002). Two retrospective integrated studies with 21 DS cases (Rozenberg et al. 2002) and 12 DS cases (Audibert et al. 2001) gave DRs of 80% and 90% respectively, with a 5% FPR.

The integrated test is best practiced with a non-disclosure approach, i.e. the patient is not given partial results after first trimester parameters have been measured. This avoids confusing patients who otherwise might get different risk estimates. Reporting partial results is not necessary even when there is a high risk, because only a very small number of women do have such a high risk after the first trimester tests which cannot be reduced and normalized by the integrated test (Hackshaw & Wald 2001a). The other method, the disclosure approach or sequential screening, avoids long waiting times because action can be taken on intermediate results. However, this approach is associated with a higher FPR than non-disclosure screening with the same combination. (Cuckle 2001, Herman et al. 2002a). Furthermore, the positive predictive value of the second trimester tests will be very low because most of the affected fetuses will already have been detected (Kadir & Economides 1997, Thilaganathan et al. 1997, Michailidis et al. 2001). If integrated screening with a single risk estimate is not established, the woman should be advised not to participate in both types of screening. This is because her age-specific risk has changed after she was screened in the first trimester, and the following second trimester test will
give an erroneous risk estimate. (Hackshaw & Wald 2001b). Although the integrated screening has a high detection rate this may be outweighed by the delay in diagnosis and the extra visits and cost, so the right time for screening is most likely to be in the first trimester (Michailidis et al. 2001).

As a summary, the detection rates of the most used or promising screening methods are shown in table 6.

Table 6. Estimations of the efficiency of the main DS screening methods at 5% FPR.

<table>
<thead>
<tr>
<th>Screening method</th>
<th>Detection of trisomy 21 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>30</td>
</tr>
<tr>
<td>2nd trimester double (AFP, ß-hCG)</td>
<td>60-65</td>
</tr>
<tr>
<td>2nd trimester triple (AFP, ß-hCG, uE3)</td>
<td>65-70</td>
</tr>
<tr>
<td>2nd trimester quadruple (AFP, ß-hCG, uE3, inhibin A)</td>
<td>70-75</td>
</tr>
<tr>
<td>1st trimester serum (PAPP-A, ß-hCG)</td>
<td>60-65</td>
</tr>
<tr>
<td>1st trimester NT</td>
<td>70-75</td>
</tr>
<tr>
<td>1st trimester combined (NT, PAPP-A, ß-hCG)</td>
<td>80-85</td>
</tr>
<tr>
<td>1st and 2nd trimesters integrated (NT, PAPP-A, AFP, ß-hCG, uE3, inhibin-A)</td>
<td>85-90</td>
</tr>
</tbody>
</table>

The future of the first trimester screening and the integrated screening depends on two major prospective trials. The SURUSS (Serum Urine and Ultrasound Screening Study) trial started in June 1996 in the UK. The FASTER (First And Second Trimester Evaluation of Risk for aneuploidy) started two years ago in the USA. These trials will allow a reliable comparison of DS detection rates between first and second trimester.

2.1.9 Cost effectiveness of screening in the first trimester

The multiple serum marker screening in the second trimester has been shown to be safer and more cost effective than screening based on maternal age alone (Sheldon & Simpson 1991, Shackley et al. 1993). However, both studies have been criticized for basing their conclusions on estimates of average rather than incremental cost effectiveness, and thus failing to inform the decision makers of the budgetary expansion required to introduce a new screening method (Petrou et al. 2000).

There are few attempts to compare economic aspects of first and second trimester screening. One study estimated the costs and savings if NT screening was introduced in the USA. The conclusion was that the present American approach (second trimester screening by maternal age and serum samples) was more cost effective. However, as NT screening decreases the number of invasive genetic procedures by almost half and thus decreases the number of fetal losses, introduction of NT screening should be seriously considered. (Vintzileos et al. 1998). Recently, an American comparison favoured first trimester screening. NT measurement alone yielded a $98,381 incremental cost per each additional identified DS fetus compared to the second trimester serum screening. The corresponding incremental costs for the first trimester serum sampling and the first
trimester combined method were $160,266 and $319,934, respectively. All these values were less than the $577,248 that each DS case was estimated to cost. (Caughey et al. 2002).

Another study compared the effects, safety, and cost effectiveness of different DS screening strategies. The NT measurement, quadruple test, first trimester combined, and integrated test represented the best options. All other strategies including screening based on maternal age, the second trimester double test, and the first trimester serum test were less effective, cost more per additional prevented birth of an affected infant, and were less safe. (Gilbert et al. 2001). Finally, contingent testing was suggested as a cost-effective alternative. In this method, biochemical samples were taken first and NT was provided only for those with an intermediate risk, in order to reduce the costs. (Christiansen & Larsen 2002).
3 Purpose of the present study

The principle aim of the present investigation was to evaluate the efficacy of the combined first trimester screening for Down syndrome in a Finnish unselected low-risk population. The following issues were of particular interest:

1. Efficacy of the different first trimester screening methods.
2. Distribution of the maternal serum biochemical markers in IVF-pregnancy
3. Distribution of the maternal serum biochemical markers in twin pregnancy
5. The influence of maternal smoking on the distribution of the screening parameters.
4 Subjects and methods

4.1 Subjects

During the years 1998 – 2000 blood samples were drawn in primary care centers and in the maternity clinics of the participating university hospitals of Oulu and Kuopio and in ten smaller hospitals. Gestational ages ranged from 7 weeks 2 days to 13 weeks 6 days and were based on ultrasound examinations in 80% of cases. The serum markers were measured prospectively in 4,982 pregnancies. The patient information was complete in 4,617 cases. Nuchal translucency was measured in 3,178 pregnancies between the 11\textsuperscript{th} and 14\textsuperscript{th} gestational weeks. The study population included 82 twin pregnancies. All women gave informed consent before being enrolled in the study. The research-ethics committee of the participating university hospitals approved the study. Alongside the study, normal population medians, standard deviations and correlation co-efficients were calculated for NT (n=3,102), maternal serum β-hCG and PAPP-A (n=4,108). A summary of the patient characteristics of the study population is presented in table 7.

Table 7. Patient characteristics of the study population.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients screened by NT</td>
<td>3,178</td>
</tr>
<tr>
<td>Patients screened by serum samples</td>
<td>4,617</td>
</tr>
<tr>
<td>Median age (range)</td>
<td>29.9 years (15 – 48)</td>
</tr>
<tr>
<td>Proportion of mothers ≥ 35 years</td>
<td>19 %</td>
</tr>
<tr>
<td>Mean weight (range)</td>
<td>66 kg (39 – 179)</td>
</tr>
<tr>
<td>Mean duration of pregnancy (range)</td>
<td>86 days (51 – 97)</td>
</tr>
<tr>
<td>Median NT (range)</td>
<td>1.4 mm (0 – 14)</td>
</tr>
</tbody>
</table>
4.2 Methods

Blood samples were frozen and sent to Wallac OY, Turku, Finland, where the maternal serum PAPP-A and β-hCG concentrations were analyzed. The serum analysis was performed using AutoDELFIA PAPP-A and β-hCG kits (PerkinElmer, Wallac, Turku, Finland). The within and between assay variation for β-hCG were both <3.4%, and for PAPP-A <1.4% and <4.8% respectively. The analytical sensitivities of β-hCG and PAPP-A were 0.2 ng/mL and 5 mU/L respectively. The results were given as multiples of medians. Quality was assessed regularly, using in-house controls. The assay laboratory was provided with the following patient information: date of birth, weight, first day of last menstrual period, gestational age by ultrasound, diabetic status, use of insulin, and the number of fetuses. The results of the serum test were not given to the patients.

Nuchal translucency was measured by specialized doctors or health care providers tutored by doctors. In clinical work, NT ≥ 3 mm was considered as a cut-off for the recommendation of further examinations in the Kuopio University region. In contrast with this, Oulu University Hospital applied the risk figure program developed by professor Nicolaides. For the purposes of the study, the risk figure was also calculated using Life Cycle, the Wallac software risk figure program. This program takes into account maternal age, fetal CRL, NT thickness, and maternal serum results. The LifeCycle program has been extended to include pregnancies before the 11th week. Women with risks greater than that of a 35-year-old at the same gestational age, i.e. 1:250, were considered to be at increased risk for Down syndrome. This cut-off was also chosen to yield a reasonable FPR of 5.8% in the combined screening. Corrections were made for maternal weight but not for diabetic status or smoking.

The Register of Congenital Malformation and The National Research and Development Center for Welfare and Health provided the information about pregnancy outcome and DS cases among screenees. The Register of Congenital Malformations contains data from 1963 onwards on congenital anomalies detected in stillborn infants and in live born infants before the age of one year. The register receives data on congenital anomalies from hospitals, health care professionals and cytogenetic laboratories as well as from the Birth and Care Registers and the Cause of Death Statistics.

Paper I is a preliminary report of the screening results from 2,515 patients screened by maternal serum samples and 1,602 patients screened by both serum samples and nuchal translucency measurement in the years 1998-1999. There were 8 cases of Down syndrome.

Paper II. The MoMs of maternal serum concentrations of PAPP-A and β-hCG in spontaneous and in in vitro fertilization pregnancies were compared by using a two-tailed pooled t-test after log10 transformation of the serum concentrations. Logarithmic transformation of the data allowed the use of parametric tests because the logarithmic serum marker levels were shown to fit Gaussian distributions.

Paper III. The MoMs of maternal serum concentrations of PAPP-A and β-hCG in singleton and twin pregnancies were compared using the same method as in paper II.

Paper IV is a retrospective comparison of the impact of two screening methods on the live born incidence of Down syndrome in two different populations. The data were
collected from the records of Kuopio University Hospital and Joensuu Central Hospital between the years 1992 – 2000. The first group comprised 47,225 newborns whose mothers underwent second trimester serum screening (AFP, ß-hCG). The second group included 10,145 newborns whose mothers attended first trimester screening based on fetal nuchal translucency measurement. The expected numbers of live born Down’s children in the absence of screening were calculated by using maternal age-specific rates. The statistical significance between the expected and actual numbers of affected children was compared using the chi-square goodness-of-fit test.

*Paper V* produces the Finnish population medians of nuchal translucency and serum markers. The correlation co-efficients between the markers were derived by a linear regression analysis. The final Down syndrome screening results of the whole study population are also presented. Again, logarithmic transformations of the MoMs were used.

*Paper VI*. The MoMs of nuchal translucency and maternal serum concentrations of PAPP-A and ß-hCG were compared between smokers and non-smokers using the same method as in paper II.
5 Results

5.1 Population medians and the performance of the screening tests

The population medians of NT, PAPP-A and β-hCG in normal pregnancies are shown in table 8. The correlation coefficients between the markers were derived by a linear regression analysis and were as follows: PAPP-A & β-hCG –0.03, PAPP-A & NT 0.04, and β-hCG & NT –0.05.

Table 8. The population medians of NT (n=3,102), PAPP-A and β-hCG (n=4,108) in normal pregnancies.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Median MoM</th>
<th>Mean log10 MoM</th>
<th>SD log10 MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT</td>
<td>1.06</td>
<td>0.0226</td>
<td>0.1529</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>1.00</td>
<td>-0.0203</td>
<td>0.2781</td>
</tr>
<tr>
<td>β-hCG</td>
<td>1.07</td>
<td>0.0382</td>
<td>0.2851</td>
</tr>
</tbody>
</table>

There were 16 DS cases in the whole study group, and 13 DS cases in the group which also underwent NT measurement. The population medians of NT, PAPP-A and β-hCG for DS-pregnancies are shown in table 9. The DS cases are presented in table 10.

Table 9. The population medians of NT (n=13), PAPP-A and β-hCG (n=16) in Down syndrome pregnancies.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Median MoM</th>
<th>Mean log10 MoM</th>
<th>SD log10 MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT</td>
<td>2.33</td>
<td>0.3316</td>
<td>0.3067</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>0.44</td>
<td>-0.4426</td>
<td>0.4800</td>
</tr>
<tr>
<td>β-hCG</td>
<td>2.14</td>
<td>0.3270</td>
<td>0.2734</td>
</tr>
<tr>
<td>Case</td>
<td>Maternal age (years)</td>
<td>CRL (mm)</td>
<td>gestation (weeks)</td>
</tr>
<tr>
<td>------</td>
<td>----------------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>1</td>
<td>44</td>
<td>66</td>
<td>13+0</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>65</td>
<td>12+6</td>
</tr>
<tr>
<td>3</td>
<td>21 **</td>
<td>10+3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>27 **</td>
<td>12+3</td>
<td>4.1</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>67</td>
<td>13+0</td>
</tr>
<tr>
<td>6</td>
<td>37</td>
<td>61</td>
<td>12+4</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>60</td>
<td>12+4</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>54</td>
<td>12+1</td>
</tr>
<tr>
<td>9</td>
<td>31</td>
<td>66</td>
<td>13+0</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
<td>61</td>
<td>12+4</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td>70</td>
<td>13+2</td>
</tr>
<tr>
<td>12</td>
<td>40 **</td>
<td>10+0</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>41</td>
<td>46</td>
<td>11+3</td>
</tr>
<tr>
<td>14</td>
<td>41</td>
<td>60</td>
<td>12+4</td>
</tr>
<tr>
<td>15</td>
<td>30</td>
<td>63</td>
<td>12+4</td>
</tr>
<tr>
<td>16</td>
<td>43</td>
<td>49</td>
<td>11+5</td>
</tr>
</tbody>
</table>

*Risk based on combining maternal age, NT, PAPP-A and β-hCG if NT was measured. Otherwise the risk is based on maternal age and serum markers only. The PerkinElmer LifeCycle risk figure program was used.
**CRL not measured.

Nuchal translucency ≥ 3 mm would have detected 46% (6/13) (95% CI 19-73%) of Down syndrome fetuses (table 10). The number of screen positive cases (NT ≥ 3 mm) was 1.2%. Combining NT with maternal age in a risk figure program (LifeCycle) would have detected 3 more affected cases, thus giving a detection of 69% (9/13) (95% CI 44-94%). The maternal age and serum screening with PAPP-A and β-hCG without NT measurement would have detected 75% (12/16) (95% CI 54-96%) of affected pregnancies.

Finally, the combination of age, NT and biochemical markers would have detected 77% (10/13) (95% CI 54-100%) of the Down’s for a 5.8% FPR when a risk cut-off of 1 in 250 was applied. The positive predictive value for the combined test was 5.1%. This indicates that 19 invasive procedures would have been needed to find one affected case and one normal fetus would have been lost for every five detected DS cases. With maternal age alone (≥ 35 years), the DR would have been much lower (50%, 95% CI 25-75%), the FPR much higher (19%), the PPV only 1.3% and 76 invasive procedures would have been needed to diagnose one Down syndrome case.

Figure 1 shows the NT values of the 13 Down syndrome cases and different NT percentiles in relation to increasing CRL.
The use of the 98th percentile of NT instead of a fixed 3 mm as a cut-off would have found two more cases of Down syndrome (fig. 1). Thus, the DR would have risen to 62% (8/13) (95% CI 36-88%). Using the 95th percentile would have detected one further affected fetus giving a detection of 69% (9/13) (95% CI 44-94%). However, this method would have increased the observed FPR from 1.2% to 2% or 5% respectively.

Figures 2 and 3 show the PAPP-A and β-hCG concentrations (after logarithmic transformation) of the 16 Down’s cases in relation to the main percentiles of the unaffected pregnancies. Figure 2 demonstrates that 4 out of 16 DS pregnancies (25%) had maternal serum PAPP-A values below the 5th percentile of the unaffected population. Figure 3 shows that 5 out of 16 DS pregnancies (31%) had maternal serum β-hCG values at or above the 95th percentile of the unaffected population.
Fig. 2. PAPP-A concentrations (as logarithmic transformations) of the 16 DS fetuses in relation to the 5th (dotted line) and the 50th (solid line) percentiles of the unaffected population.

Fig. 3. ß-hCG concentrations (as logarithmic transformations) of the 16 DS fetuses in relation to the 50th (dotted line) and the 95th (solid line) percentiles of the unaffected population.
5.2 Serum concentrations in IVF pregnancies

In a comparison between 49 consecutive in vitro fertilization pregnancies and 4,265 spontaneous singleton control pregnancies, elevated levels of β-hCG were detected. PAPP-A concentrations were not different between the groups. The results are shown in table 11. Figure 4 presents the distributions of β-hCG values in the IVF and in the control group.

Table 11. Geometric means of the serum concentrations in IVF pregnancies.

<table>
<thead>
<tr>
<th>Marker</th>
<th>IVF pregnancies</th>
<th>Controls</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ß-hCG</td>
<td>1.25 MoM</td>
<td>1.03 MoM</td>
<td>0.029</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>1.03 MoM</td>
<td>0.99 MoM</td>
<td>0.451</td>
</tr>
</tbody>
</table>

Fig. 4. Maternal serum β-hCG values (MoM) in natural and IVF pregnancies.

Among the screenees there were no pregnancies affected by Down syndrome. In the control group the FPR with combined screening was 5.4% whereas in the IVF group it was 12.2%. This difference is partly explained by the higher number of women over 34 years of age in the study group (26% versus 19%). However, after reducing the control group to include a 26% proportion of women over 34 years, the FPR was still higher among the IVF pregnancies (12.2% versus 7.2%). The mean duration of pregnancy in the study group was 90 days (SD 5.1 days) and in the control group it was 86 days (SD 7.8). The mean weight of 66.2 kg (SD 12.0 kg) was comparable to the controls (66.5 kg, SD 12.7 kg).
5.3 Serum concentrations in twin pregnancies

In this study, the subjects comprised 4,361 women. There were 82 twin pregnancies, a frequency of 1 in 54. However, in 15 twin pregnancy cases, the maternal weight was missing, and these cases were excluded from the analysis. Thus, a total of 67 women with twin pregnancies were enrolled in the study. The geometric means of ß-hCG and PAPP-A in the twin pregnancies were calculated along with the 5th, 50th, and 95th percentiles of the singleton controls.

The median age of the women with twin pregnancies was 30.8 years (range 21 - 44 years) compared to the median of 29.9 years (range 15 – 48) of singleton controls. The mean maternal weight was similar in the twin group and in the singleton group, 66.0 kg (SD 11.6 kg) and 66.5 kg (SD 12.7 kg), respectively. Women with a twin pregnancy participated somewhat later in the screening (mean 88 days, SD 6 days) than those with a singleton pregnancy (mean 86 days, SD 8 days). All twin pregnancies in the study were chromosomally normal.

The geometric means of ß-hCG for singleton and twin pregnancies were 1.08 MoM and 1.85 MoM (95% CI 1.58-2.26) and for PAPP-A 1.01 MoM and 2.36 MoM (95% CI 2.02-2.56) respectively (figure 5). The SDs of geometric means for singleton and twin pregnancies were comparable (p=0.46) for ß-hCG concentrations (log_{10} SD 0.29 MoM and 0.30 MoM respectively). In contrast, those for PAPP-A were significantly different (p<0.001) (log_{10} SD 0.28 MoM and 0.20 MoM respectively). The twin to singleton ratio of the markers was not constant across the range (table 12). The serum concentrations of ß-hCG and PAPP-A were not correlated to each other (correlation coefficient 0.07), or with maternal age (correlation coefficients 0.13 and 0.06 respectively). There was no significant difference in serum levels between dichorionic (n=54) and monochorionic (n=13) twins.

<table>
<thead>
<tr>
<th>Centile</th>
<th>Singleton</th>
<th>Twins</th>
<th>Twin: singleton ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAPP-A</td>
<td>ß-hCG</td>
<td>PAPP-A</td>
</tr>
<tr>
<td>5⁰</td>
<td>0.34</td>
<td>0.39</td>
<td>0.89</td>
</tr>
<tr>
<td>50⁰</td>
<td>1.00</td>
<td>1.07</td>
<td>2.32</td>
</tr>
<tr>
<td>95⁰</td>
<td>2.36</td>
<td>3.34</td>
<td>5.17</td>
</tr>
</tbody>
</table>

Table 12. Marker MOMs and twin to singleton ratios of the main centiles.
5.4 Impact of screening on the live born incidence of Down syndrome

In the first group, maternal midtrimester serum screening (AFP, ß-hCG) was offered as a public policy to all pregnant women (n = 54,097), with an uptake of 87%. The mean maternal age was 29.6 years (SD 5.4) with 16% being older than 35 years. Thus a trisomy 21 birth prevalence rate of 97 cases (1: 490) could be expected on the basis of maternal age distribution (Hecht & Hook 1994).

In the other group, first trimester ultrasonographic screening was offered to all pregnant women (n=10,352), and 98% of them elected to participate. The mean maternal age was also 29.6 years (SD 5.7), and 19% were older than 35 years, with the expected live-born trisomy 21 rate thus being 22 cases (1: 460) (Hecht & Hook 1994).

There were 118 DS cases in the second trimester screening group, of which 69 cases were detected; thus the detection rate of DS was 58%. In the first trimester screening group there were 33 affected cases, and 18 were detected, giving the DR of 55%. The parents decided to continue the pregnancy in one case in both groups, and these were not counted as live-borns. The gestational age-corrected expected incidence of DS at the time of the screening was in the serum group 123 cases (assuming 21% fetal loss rate) and in the ultrasound group 32 cases (assuming 31% fetal loss rate). Thus, the real detection rate of trisomy 21 was 56% (69/123) in the serum group and 56% (18/32) in the ultrasound group.
The midtrimester screening was potentially able to reduce the number of live-born affected children by 52% (50 cases), and the first trimester screening by 32% (7 cases). This difference between expected numbers of DS newborns in the absence of any screening and actual observed cases was statistically significant in the serum screening group \((p < 0.001)\), but not in the ultrasound group \((p = 0.135)\). The rate of CVS or AC was 6.4% in the serum group and 4.2% in the ultrasound group. The results are summarized in table 13.

**Table 13. Impact of screening on the live born incidence of Down syndrome.**

<table>
<thead>
<tr>
<th>Down syndrome</th>
<th>2nd trimester biochemical screening</th>
<th>1st trimester ultrasound screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 47,225</td>
<td>n = 10,145</td>
<td></td>
</tr>
<tr>
<td>Expected live-born*</td>
<td>97</td>
<td>22</td>
</tr>
<tr>
<td>Actual live-born</td>
<td>47</td>
<td>15</td>
</tr>
<tr>
<td>Decrease of live-born incidence</td>
<td>52% (50/97)</td>
<td>32% (7/22)</td>
</tr>
<tr>
<td>All cases</td>
<td>118</td>
<td>33</td>
</tr>
<tr>
<td>Detected</td>
<td>58% (69/118)</td>
<td>55% (18/33)</td>
</tr>
<tr>
<td>Rate of invasive testing</td>
<td>6.4%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

*Counted on the basis of natural live-born DS ratio (Hecht & Hook 1994) and the maternal age distribution.

### 5.5 The influence of smoking on the first trimester screening parameters

For this study, only normal singleton pregnancies were included; thus the NT group and the serum screening group constituted 3,115 and 4,436 women respectively. Smoking habits were assessed by a self-reporting procedure during testing.

Smoking habits were recorded in 4,279 women (96%) in the serum screening group and in 2,903 women (93%) in the NT group. In the two groups, there were 454 (11%) and 284 (10%) smokers respectively. The mean weight was 66.4 kg (SD 12.6 kg) among non-smokers and 67.0 kg (SD 13.7 kg) among smokers respectively. The mean duration of gestation was 86 days (SD 7.7 days) in the non-smoking group and 85 days (SD 8.4 days) in the smoking group respectively. The smokers were slightly younger, with a mean age of 28.4 years (SD 6.2 years) compared to the non-smokers’ mean age of 29.7 years (SD 5.4 years). The reduction of PAPP-A among the women who smoked during pregnancy was statistically significant. ß-hCG levels were unaltered. NT was significantly thicker in the smoking group. The results of the study are shown in table 14.
Table 14. Median MoM, mean log_{10} MoM, log_{10} SD and statistical significance of PAPP-A, ß-hCG and NT, classified by smoking status.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Smokers, n = 454</th>
<th>Non-smokers, n = 3,825</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median MoM</td>
<td>Mean log_{10} MoM</td>
<td>log_{10} SD</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>0.81</td>
<td>-0.10</td>
<td>0.29</td>
</tr>
<tr>
<td>ß-hCG</td>
<td>1.06</td>
<td>0.03</td>
<td>0.31</td>
</tr>
<tr>
<td>NT*</td>
<td>1.13</td>
<td>0.05</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*The NT group comprised 284 smokers and 2619 non-smokers.

Lower concentrations of PAPP-A could logically lead to an increased false positive rate among smokers. Indeed, those with a positive screening risk ≥ 1:250 for Down syndrome constituted 5.4% of the non-smoking group and 7.1% of the smoking group, although the smokers were younger than the non-smokers and thus had a lower background risk for Down syndrome.
6 Discussion

6.1 Efficacy of the screening tests

In the present study of 4,617 pregnant women including 16 cases of DS, the sensitivity of the combined screening (77%) is in keeping with earlier reports. The serum screening without measurement of NT performed excellently, achieving a high detection rate of 75%. Earlier studies have given sensitivities in the order of 60%. This difference may be due to chance because the number of affected cases was limited. On the other hand, laboratory assays have developed as has ultrasound technology during recent years. All the assays in this study were done in a single research laboratory.

In the present series of 3,178 NT measurements and 13 DS cases, the DR for maternal age and NT was 69%. This result is in agreement with previous reports. Using the 98th or 95th percentile of NT instead of fixed 3 mm as a cut-off limit would have improved the DR from 46% to 62% or 69% respectively. However, the FPR would have increased as well. It is an advantage of computerized risk figure programs that they calculate a personal risk figure to all examined women. Recently, tables for “bedside” estimation of DS risk based on maternal age and NT measurements have been constructed (Herman et al. 2002b). Those health care units which cannot acquire computer software might find that kind of application of a risk figure program a more appropriate solution than using NT percentiles instead of a fixed cut off value.

A few prospective combined first trimester screening studies have been carried out (table 15), not all of them comprising an unselected low-risk population. There are 4 prospective studies (by Schuchter et al, by Crossley et al, by Spencer et al and the present study) where the age distribution of the study population reflects that of the general population and these include 97 pregnancies affected by DS. Due to the small number of cases in the studies, the confidence intervals of detection rates of the different screening methods overlap. This means that the observed superiority of combined screening over plain NT or serum screening lacks statistical significance. However, all the studies in table 15 point out that the combined method yields better results than either NT or biochemistry alone, although the difference is rather small in the well planned study by Crossley et al. As regards trisomy 21, there is not enough prospective data from low-risk
populations at present to conclude that first trimester screening yields higher DRs than the 65-75% published in the second trimester studies. The many advantages of an early ultrasound examination will, however, stress the need to shift the screening to the 11th – 14th gestational weeks. Those centers which have already established the NT screening, should consider adopting the combined approach.

The combined screening method in the first trimester should be studied more extensively within large non-selected low risk populations before it can be recommended as a primary means of DS screening. This question will probably be answered soon when the results of the two major prospective trials, the SURUSS and the FASTER, are published. In clinical practice, the success of NT and thus combined screening will depend on the arrangements for the education of the sonographers and the auditing of the NT measurements.

If the present finding of a 75% DR could be later confirmed, screening might even be based on plain serum sampling at around 10 weeks of pregnancy, with further investigations then being offered to those with an increased or intermediate risk. The risk could be recalculated by CRL and NT measurement. Counseling and diagnostic tests would be offered only for those who still stay at high risk. This might reduce the cost of screening (Christiansen & Larsen 2002).

Table 15. The prospective studies of combined first trimester screening.

<table>
<thead>
<tr>
<th>Study</th>
<th>Population (n)</th>
<th>Mean/median maternal age (years)</th>
<th>≥ 35 years (%)</th>
<th>Down cases (n)</th>
<th>DR NT (%)</th>
<th>DR serum screening (%)</th>
<th>DR combined screening (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Orlandi et al. 1997)</td>
<td>2,010*</td>
<td>NS</td>
<td>35</td>
<td>11*</td>
<td>73</td>
<td>61</td>
<td>87</td>
</tr>
<tr>
<td>(De Biasio et al. 1999)</td>
<td>1,467</td>
<td>32</td>
<td>NS</td>
<td>13</td>
<td>61**</td>
<td>69</td>
<td>85**</td>
</tr>
<tr>
<td>(Krantz et al. 2000)</td>
<td>10,251***</td>
<td>32</td>
<td>NS</td>
<td>50***</td>
<td>74</td>
<td>63</td>
<td>91</td>
</tr>
<tr>
<td>(Schuchter et al. 2002)</td>
<td>4,939</td>
<td>NS</td>
<td>13</td>
<td>14</td>
<td>71****</td>
<td>NS</td>
<td>86</td>
</tr>
<tr>
<td>(Crossley et al. 2002b)</td>
<td>17,229</td>
<td>30</td>
<td>15</td>
<td>45</td>
<td>54</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>(Bindra et al. 2002)</td>
<td>15,030</td>
<td>34</td>
<td>47</td>
<td>82</td>
<td>79</td>
<td>60</td>
<td>90</td>
</tr>
<tr>
<td>(Spencer et al. 2003)</td>
<td>11,105</td>
<td>30</td>
<td>17</td>
<td>25</td>
<td>76</td>
<td>68</td>
<td>92</td>
</tr>
<tr>
<td>Present study</td>
<td>4,617*****</td>
<td>30</td>
<td>19</td>
<td>16*****</td>
<td>69</td>
<td>75</td>
<td>77</td>
</tr>
</tbody>
</table>

The DRs are combined with maternal age at a 5% FPR

*The combined screening group comprised 744 women with 7 DS cases
**FPR 6.7% for NT and 3.3% for the combined method
***The combined group comprised 5,809 women with 33 DS cases
****FPR 10%
*****The combined group comprised 3,178 women with 13 DS cases
NS= not stated
6.2 Serum concentrations in IVF-pregnancies

Several publications have indicated in recent years that mid gestation maternal serum levels of hCG, AFP and uE3 are altered in different forms of assisted conception (Ribbert et al. 1996, Heinonen et al. 1996, Barkai et al. 1996, Frishman et al. 1997, Maymon et al. 1999, Wald et al. 1999b, Räty et al. 2002). Of all the serum markers, high hCG / ß-hCG levels were the most consistent, although not all studies have confirmed this finding (Barkai et al. 1996, Lam et al. 1999, Maymon & Shulman 2002). The practical consequence of high levels of hCG in IVF-pregnancies is an almost doubled FPR. Because many women who undergo IVF treatment have suffered from infertility for years and are generally at the end of their reproductive age, the high FPR and subsequent high rate of invasive testing with a threat of pregnancy loss are especially harmful.

The cause of the high hCG levels in IVF pregnancies in the second trimester is unclear. Studies on women who conceived after oocyte donation (Maymon & Shulman 2001) and frozen embryo transfer (Perheentupa et al. 2002) without prior ovarian stimulation therapy showed that ovulation induction is not a likely reason.

Vaginal bleeding (Koivurova et al. 2002) and low-lying placentas (Tan et al. 1992) seem to occur more commonly in IVF pregnancies, suggesting that placentation might be more often poor in IVF cases. This may logically lead to diminished blood flow and reduced oxygen supply to the placenta. The hypothesis that a reduced oxygen supply to the trophoblast may result in increased production of hCG in complicated pregnancies has been supported by the results of studies carried out in vitro (Fox 1970). However, in a study of 46 consecutive spontaneous pregnancies affected by placenta previa, no differences in AFP and hCG concentrations were found (Heikkilä et al. 2000). Furthermore, a recent report suggested that term trophoblasts (Esterman et al. 1996) secrete less hCG under hypoxic conditions.

A positive correlation between the number of embryos transferred and the marker levels has been observed in IVF pregnancies (Raty et al. 2002). In this material, one embryo was transferred in about half of the cases and two in the rest.

Studies in the first trimester are still sparse (table 16).

**Table 16. Studies on serum markers in IVF-pregnancies in the first trimester.**

<table>
<thead>
<tr>
<th>Study</th>
<th>IVF- cases (n)</th>
<th>FPR (%)</th>
<th>PAPP-A MoM</th>
<th>ß-hCG MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Wojdemann et al. 2001)</td>
<td>47</td>
<td>4.7</td>
<td>4.9</td>
<td>1.02</td>
</tr>
<tr>
<td>(Liao et al. 2001)</td>
<td>220</td>
<td>15.9</td>
<td>7.0</td>
<td>1.00</td>
</tr>
<tr>
<td>(Maymon &amp; Shulman 2002)</td>
<td>71</td>
<td>15.9</td>
<td>9</td>
<td>0.96</td>
</tr>
<tr>
<td>(Orlandi et al. 2002)</td>
<td>32</td>
<td>13.8</td>
<td>6.1</td>
<td>0.79</td>
</tr>
<tr>
<td>Present study</td>
<td>49</td>
<td>12.2</td>
<td>7.2</td>
<td>1.03</td>
</tr>
</tbody>
</table>
Reports on the serum concentrations in IVF pregnancies seem to be more controversial in the first than in the second trimester. The present findings are in keeping with the largest study by Liao et al., that ß-hCG is already elevated in early pregnancy and the FPR is increased in IVF pregnancies. Results of PAPP-A levels are contradictory as well. However, FPRs between IVF pregnancies and spontaneous controls possibly differ less than in mid pregnancy. This might allow the use of combined first trimester screening in assisted pregnancy. Moreover, if larger trials confirm that IVF affects second trimester more than first trimester screening, those of us who explore the reasons for these serum alterations might preferably bet on poor placentation than the IVF procedure itself. Meanwhile, NT screening can be recommended to women who conceive by IVF, because NT measurement is not affected (Liao et al. 2001).

### 6.3 Serum concentrations in twin pregnancies

With the increased use of assisted reproduction, the last decade has seen an increase in multiple births from 1.2% to 1.5% (The National Research and Development Center for Welfare and Health). In addition, the rates of twinning are related to maternal age, such that women over 35 are three times more likely to conceive twins than are women under the age of 20.

The serum marker levels in chromosomally normal twin pregnancies are approximately double those of singletons in the second trimester (Wald et al. 1991). Based on this finding, mathematical risk prediction from singleton pregnancies has been extrapolated to twins using a twin correction, where the actual MoM value is divided by the median MoM of the twin population, and the risk calculation is then treated as for a singleton pregnancy. With such a policy, second-trimester serum screening has been estimated by modeling to yield a DR of about 50% at a 5% screen positive rate (Spencer et al. 1994, Neveux et al. 1996). However, the basic assumption that biochemical markers are twice as high as in singletons has been challenged (O’Brien et al. 1997) and therefore the question of whether a mere mathematical conversion of singleton trisomy risk can reasonably be applied to twins is not yet fully resolved.

Studies of twin pregnancies in the first trimester are still few (table 17).

<table>
<thead>
<tr>
<th>Study</th>
<th>Normal twins (n)</th>
<th>Down cases (n)</th>
<th>ß-hCG*</th>
<th>PAPP-A*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Berry et al. 1997)</td>
<td>50</td>
<td></td>
<td>1.97</td>
<td></td>
</tr>
<tr>
<td>(Noble et al. 1997a)</td>
<td>136</td>
<td>12</td>
<td>1.94</td>
<td></td>
</tr>
<tr>
<td>(Spencer 2000)</td>
<td>159</td>
<td></td>
<td>2.10</td>
<td>1.86</td>
</tr>
<tr>
<td>(Spencer &amp; Niclaides 2003)</td>
<td>206</td>
<td>4</td>
<td>2.15</td>
<td>1.93</td>
</tr>
<tr>
<td>Present study</td>
<td>67</td>
<td></td>
<td>1.85</td>
<td>2.36</td>
</tr>
</tbody>
</table>

*Median MoMs or geometric mean MoMs.
Noble et al. measured ß-hCG in 148 twin pregnancies and in 12 cases one or both fetuses had trisomy 21. In the DS group the median ß-hCG was significantly higher, but only one of the trisomic pregnancies had a level above the 95th percentile. Thus it was concluded that in twin pregnancies ß-hCG is unlikely to be useful in the prediction of trisomy 21. (Noble et al. 1997a). However, Spencer estimated that at a 5% FPR, the detection rate of serum screening is 52% in twins discordant and 55% in twins concordant for trisomy 21. Furthermore, it was predicted that combining NT and biochemistry would give DRs approaching 80%, with the benefit of ultrasound being able to locate the affected twin. (Spencer 2000). In a more recent report, the same author presented an observed DR of 75% (3/4) using the combined method in twin pregnancy (Spencer & Nicolaides 2003).

In the present study, it was found that the geometric means of ß-hCG and PAPP-A in twin pregnancies are about twice as high as in singleton controls in the first trimester, the geometric mean for ß-hCG being 1.85 MoM and for PAPP-A 2.36 MoM. The SDs of the serum markers were comparable for singletons and twins, the SD of PAPP-A was even tighter in twins. Thus far the results were consistent with the earlier report (Spencer 2000). However, the level of PAPP-A was significantly higher than the expected 2.0 MoMs (p=0.034) and the twin to singleton ratio of serum concentrations was not constant across the range, as can be seen in table 12. These findings suggest that the criticism towards the mathematical risk predictions extrapolated from singleton pregnancies to twins is not without basis (O’Brien et al. 1997).

It has been shown that chorionicity has no impact on maternal serum ß-hCG and PAPP-A levels (Spencer 2001a), as also observed in the present study. Further data are required in normal and affected twin pregnancies before combined screening can be recommended. Meanwhile, screening by NT is the method of choice in twin pregnancies (Sebire et al. 1996).

### 6.4 Impact of fetal loss on the screening

Fetal loss in pregnancies complicated by a chromosomal anomaly is a well-recognized phenomenon, as is the fact that the majority of these losses are assumed to occur in the first trimester. Thus when considering different screening programs for trisomy 21, due consideration needs to be given to the presence of more cases of trisomy 21 in the first trimester. Most studies of fetal losses have used data collected predominantly in women aged over 35 years undergoing CVS or AC in the period 1970-1990 for reasons of advanced maternal age. Table 18 summarizes some of these studies and the fetal loss rates between the time of the invasive procedure and term.
Table 18. Summary of studies of fetal loss rates between the time of the invasive procedure and term (Spencer 2001b).

<table>
<thead>
<tr>
<th>Study</th>
<th>Age range (years)</th>
<th>First trimester fetal loss rate (cases)</th>
<th>Second trimester fetal loss rate (cases)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hook et al. 1995</td>
<td>16-49</td>
<td>75% (8)</td>
<td>50% (168)</td>
</tr>
<tr>
<td>Halliday et al. 1995</td>
<td>36-43</td>
<td>31% (39)</td>
<td>18% (73)</td>
</tr>
<tr>
<td>Macintosh et al. 1995</td>
<td>35-48</td>
<td>48% (302)</td>
<td>24% (610)</td>
</tr>
<tr>
<td>Bray &amp; Wright 1998</td>
<td>35-50</td>
<td>31% (341)</td>
<td>12% (1159)</td>
</tr>
<tr>
<td>Morris et al. 1999</td>
<td>16-49</td>
<td>31% (441)</td>
<td>24% (2035)</td>
</tr>
<tr>
<td>Snijders et al. 1999</td>
<td>35-45</td>
<td>31% (221)</td>
<td>21% (317)</td>
</tr>
</tbody>
</table>

Screening for aneuploidy in the first rather than the second trimester contains a potential disadvantage in that earlier screening may preferentially identify those chromosomally abnormal fetuses that are destined to die in utero. If this is true, the overall screening impact on live birth prevalence of DS could be much less impressive than the published high DR figures.

Because most affected and detected fetuses are terminated, it is not known how many of them would be born alive. Thus the real impact of screening on live born prevalence of affected children is not clear. If spontaneous fetal loss is similar among screen positive and negative pregnancies, then the reduction of birth prevalence is equal to the detection rate. However if fetal loss is more common among screen positive than screen negative conceptuses, the reduction of live born Down’s children is less than the detection rate suggests.

In the present study, the second trimester serum screening (AFP, β-hCG) was more effective in reducing the live born incidence of DS compared to the first trimester NT screening, when the DRs were at the same level. Calculated per 10,000 deliveries in a Finnish population, NT screening would allow 4.8 more Down’s children to be born, although firm conclusions should not be drawn due to the rather limited number of affected cases.

The 52% reduction of DS birth prevalence in the serum group achieved statistical significance ($p<0.001$) but the reduction in NT screening was not as efficient (32%, $p=0.135$). This difference can be understood in three ways. Firstly, it has to be mentioned in favor of ultrasound screening that the rate of invasive testing was significantly lower compared to the serum screening; 4.2% vs 6.4% respectively. Unfortunately, the data does not allow the FPR to be fixed. It is known that the DR and the FPR are correlated, i.e., a higher FPR means a higher DR. Thus, if the FPR for both measurements were fixed at 5%, for example, it is possible that the conclusion of the impact on the live-born ratio would be different. The same DRs of Down syndrome in the present study in both groups alongside the difference of live-born cases indicates that NT screening will lead to termination of more nonviable fetuses. Secondly, the results included the early years of NT screening; the performance of the ultrasound examiners could have been improved since then by means of training and external auditing. Thirdly, there is a concern that the first trimester NT screening preferentially identifies those fetuses which would die in utero anyway.
Increased NT has been associated with higher fetal loss rates both in chromosomally normal (Pajkrt et al. 1999) and in trisomy 21 fetuses (Hyett et al. 1996b). If the results of the present study can be confirmed, a re-evaluation of the cost effectiveness of various Down syndrome screening methods would be warranted.

### 6.5 Impact of maternal smoking on the first trimester screening parameters

In the second trimester, expectant mothers who smoke have higher levels of AFP, and lower levels of uE3 and both total HCG and ß-hCG AFP was increased by 7% in normal pregnancies (Crossley et al. 2002a) and 10% in Down pregnancies (Spencer 1998). Estriol was reduced by 3% (Palomaki et al. 1993), and the effect is stronger in Down cases (24%) (Cuckle et al. 1990). The reduction in HCG was reported as 29% in normal pregnancies and as much as 39% in affected pregnancies (Crossley et al. 2002a). The only study on ß-hCG reported a 14% reduction in normal pregnancies and 16% decrease in DS cases (Spencer 1998). Inhibin-A might be strongly elevated due to smoking (Ferriman et al. 1999). Others have considered these effects large enough to propose an adjustment of the DS risk evaluation algorithm according to smoking habits (Crossley et al. 2002a), at least in a population where smoking is common (Spencer 1998).

The findings of the present study on the influence of smoking on serum screening analytes in the first trimester are in keeping with the two previous studies (Spencer 1999, de Graaf et al. 2000). The data show that PAPP-A is reduced by about 20% in the first trimester if the mother smokes. In contrast to the evidence in the second trimester, smoking is not associated with reduced ß-hCG in early pregnancy. Using modeling techniques, Spencer suggested that in smokers the detection of trisomy 21 by ß-hCG, PAPP-A and maternal age is reduced by 5-6% compared with that of the general population (Spencer 1999).

The results of the first trimester serum studies are summarized in table 19.

**Table 19. Studies of the influence of smoking on the first trimester serum analytes.**

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>PAPP-A MoM</th>
<th>ß-hCG MoM</th>
<th>AFP MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>S</td>
<td>NS</td>
<td>S</td>
</tr>
<tr>
<td>(Spencer 1999)</td>
<td>2,287</td>
<td>600</td>
<td>1.00</td>
<td>0.85</td>
</tr>
<tr>
<td>(de Graaf et al. 2000)</td>
<td>1,247*</td>
<td>117*</td>
<td>1.07</td>
<td>0.81</td>
</tr>
<tr>
<td>Present study</td>
<td>3,825</td>
<td>454</td>
<td>1.03</td>
<td>0.81</td>
</tr>
</tbody>
</table>

*For PAPP-A there were 703 non-smokers and 52 smokers. NS = non-smokers, S = smokers.
The reasons why smoking affects the serum marker concentrations are still poorly known. Smoking has been shown to damage the placental barrier and to disturb transportation across the placenta (Demir et al. 1994). The placental syncytiotrophoblast undergoes apoptosis, and this process is inhibited by smoking, which in turn might modify the maternal-fetal exchange (Marana et al. 1998). The number of placental areas presenting with syncytiotrophoblastic necrosis was significantly higher in the smoking than in the non-smoking group (Jauniaux & Burton 1992).

In this study, the increase of NT was significant among smokers. This finding may not be relevant clinically due to the small difference between the groups. However, the impact of smoking on NT raises more interest when the reasons for increased NT among normal pregnancies are considered. An impaired or delayed development of the fetal lymphatic system is mentioned as a possible explanation (Greco et al. 1996), and smoking might affect this process. Increased NT is associated with higher intrauterine mortality in normal pregnancies (Souka et al. 2001) but smoking is not (Pandya et al. 1996). Other factors affecting NT through unknown mechanisms include fetal sex (Lam et al. 2001) and maternal ethnic origin (Chen et al. 2002).

The best antenatal screening method seems to be the combination of NT, PAPP-A and ß-hCG in the first trimester of pregnancy. With PAPP-A being the strongest of the serum markers, the 15-20% reduction of PAPP-A levels due to smoking may have an impact on the performance of the screening. Smokers may be more often classified as being at high risk for Down syndrome. Although the influence may be small on the overall screening for DS, reducing the PAPP-A median for smokers by 20% might improve the accuracy of the risk evaluations given to individual women.
7 Conclusions

The following conclusions can be drawn from the results of the present study:

1. Screening for Down syndrome in the first trimester by combining the measurement of fetal nuchal translucency and maternal serum β-hCG and PAPP-A is an efficient method, also among unselected low risk women. Those centers which have established the NT screening should consider adopting the combined approach.

2. In IVF-pregnancies, β-hCG is elevated in the first trimester due to unknown reasons, increasing the false positive rate. Serum screening is not recommended in these pregnancies.

3. In twin pregnancies, the serum marker concentrations are approximately double those in singletons. PAPP-A production may be slightly elevated in twin pregnancies. The standard deviations for singletons and twins are comparable for β-hCG concentrations, whereas those for PAPP-A are significantly different. The twin to singleton ratio of the marker values is not constant over the range. These results imply that the mathematical simulation models, where the DS risks of singleton pregnancies are converted into those of twin pregnancies, may not be reliable.

4. The first trimester ultrasound screening based on measurement of nuchal translucency seems to decrease less the live born incidence of Down’s children, compared with the second trimester maternal serum double screening, when the detection rate of the methods is similar. There is a concern that NT screening identifies preferentially those DS fetuses which are destined to miscarry.

5. Mothers who smoke during pregnancy have lower values of PAPP-A. This increases the likelihood of a positive screening result, i.e. smoking mothers are more often classified as being at high risk for Down syndrome. Smoking seems to influence NT as well, but without clinical relevance.
References


Spencer K (2001b) What is the true fetal loss rate in pregnancies affected by trisomy 21 and how does this influence whether first trimester detection rates are superior to those in the second trimester? Prenat Diagn 21: 788-789.


