Review article

Advances in prenatal screening for Down syndrome: I. General principles and second trimester testing

Peter A. Benn*

Division of Human Genetics, Department of Pediatrics, University of Connecticut Health Center, 263 Farmington Avenue, Farmington, CT 06030-6140, USA

Received 5 March 2002; received in revised form 22 May 2002; accepted 22 May 2002

Abstract

Background: Down syndrome is one of the most important causes of mental retardation in the population. In the absence of prenatal screening and diagnosis, prevalence at birth in the United States would currently exceed 1:600. The purpose of prenatal screening is to identify those women at the increased risk for an affected pregnancy and to maximize the options available to these women. Tests available: Second trimester serum screening involves combining the maternal age-specific risk for an affected pregnancy with the risks associated with the concentrations of maternal serum alpha-fetoprotein (MSAFP), un conjugated estriol (uE3), and human chorionic gonadotropin (hCG) (triple testing). A fourth analyte, inhibin-A (INH-A), is increasingly being utilized (quadruple testing). Optimal second trimester screening requires the integration of a number of clinical variables, the most important of which is an accurate assessment of gestational age. In addition to Down syndrome, the triple and quadruple tests preferentially identify fetal trisomy 18, Turner syndrome, triploidy, trisomy 16 mosaicism, fetal death, Smith–Lemli–Opitz syndrome, and steroid sulfatase deficiency. Some programs modify the Down syndrome risks generated through maternal serum screening tests with fetal biometric data obtained by ultrasound. Other second trimester tests have shown promise, including the analysis of maternal urine and fetal cells in the maternal circulation, but none are in routine clinical use. Conclusion: The second trimester triple and quadruple tests provide benchmarks for evaluating new screening protocols. The combination of fetal biometry, new test development as well as clarification of the role of co-factors that affect the concentrations of analytes in existing tests should lead to greater efficacy in second trimester screening for Down syndrome.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Down syndrome; Second trimester; Screening; Serum tests; Alpha-fetoprotein; Aneuploidy

1. Introduction

1.1. Prenatal screening for Down syndrome: general principles

Screening has been defined as 'the systematic application of a test or inquiry, to identify individuals at sufficient risk of a specific disorder to benefit from further investigation or direct preventative action,
among persons who have not sought medical attention on account of symptoms of the disorder’ [1]. Criteria for worthwhile screening programs have been specified [2]. These include a well-defined medically important disorder with known prevalence and tests that are cost-effective, safe, accessible, and have well-defined performance.

Identification of women who are at high risk for fetal Down syndrome is consistent with this definition and the measures of utility. Patients who are screen-positive are generally offered ultrasound evaluations of the fetus, genetic counseling, and definitive diagnosis through cytogenetic analysis of either chorionic villi in the first trimester or amniotic fluid cells in the second trimester. Women with a Down syndrome affected fetus frequently choose to terminate their pregnancy [3]. However, this is not the goal of prenatal screening and diagnosis. The goal has been clearly defined by Peter Rowley [4] who notes that ‘the aim of genetic screening programs and prenatal diagnosis should be to maximize the options available to families rather than to reduce the prevalence of the disease’.

Screening tests are generally evaluated in terms of detection rate (sensitivity), false-positive rate, and odds of being affected given a positive result (OAPR). These performance-based measures are frequently used to justify particular protocols. However, maximizing options for individual patients must be paramount. Timely transfer of information together with a respect for patients’ ethical values, sensitivities, and decision options at every step in the prenatal testing pathway are some of the most important aspects of screening and diagnosis of Down syndrome.

1.2. Clinical aspects of Down syndrome

Down syndrome is clinically characterized by mental retardation, birth defects, and specific physical features that are identifiable at birth [5]. Mental retardation ranges from mild to severe with most cases showing a moderate level. At birth, cardiac defects are present in approximately 56% of Down syndrome individuals, 11% show digestive tract anomalies and a diverse group of other anatomical defects may also be present [6].

In the absence of prenatal screening and diagnosis, it can be estimated that Down syndrome prevalence at birth in the United States would currently exceed 1 in 600 [7]. Approximately 85–90% of individuals born with Down syndrome can be expected to survive to 1 year of age [8,9] and over 50% will be expected to survive beyond 50 years [10]. This disorder is therefore one of the most important potential causes of mental retardation in the population.

Down syndrome is caused by the presence of an additional copy of chromosome 21 [5]. This additional copy is usually the result of a maternal meiotic nondisjunction event but approximately 4% of cases are attributable to the unbalanced segregation of a Robertsonian translocation. In approximately 1% of cases, mosaicism is present and these individuals may show a milder phenotype.

1.3. Maternal age

Penrose [11] first noted the association between Down syndrome prevalence at birth and maternal age in 1933. Numerous studies have confirmed the association using data derived from birth certificates, hospital records, cytogenetic laboratories, and other sources. Prior to the introduction of biochemical screening tests, maternal age alone was used as a screening test. For example, for the United States population in the 1970s, it was estimated that offering amniocenteses to all women aged 35, or more, potentially allowed 20% of Down syndrome cases to be identified [12]. More recent data for the United States shows that nearly half of the Down syndrome affected pregnancies are present in women aged 35, or more [7].

With the introduction of maternal serum screening, there was a need for maternal age-specific prevalence rates that could be applied to any woman and a regression curve was developed by Cuckle et al. [13]. This rate schedule remains embedded in many computer programs used in screening. A critical reevaluation of the data [14] suggested that these widely used rates may somewhat underestimate prevalence, particularly for older women, and alternate rate schedules have been proposed [15,16]. For example, the incidence at birth among 35-year-old women is 1:384 according to the curve developed by Cuckle et al. [13] and is 1:336 using the eight-study schedule of Bray et al. [16]. Cuckle [17] considered the effect of using different maternal age-specific prevalence curves on the efficacy of various second trimester
screening protocols and concluded that use of alternative schedules had a rather minor effect on the detection rate. Maternal age adjusted Down syndrome birth prevalence appears to be consistent in different populations [18,19], indicating the widespread applicability of the published rates.

While birth prevalence can be considered to be reasonably well defined, there is much more uncertainty as to prevalence at various times in pregnancy. A substantial proportion of Down syndrome affected fetuses do spontaneously abort but determining this proportion has been problematic. Surveys have been carried out to determine the outcome of prenatally detected cases [20], but these studies have been based on relatively small numbers and data collection may have been subject to ascertainment bias. Use of more recent data from cytogenetic registries has also been analyzed but this also assumes that Down syndrome pregnancies identified prenatally are representative of all cases [21]. The most widely used estimates for prenatal survival of affected pregnancies have been derived by comparing birth prevalence and prenatal prevalence for cases in older women [22]. A constant loss rate across all maternal ages is assumed. Using this approach, estimates for affected pregnancy loss rates from the time of amniocentesis to birth are 10–24%, depending on the birth prevalence curve selected. Standardization of the maternal-age specific risks against a single source of prenatal data minimizes differences arising from the use of the various birth prevalence curves and is therefore preferable from a counseling standpoint [23]. Although Down syndrome prevalence at various times in utero is not well defined, it is a common practice to report risks at the time of the screening test. Those pregnancies that have a risk greater than a predetermined cut-off, for example, 1:270 in the second trimester, are identified as the high-risk, screen-positive, group.

2. Second trimester serum markers

2.1. Alpha-fetoprotein

In 1984, Merkatz et al. [24] reported that maternal serum alpha-fetoprotein (MSAFP) levels were lower in pregnancies in which fetal chromosomal abnormalities (primarily Down syndrome) were present. In a highly perceptive comment, they suggested that MSAFP, maternal age, and other relevant parameters could be combined to construct a risk profile that would improve the detection of the most serious chromosome abnormalities.

Because MSAFP assay was already well established as a screening tool for open neural tube defects, development of a biochemical-based Down syndrome screening protocol could be implemented relatively easily. Initial strategies used fixed MSAFP cut-offs (e.g. ≤ 0.5 multiples of the median (MoM)) in combination with maternal age to identify high-risk pregnancies [25,26]. This was replaced by a method in which the maternal age-specific risk is multiplied by a likelihood ratio determined by the heights of the MSAFP Gaussian distributions in affected and unaffected pregnancies [27,28]. Fig. 1 illustrates the principle of using Gaussian distributions to modify risk. Using a 1:270 second trimester Down syndrome risk cut-off (equivalent to maternal age 35 in the absence of serum screening), it was estimated that MSAFP screening would allow an additional 20% of all affected pregnancies to be identified [28]. Prospective studies confirmed the efficacy of this screening [29]. The risk-based method of screening using Gaussian

![Fig. 1. Use of Gaussian distributions to adjust risk. In this example, the test result is 1.5 MoM. The relative probability that the result is from the unaffected population is given by the height, \( n \), of the unaffected distribution at 1.5 MoM and the relative probability that the result is from the affected population is given by the height, \( d \), of the affected distribution. The likelihood ratio is \( d/n = 3.2/1.1 = 2.91 \). If the maternal age specific risk for Down syndrome is 1:500, the risk following the screening test is 2.91:500 or 1:172. Likelihood ratios derived from independent tests can be multiplied together. When tests are not completely independent, the correlation factors need to be factored into the calculations.](image-url)
distributions of markers has become the established model for Down syndrome screening.

The biological function of AFP in the fetus remains poorly defined [30] and the reason why MSAFP levels are lower in Down syndrome pregnancies is also unclear. Studies on the synthesis of AFP by fetal liver in normal and Down syndrome fetuses have yielded inconsistent results [31,32]. High levels of AFP have been found in the placentas of affected pregnancies suggesting a defect in the secretion of AFP into the maternal circulation [32].

2.2. Human chorionic gonadotropin

Bogart et al. [33] showed that second trimester maternal serum human chorionic gonadotropin (hCG) levels are generally higher in maternal serum when fetal Down syndrome is present. They noted that hCG appeared to be superior to MSAFP in detecting fetal chromosome abnormalities. Because of the widespread availability of hCG assays for pregnancy detection and monitoring, a rapid introduction of the testing as an adjunct to Down syndrome screening was possible.

HCG is a glycoprotein composed of two dissimilar subunits, α and β, produced by the placenta. In addition to intact hCG, maternal serum contains free α, free β, and degradation products (nicked hCG) [34]. Intact hCG and free β-hCG show peak concentrations at 8–10 weeks gestation while free α-hCG does not peak until much later in pregnancy [35]. In the second trimester, assays to both α- and β-subunits will help identify Down syndrome pregnancies. However, testing with an antibody that identifies all β-subunits (intact hCG and free β-hCG combined) appears to be superior. Assays also exist that identify only the free β-hCG but these appear to have little advantage over those that measure total hCG. For example, it is estimated that, with a 5% false-positive rate, detection rates for Down syndrome can be as high as 49% with a total hCG assay, 38% with free α-hCG testing, and 48% with free β-hCG [36].

The concentrations of hCG in maternal serum are markedly increased when fetal hydrops (generalized edema) and/or a cystic placenta is present. This is true not only for hydropic Down syndrome [37], but also for trisomy [38], Turner syndrome [39], and other causes of hydrops fetalis [40]. Although most cases of Down syndrome are not associated with hydrops, enlarged nuchal translucency and thickening is common and this has been attributed to fluid accumulation [41]. Elevated hCG may therefore be related to a disturbance in fluid homeostasis. However, the details of the regulation of hCG in maternal serum remain unknown [30]. It has been suggested that in Down syndrome pregnancies there is an increase in the nicking of hCG which results in a reduction in an inhibitory feedback mechanism that hCG has on its own production [42].

2.3. Unconjugated estriol

Isolated case reports and an early study by Jørgensen and Trolle [43] noted lower than normal levels of estriol in maternal urine when fetal Down syndrome was present. Subsequent analysis of second trimester maternal serum indicated that a reduction of unconjugated estriol (uE3) was also present and that this marker could also be used for Down syndrome screening [44,45]. Although there has been some controversy as to the value of this marker [46], the cumulative data from multiple studies indicated that uE3 is nearly as useful as hCG and is more powerful than MSAFP in distinguishing between affected and unaffected pregnancies (Table 1) [47].

Table 1  
Expected detection rates for second trimester tests when the false-positive rate is held at 5%

<table>
<thead>
<tr>
<th>Screening test</th>
<th>Detection rate (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age alone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 38 years at delivery</td>
<td>32</td>
<td>[136]</td>
</tr>
<tr>
<td>Maternal age, plus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSAFP</td>
<td>36</td>
<td>[36]</td>
</tr>
<tr>
<td>Total hCG</td>
<td>49</td>
<td>[36]</td>
</tr>
<tr>
<td>uE3</td>
<td>48</td>
<td>[36]</td>
</tr>
<tr>
<td>INH-A</td>
<td>45</td>
<td>[55]</td>
</tr>
<tr>
<td>MSAFP + hCG</td>
<td>63</td>
<td>[82]</td>
</tr>
<tr>
<td>MSAFP + uE3 + hCG</td>
<td>71</td>
<td>[82]</td>
</tr>
<tr>
<td>MSAFP + uE3 + hCG + INH-A</td>
<td>79</td>
<td>[82]</td>
</tr>
</tbody>
</table>

Detection rates for maternal age plus serum tests are based on pregnancies dated by ultrasound, with correction for maternal weight.
pregnancies, both uE3 and DHEAS appear to be lower than normal in the fetal liver, placental tissue, and maternal serum [48]. This indicates that Down syndrome pregnancies are characterized by a diminished supply of DHEAS. UE3 is thought to stimulate the endometrium and augment uterine blood flow [49]. However, this activity may not be too important because absence of uE3, as occurs in steroid sulfatase deficiency, does not appear to significantly affect pregnancy outcome [50]. The concentration of uE3 rises very rapidly during the second trimester and this analyte may therefore be particularly sensitive in identifying those pregnancies where the fetus is small or underdeveloped at the time of screening.

2.4. Inhibin-A

The use of inhibin as an additional marker for Down syndrome screening was first suggested by Van Lith et al. [51]. Inhibins are dimeric glycoproteins synthesized by gonads and placental tissue. There are two subunits, α and β, the latter existing in two forms, βA and βB, to form inhibin-A or inhibin-B [52]. It is the inhibin-A (INH-A) form that has been shown to have the greatest practical utility in Down syndrome screening. INH-A and hCG secretions appear to be interdependent [53], with increased production of INH-A by placental trophoblasts in pregnancies complicated by Down syndrome [54]. There is a moderately strong correlation between the maternal serum concentrations of hCG and INH-A in both affected and unaffected pregnancies [55]. Nevertheless, INH-A still provides good distinction between affected and unaffected pregnancies, alone, or in combination with other tests and this can include hCG (Tables 1 and 2).

Procedures for performing the INH-A immunoassay were initially complex [56], limiting its practical adaptation to a routine clinical chemistry setting. However, the availability of simplified procedures and commercially available kits has facilitated its introduction [57, 58]. In the second trimester, INH-A concentrations vary less with gestational age than those seen for many other markers making risks based on this test somewhat less susceptible to gestational age inaccuracy [59].

3. Multiple marker screening

3.1. Optimal second trimester screening

The serum markers described above have been combined to produce double (usually MSAFP and hCG) [60], triple (MSAFP, uE3, and hCG) [61], or quadruple (MSAFP, uE3, hCG, and INH-A) [55] tests. The mathematical methods used to develop these combinations are based on a multivariate Gaussian model [62, 63] that can readily accommodate additional tests. The analyte values do not necessarily have to be completely independent markers for Down syndrome; any correlation that exists between the variables can be entered into the risk algorithm. Other
mathematical models have been suggested but these have not been widely adopted [64].

Currently, second trimester serum screening is generally carried out at 15 to 18 weeks gestational age but the testing is considered to be valid up to 22 weeks. Reliable Down syndrome risks can also be generated using these same combinations of markers at 14 weeks but this is not widely advocated because at this gestation, the MSAFP test loses its discriminatory power to identify open neural tube defects [65].

Fig. 2 shows a typical testing pathway for patients choosing second trimester screening. Protocols and timing tend to vary in different locations. Additionally, the use of ultrasound examinations prior to screening, following a positive result, or as a follow-up procedure for women declining amniocentesis is highly variable (see Section 3.5).

Optimal screening requires accurate estimation of gestational age and efficacy is maximized when results are based on an ultrasound determination of gestational age [66,67]. Crown-rump measurement between 8 and 11 weeks provides the greatest accuracy. However, most laboratories will only recalculate the Down syndrome risk when there is a substantial (e.g. greater than 10 day) discrepancy between dating based on time from the last menstrual period and a post-test ultrasound dating because of the imprecision associated with the ultrasound measurement of gestational age. Table 3 illustrates the importance of a correct gestational age in developing patient-specific Down syndrome risks. It can be seen that individual patients’ risks need to be interpreted with considerable caution.

Adjustments are made to analyte concentrations to allow for some known factors that independently

<table>
<thead>
<tr>
<th>Weeks</th>
<th>Days</th>
<th>Risk 1:1</th>
<th>Risk 1:n</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>0</td>
<td>342</td>
<td>518</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>291</td>
<td>467</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>210</td>
<td>380</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>178</td>
<td>341</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>127</td>
<td>275</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>108</td>
<td>247</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>76</td>
<td>197</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>64</td>
<td>176</td>
</tr>
</tbody>
</table>

In Case 1, a patient aged 27.5 years has MSAFP = 0.72 MoM, uE3 = 0.72 MoM, hCG = 2.00 MoM (typical for an affected pregnancy) at 17 weeks and 4 days, has a computed Down syndrome risk of 1:127. The effect of adding or subtracting days on the risk can be seen. In Case 2, a patient aged 42.5 years with MSAFP = 1.00 MoM, uE3 = 1.00 MoM, hCG = 1.00 MoM (typical for an unaffected pregnancy) is considered. In each case, an error of 3–4 days can approximately double, or halve, the risk. Based on pregnancies dated by ultrasound, with correction for maternal weight.
affect observed levels. These include maternal weight [68,69], race/ethnicity [69,70], and diabetic status [71]. The presence of a twin pregnancy poses a special problem because prevalence of Down syndrome is less well established and there is uncertainty as to the expected concentrations when one, or both, fetuses are affected. At least in unaffected twin pregnancies, the observed concentrations of the maternal serum analytes are approximately twofold higher than that seen in unaffected singleton pregnancies. Adjusting the observed concentrations in twin pregnancies to that expected for a singleton pregnancy allows a “pseudo-risk” to be calculated. These “pseudo-risks” must be considered to be imprecise and the detection rate in twin pregnancies is expected to be lower than that for singleton pregnancies [72].

Other factors that appear to affect serum markers include cigarette smoking [73], parity [74], in vitro fertilization [75], intra-uterine insemination [76], analyte concentrations in a previous pregnancy [77], sex of the fetus [78], maternal rhesus blood type [79], and maternal systemic lupus erthematosus [80]. While some of these factors appear to have a relatively large effect, well-defined policies to control for these factors, generally, do not yet exist. It is often difficult for laboratories to have access to all aspects of the clinical data that might need to be entered into the screening algorithm. The improvement in screening that can be achieved by incorporating additional clinical factors therefore needs to be balanced against the practical realities involved with the data collection.

3.2. Expected performance

The triple test was first proposed by Wald et al. [61] in 1988 and, because of the existing availability of the component assays, was rapidly adopted into routine prenatal care. A 1995 United States survey indicated that approximately 63% of women were receiving multiple marker Down syndrome screening, usually the triple test [81]. The statistical parameters used in the algorithm to calculate risks have been updated [55,82,83] and methods have been described to customize these for individual screening programs [84].

Fig. 3. Receiver operating characteristic curves for the triple test (MSAFP, uE3, and hCG) and quadruple test (MSAFP, uE3, hCG, and INH-A) using a 1:270 second trimester risk for Down syndrome as a cut-off. Each point on the curve represents maternal age at delivery in 2-year intervals (13 –47). Rates were established by computer simulation using the statistical parameters in Refs. [16,23,36,55,82,136] for pregnancies dated by ultrasound.
Evaluating the performance of screening from the standpoint of the detection rate for a given fixed false-positive rate (e.g. 5% as in Table 1) is useful for comparing tests and combinations of tests. In practice, screening is usually carried out using a predetermined risk threshold, typically 1:270 in the second trimester. Fig. 3 shows the receiver operating characteristic curve (plots of detection rate against false-positive rate) for the triple and quadruple tests. These data are based on computer simulations using the statistical parameters in Refs. [16,23,36,55,82] for pregnancies dated by ultrasound, with maternal weight correction. Table 2 provides some estimates for the net effect of the various screening protocols applied to a population of women with the maternal ages seen in the United States in 1999.

Even when programs use the same cut-off to define a positive test result, individual screening programs can still expect to see different overall detection and false-positive rates, depending on the maternal age distribution of the women served and extent of the use of ultrasound to date pregnancies.

3.3. Observed performance

Numerous studies have verified that the second trimester triple test does perform as predicted [85]. Verifying detection rates in prospective trials requires pregnancy outcome information for both screen-positive and screen-negative cases with adjustment to account for the fact that a proportion of affected pregnancies with screen-negative results will spontaneously abort and may not come to the attention due to absence of cytogenetic testing [86,87]. A similar adjustment may also be required for screen-positive women who elect not to receive amniocentesis.

There appears to be no significant difference in the incidence of major anomalies in Down syndrome cases identified through the triple test versus those with screen-negative results [88,89]. With the exception of rare cases of Down syndrome in which fetal hydrops is present [37], there is little evidence that screening preferentially identifies those Down syndrome pregnancies most likely to experience fetal death [89].

Although the INH-A test has been available for some time, there are little prospective data available for the performance of the quadruple test [90]. Incremental gains in detection when comparing triple and quadruple testing are difficult to demonstrate without very large studies and, because of the increasing availability of first trimester testing, the quadruple test is often offered to a nonrandom population.

Programs with large pregnancy follow-up databases can evaluate screening performance by grouping cases according to ranges of reported risk and comparing observed prevalence of Down syndrome with that expected in each group [91]. Such studies (which are equivalent to using different cut-offs) establish that risks, when considered as groups, are accurate. However, it is not possible to prove accuracy for an individual patient’s reported numerical risk; each pregnancy is either affected or it is not.

3.4. Other disorders identified

The triple and quadruple tests are effective in identifying fetal trisomy 18 (Edwards’ syndrome), the second most common autosomal trisomy (trisomy 21 being most common). These pregnancies are characterized by low concentrations of MSAFP, uE3, hCG, and INH-A [92,93]. An initial protocol that identified high-risk pregnancies on the basis of analyte levels falling below fixed cut-offs [92] has been replaced by an approach that calculates patient-specific risk using a multivariate normal model similar to that employed in Down syndrome screening [94,95]. This risk-based approach has been shown to be superior to the fixed cut-off method [96,97]. Detection rates for trisomy 18 are comparable to those achieved in Down syndrome screening but with only a modest incremental rise in the total (trisomy 18 plus trisomy 21) false-positive rate [98].

Trisomy 13, the next most common autosomal trisomy, does not appear to be preferentially identified through the triple and quadruple tests [99] although those cases with open neural tube defects may show elevated MSAFP. There are a number of other aneuploidies that are preferentially identified but formal screening protocols to specifically identify these anomalies are not used. Essentially all cases of triploidy [38], an unknown proportion of cases with trisomy 16 mosaicism [100] and possibly trisomy 20 mosaicism [101], may be identified.

Among the sex chromosome abnormalities, Turner syndrome with fetal hydrops generally shows low
MSAFP and uE3 but elevated hCG and INH-A, while all four markers are often present at lower than normal concentrations in non-hydric cases [39,93]. Other sex chromosome abnormalities, many of which are associated with minor clinical significance, do not appear to be preferentially identified [102].

The association between fetal hydrops and elevated hCG extends to nonimmune hydrops with etiologies other than chromosomal [40]. These disorders are therefore likely to be preferentially identified through screening. Genetic disorders such as steroid sulfatase deficiency [50], Smith–Lemli–Opitz syndrome [103], and others associated with impaired uE3 synthesis can also be identified through these screening tests.

A very low level of all analytes is usually indicative of fetal death [50]. Given that analyte values are strongly dependent on developmental stage and normal placental function, the finding of anomalous values seems likely if there is fetal growth restriction or abnormal placental function. Low MSAFP (and also unexplained elevation of MSAFP) has been associated with fetal loss [104], low uE3 with fetal growth restriction [105], elevated hCG with diverse pregnancy complications and poor outcomes [106,107], and elevated INH-A with pre-eclampsia [108]. It is reasonable to hypothesize that false-positives in Down syndrome screening may therefore be preferentially associated with a broad range of abnormalities and poor pregnancy outcomes [109]. However, thus far, these associations do not provide a basis for altered pregnancy management [110].

3.5. Combining second trimester ultrasound and serum screening

An ultrasound examination is commonly performed for patients with maternal serum screen-positive results. This ultrasound, minimally, may be used to correct a major error in gestational age that may have been sufficient to explain the screen-positive result. Second trimester ultrasound may identify specific anatomic anomalies and/or “markers” that have been associated with Down syndrome. These markers include increased nuchal fold thickness, short femur and humerus, echogenic cardiac foci, renal pyelectasia, echogenic bowel and presence of choroid plexus cysts [111,112].

In most studies, detection rates and false-positive rates have been established on the basis of the presence or absence of these markers rather than treating fetal measurements as continuous variables. Applied to serum screen-positive patients, the ultrasound identification of a marker will increase the risk presented to a patient and potentially increase the chance that she will accept amniocentesis. Absence of these markers can reduce the false-positive rates but may result in true-positives being missed. Using currently available likelihood ratios, failure to visualize an ultrasound marker only reduces risk by 50–70% [113].

Modification of risk using ultrasound-derived likelihood ratios that reflect the presence or absence of specific markers needs to be approached cautiously. In Down syndrome fetuses, the presence of more than one marker occurs more often than expected by chance [113], and therefore, likelihood ratios for each marker cannot be treated as independent factors. Additionally, biochemical tests and ultrasound findings are not necessarily independent [114]. In their meta-analysis, Smith-Bindman et al. [112] concluded that none of these ultrasound markers alone is sufficient to be clinically useful.

Wald et al. [47] have expressed the opinion that modifying a positive second trimester maternal serum screening result by ultrasound should be avoided because true-positives will be missed. However, the policy also needs to be viewed in the context of the choice of cut-off selected for serum screening and the fact that an ultrasound examination will be an integral component of the management of screen-positive patients. Use of ultrasound data to modify risk might, with an appropriate risk cut-off, result in a substantial reduction in the number of amniocenteses performed with only a small reduction in the detection rate [115].

Because nuchal fold thickness and proximal long bone measurements can be treated as continuous variables, it is possible to combine these markers into a multivariate Gaussian marker model that can include maternal serum analytes. Based on one relatively small study, improved efficacy of screening could be demonstrated (relative to the usual triple test) when humerus length was substituted for uE3 [116]. Recently, a protocol that combined the quadruple test with nuchal thickness and long bone
measurements was developed [117]. This provisional study indicated combined second trimester screening might achieve an approximately 90% detection rate at the 5% false-positive standard.

Currently, the use of second trimester ultrasound to modify risk for Down syndrome is controversial. Longer term, there is considerable potential for second trimester ultrasound biometry to help identify fetal Down syndrome. Given the level of detail available in the second trimester relative to that available in the first trimester, it seems reasonable to expect that Down syndrome identification through a second trimester scan may ultimately prove to be even more effective than that currently achievable in the first trimester [118]. Further blending of second trimester serum analyte testing and biometry can therefore be expected.

### 4. Laboratory issues

MSAFP, uE3, hCG, and INH-A serum tests are available in immunoassay formats suitable for typical clinical chemistry laboratories. There appear to be no major problems with analyte stability that significantly complicate screening. Testing should be carried out on fresh (not frozen) serum specimens within 7 days of collection.

There are some aspects of the quality control of testing that require special consideration. Reference values for each serum analyte need to be determined. A representative set of maternal serum specimens from pregnant women should receive the analyses using test formats that will reflect the protocol used in the subsequent screening. Concentrations for all four analytes are dependent on gestational age and there-

---

### Table 4

Maternal age-specific detection rates (DR), false-positive rates (FPR), minimum likelihood ratios needed for a positive result (min LR), and odds of being affected given a positive result (OAPR) for the triple and quadruple tests

<table>
<thead>
<tr>
<th>Maternal age</th>
<th>Second trimester risk (1:270)</th>
<th>min LR</th>
<th>Quadruple DR (%)</th>
<th>FPR (%)</th>
<th>OAPR (1:n)</th>
<th>Triple DR (%)</th>
<th>FPR (%)</th>
<th>OAPR (1:n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>1289</td>
<td>4.77</td>
<td>66.0</td>
<td>3.0</td>
<td>58</td>
<td>56.9</td>
<td>3.1</td>
<td>70</td>
</tr>
<tr>
<td>15</td>
<td>1282</td>
<td>4.75</td>
<td>66.0</td>
<td>3.0</td>
<td>59</td>
<td>57.1</td>
<td>3.1</td>
<td>70</td>
</tr>
<tr>
<td>17</td>
<td>1271</td>
<td>4.71</td>
<td>66.2</td>
<td>3.0</td>
<td>58</td>
<td>57.3</td>
<td>3.2</td>
<td>70</td>
</tr>
<tr>
<td>19</td>
<td>1251</td>
<td>4.63</td>
<td>66.4</td>
<td>3.1</td>
<td>58</td>
<td>57.5</td>
<td>3.2</td>
<td>70</td>
</tr>
<tr>
<td>21</td>
<td>1218</td>
<td>4.51</td>
<td>66.8</td>
<td>3.2</td>
<td>58</td>
<td>57.9</td>
<td>3.3</td>
<td>70</td>
</tr>
<tr>
<td>23</td>
<td>1163</td>
<td>4.31</td>
<td>67.5</td>
<td>3.3</td>
<td>57</td>
<td>58.7</td>
<td>3.5</td>
<td>69</td>
</tr>
<tr>
<td>25</td>
<td>1077</td>
<td>3.99</td>
<td>68.6</td>
<td>3.6</td>
<td>57</td>
<td>60.0</td>
<td>3.8</td>
<td>69</td>
</tr>
<tr>
<td>27</td>
<td>953</td>
<td>3.53</td>
<td>70.4</td>
<td>4.1</td>
<td>55</td>
<td>62.2</td>
<td>4.4</td>
<td>67</td>
</tr>
<tr>
<td>29</td>
<td>792</td>
<td>2.93</td>
<td>72.9</td>
<td>4.9</td>
<td>53</td>
<td>65.3</td>
<td>5.4</td>
<td>65</td>
</tr>
<tr>
<td>31</td>
<td>610</td>
<td>2.26</td>
<td>76.5</td>
<td>6.2</td>
<td>49</td>
<td>69.4</td>
<td>7.2</td>
<td>63</td>
</tr>
<tr>
<td>33</td>
<td>434</td>
<td>1.61</td>
<td>80.7</td>
<td>8.4</td>
<td>45</td>
<td>74.8</td>
<td>10.0</td>
<td>58</td>
</tr>
<tr>
<td>35</td>
<td>288</td>
<td>1.07</td>
<td>85.4</td>
<td>12.0</td>
<td>40</td>
<td>80.8</td>
<td>14.4</td>
<td>51</td>
</tr>
<tr>
<td>37</td>
<td>180</td>
<td>0.67</td>
<td>89.6</td>
<td>17.1</td>
<td>34</td>
<td>86.6</td>
<td>21.2</td>
<td>44</td>
</tr>
<tr>
<td>39</td>
<td>109</td>
<td>0.40</td>
<td>93.1</td>
<td>23.8</td>
<td>28</td>
<td>91.6</td>
<td>30.5</td>
<td>36</td>
</tr>
<tr>
<td>41</td>
<td>64</td>
<td>0.24</td>
<td>95.8</td>
<td>32.2</td>
<td>21</td>
<td>95.2</td>
<td>42.3</td>
<td>28</td>
</tr>
<tr>
<td>43</td>
<td>37</td>
<td>0.14</td>
<td>97.6</td>
<td>41.8</td>
<td>15</td>
<td>97.5</td>
<td>55.0</td>
<td>20</td>
</tr>
<tr>
<td>45</td>
<td>21</td>
<td>0.08</td>
<td>98.8</td>
<td>52.2</td>
<td>11</td>
<td>98.9</td>
<td>67.2</td>
<td>14</td>
</tr>
<tr>
<td>47</td>
<td>12</td>
<td>0.04</td>
<td>99.3</td>
<td>62.1</td>
<td>7</td>
<td>99.5</td>
<td>77.3</td>
<td>9</td>
</tr>
</tbody>
</table>

Based on 1:270 second trimester cut-off, for pregnancies dated by ultrasound. Rates were established by computer simulation using the statistical parameters in Refs. [16,23,36,55,82,136].

Screening programs using different cut-offs can expect different detection rates and false-positive rates but should know their program’s expected screen-positive rates. An estimate can be made from this data. Using the alternative cut-off, calculate at each maternal age the minimum likelihood ratio (min LR) needed for a positive result. Then find the detection rates and false-positive rates that correspond to the revised set of min LR values. For example, consider a program performing the triple test and using a 1:250 second trimester cut-off. For women aged 25, the second trimester age-specific risk for Down syndrome is approximately 1:1077 and the minimum likelihood ratio now needed for a screen-positive result is 1077/250 = 4.31 (post-test risk is = 4.31:1077 or = 1:250). From the above table, this min LR value of 4.31 corresponds to a triple test detection rate of approximately 58.7% and false-positive rate of 3.5% (equivalent to women aged 23 with the 1:270 cut-off). By interpolating, at each maternal age, the alternative set of detection rates and false-positive rates can be similarly estimated and the net rates can be calculated for the overall population by calculating the average, weighted for the number of women screened at each age.
fore samples that substantially cover 15 to 22 weeks of gestation are needed. Sufficient numbers of samples (up to 100 for each week) are required to obtain reliable estimates of the weekly median values. Once the weekly medians have been determined, regression analyses, weighted for the number of samples at each week, are carried out to derive median values that are appropriate for each day of gestation. The models used to relate median with gestation are log-linear for MSAFP [119] and uE3 [45], exponential for hCG [61], and log-quadratic for INH-A [59].

Prior to initiating a maternal serum screening program, cut-offs and other policies that will be used in the program need to be defined. Software is commercially available for risk calculations. The accuracy of the risks generated by these packages should be verified for various combinations of test and clinical situations. Verification of the accuracy of software is particularly important following the release of significant upgrades in these computer programs.

Because all serum analytic values will be expressed as multiples of the normal median values, stability of the assays with minimal drift is particularly important. Elements of quality control that help ensure reliable assays include: validation of new kits prior to use in screening, inclusion of pooled control samples (high medium and low concentrations) in every run, checks for within-run assay drift, participation in proficiency testing, and periodic review of analytic values. The importance of closely monitoring observed median MoMs and screen-positive rates has been stressed [120]. Screen-positive rates are highly dependent on the maternal ages of the populations screened, and using weighted averages of the detection rates and false-positive rates may be helpful in establishing individual program’s expected performance (Table 4).

Methods have been described to analyze the effect of modest levels of analytical imprecision on risk [121]. Relatively small coefficients of variation in analytes become compounded when risk figures are computed. For any particular level of imprecision in the analytical test values, the imprecision in the associated Down syndrome risks is not the same for all screened. In particular, imprecision should be of particular concern when women of advanced maternal age receive screening. Table 5 illustrates the impact of analytical imprecision on the risk figures that may be presented to women receiving the triple test. Down syndrome screening laboratories should select assays that are highly reproducible and pay close attention to both drift and precision.

The Foundation for Blood Research provides a useful guide for laboratories that contains detailed information on establishing tests, selecting appropriate policies for screening, and quality control [122].

### Table 5
Illustrations of the effect of analytical precision on the Down syndrome risk figures reported to patients receiving the triple test (MSAFP, uE3, hCG)

<table>
<thead>
<tr>
<th>Test CV (%)</th>
<th>Case 1: Age 27.5, MSAFP = 0.72 MoM, uE3 = 0.72 MoM, hCG = 2.00 MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean risk 1:n</td>
</tr>
<tr>
<td>3</td>
<td>127</td>
</tr>
<tr>
<td>5</td>
<td>127</td>
</tr>
<tr>
<td>7</td>
<td>127</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test CV (%)</th>
<th>Case 2: Age 42.5, MSAFP = 1.00 MoM, uE3 = 1.00 MoM, hCG = 1.00 MoM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean risk 1:n</td>
</tr>
<tr>
<td>3</td>
<td>275</td>
</tr>
<tr>
<td>5</td>
<td>275</td>
</tr>
<tr>
<td>7</td>
<td>275</td>
</tr>
</tbody>
</table>

When the individual tests (MSAFP, uE3, hCG) are associated with a large coefficient of variation (CV), the range of risks that may be reported by a laboratory is large. For example, in Case 1 (with analytic values typical for Down syndrome), the mean risk that would be reported is 1:127. However, when each test is associated with a CV of 7%, the risk reported may depart substantially from 1:127 with 95% of the values falling in the range 1:67 to 1:1222. Case 2 demonstrates variability in risk for an advanced maternal age patient with analytic values typical for an unaffected pregnancy. Results are based on pregnancies dated by ultrasound with weight correction.

5. Other second trimester approaches

5.1. Serum tests

Other biochemical markers have been noted to distinguish between Down syndrome affected and unaffected pregnancies, but these are not in widespread practical use. Some have relatively poor separation of affected and unaffected pregnancies, require standardization or complex protocols, have not been extensively evaluated, or may be otherwise uneconomically feasible. Some of the proposed markers have been reviewed elsewhere [47]. Potentially useful tests
include neutrophil alkaline phosphatase [123], α-Inhibin [51] schwangerschaftsprotein 1 (SP1, also known as pregnancy-specific β1 glycoprotein) [124], the proform of eosinophil major basic protein (proMBP) [125], placental isoferritin p43 component [126], and hyperglycosolated hCG (also known as Invasive Trophoblastic Antigen or ITA) [127]. Analysis of the sub-types of MSAFP that bind to lectins (AFP microheterogeneity assay) may also be useful [128].

5.2. Urine tests

Screening tests using maternal urine have also been proposed. Initial studies indicated that urinary β-core hCG concentrations were substantially elevated in affected pregnancies. More recent data has suggested that the power of this marker is not as high as hoped and there are difficulties with sample stability or reproducibility [129]. Urinary total hCG, free β-hCG, and total (conjugated plus unconjugated) estriol appear to show differences in affected, relative to unaffected, pregnancies [130] and there is also optimism that hyperglycosolated hCG in urine will be clinically useful [131].

5.3. Fetal cell sorting

Considerable efforts have been made to try to isolate fetal cells present in maternal circulation [132]. Fetal lymphocytes, granulocytes, nucleated red blood cells, and trophoblast cells are present but in very low concentrations. An additional complication is the long-term persistence of fetal progenitor cells that may complicate diagnosis of a current pregnancy with information pertaining to previous pregnancies [133]. There is also the ability to detect fetal DNA in maternal circulation [134]. If these technologies could be perfected, the techniques would have some obvious potential advantages in the diagnosis of many genetic disorders. However, the current status of this research suggests that a clinical screening or diagnostic test for aneuploidy is not imminent [135].

Acknowledgements

This paper was supported by a grant from the Patrick and Catherine Weldon Donaghue Medical Research Foundation. I thank JFX Egan for helpful suggestions.

References


Ozturk M, Bellet D, Manil L, Hennen G, Frydman R, Wands J. Physiological studies on human chorionic gonadotropin (hCG), αhCG, and βhCG as measured by specific monoclo-


[71] Sandeen U, Bartels I. Preliminary data on an association between blood groups and serum markers used for the so-called “triple screening”: free oestriol MoM values are decreased in rhesus-negative (Rh−) women. Prenat Diagn 2001;21:194–5.


