The triple test as a screening technique for Down syndrome: reliability and relevance

Abstract: The triple test is a second trimester screening test used to identify those pregnant women who should be offered a diagnostic test to identify whether their fetus has an aneuploidy. It was first described in 1988, but has largely been superseded by newer tests either conducted earlier in the first trimester (i.e., the combined test, using ultrasound measurement of nuchal translucency, pregnancy-associated plasma protein A, and human chorionic gonadotrophin [hCG]) or in the second trimester (i.e., the quadruple test, using α-fetoprotein, hCG, uE3, and inhibin). These newer tests have been introduced because they offer greater detection and lower screen positive results thereby enhancing diagnosis rates, while decreasing the risk of iatrogenic harm caused by the invasive testing required when collecting suitable sample tissue. Noninvasive alternatives to the triple test have been identified, but these have not been adopted despite 13 years of development. It is likely, therefore, that the triple test (or variants thereof) will continue to be used in routine antenatal care for the foreseeable future.

Keywords: pregnancy, screening test, antenatal, Down syndrome

Introduction
The triple test is one of a range of screening tests that are used to identify pregnant women whose fetus is likely to be affected by trisomy 21 (Down syndrome) and who should then be offered a diagnostic test. All of the tests similar to the triple test are based on the same mathematical principle (Bayes theorem) and work by combining a prior probability derived from maternal age at expected date of delivery with a likelihood ratio usually based on two multivariate Gaussian distribution functions. This combination results in a reasonably accurate risk estimate of the probability that the fetus has Down syndrome. Women whose risk exceeds a specified cutoff are then offered a diagnostic test (i.e., amniocentesis or chorionic villus biopsy), which allows a cytogenetic diagnosis to be determined. This may be done either by cell culture and karyotyping or by fluorescent in situ hybridization (FISH).

The triple test is used only in the second trimester of pregnancy and now has a range of competitors (Table 1). As one of the first entrants into the serum screening arena, it is therefore legitimate to question whether it remains relevant more than 20 years after it was developed. There are a number of factors that affect the decision about which screening test to use: screening test effectiveness; cost-effectiveness; cost-benefit/cost-hazard.

How reliable are screening protocols?
When screening, reliability is measured by assessing the effectiveness of different screening protocols by measuring the detection rate and corresponding screen...
positive rate. To allow comparison, it is usual to fix the screen positive rate and assess the detection rate using a computer model rather than gleaning this directly from patient data. Different studies generate different basic data sets and modeling can generate controversy over the value of different protocols. The original description of the triple test estimated that adding unconjugated estriol to a double test increased the detection rate from 55% to approximately 60% for a 5% screen positive rate. A later estimate claimed that for ultrasound-dated pregnancies, the double test had a detection rate of 58% and the triple test 67% for a screen positive rate of 5% and suggested that without ultrasound there was only a 4% difference in detection rates between the two tests. In the early days of Down syndrome screening, it was often felt by laboratory managers that the slight increment in detection was not worth the extra reagent and staff costs, leading them to opt for the double test.

Table 2 shows estimates of detection rates at specific screen positive rates for different screening strategies and makes it clear that the triple test is now outclassed by other test variations.

### Logical choices and consumer behavior

The triple test is, thus, no longer the most effective screening test for antenatal Down syndrome and consequently many national guidelines recommend other screening tests instead. In practice, however, it is still in common use, at least in the United Kingdom (UK). There must, therefore, be other factors that influence choice of screening tests to use.

Over the last 10 years in the UK, the triple test was the test routinely offered. It would be logical to expect there would be a move to the quadruple test because this would allow improved detection and lower screen positive rates without the need to redesign the way in which patient services were provided. This did not occur, partly because the only commercially available assay for inhibin-A (the fourth analyte in the quadruple test) was not suitable for use in a routine laboratory because it was insufficiently stable and the intrabatch assay variation was excessive (coefficient of variation [CV], 17%). This lead to an excessively high screen positive rate when compared with the computer simulation models of quadruple screening. Consequently, although superior in a research setting, the quadruple test was not practical for use in a routine laboratory. More recently, the inhibin assay has been automated, leading to substantial improvement in performance. In the UK, this has not resulted in wide uptake of the extra test.

Another reason for the reluctance to add extra tests is the law of diminishing returns meaning as each extra analyte is added to the basic double test the incremental improvement in detection rate is less. Furthermore, there is a tendency for the newer tests to be more expensive so the cost-benefit equation becomes harder to justify. In the UK National Health Service (NHS) centrally-funded health service model, it is difficult to persuade the commissioners to pay yet more for a tiny improvement in a screening program when there are pressing health needs that must be addressed elsewhere.

Yet another factor was consumer behavior. In the UK, the standard service provided by the NHS was the second trimester triple test. One of my roles is the Director of Prenatal Screening for the South Yorkshire Sub-Regional Down’s Screening Programme run by the laboratory of the Northern General Hospital, Sheffield. It was obvious from the pattern of cytogenetic reports in our regular audit meetings that a significant proportion of patients were exercising their consumer choice and paying for private first trimester

### Table 1: Strategies for antenatal Down syndrome screening

<table>
<thead>
<tr>
<th>Test name</th>
<th>Used in</th>
<th>Analytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double test</td>
<td>Second trimester</td>
<td>AFP + hCG (total or free-β)</td>
</tr>
<tr>
<td>Triple test</td>
<td>Second trimester</td>
<td>As double test + unconjugated estriol</td>
</tr>
<tr>
<td>Quadruple test</td>
<td>Second trimester</td>
<td>As triple test (using free-β hCG) + inhibin-A</td>
</tr>
<tr>
<td>Combined test</td>
<td>First trimester</td>
<td>Ultrasound measurement of NT + PAPP-A + free-β hCG</td>
</tr>
<tr>
<td>Serum integrated test</td>
<td>Both first and second trimester</td>
<td>PAPP-A (first trimester) + quadruple test (or triple test)</td>
</tr>
<tr>
<td>Integrated test</td>
<td>Both first and second trimester</td>
<td>As serum integrated test + NT in first trimester</td>
</tr>
<tr>
<td>Contingent test</td>
<td>Both first and second trimester</td>
<td>Dependent on structure of contingent screen chosen</td>
</tr>
</tbody>
</table>

**Abbreviations:** AFP, α-fetoprotein; hCG, human chorionic gonadotropin; NT, nuchal translucency; PAPP-A, pregnancy-associated plasma protein A.

### Table 2: Detection rates for different antenatal Down syndrome screening strategies from the SURUSS study

<table>
<thead>
<tr>
<th>Test name</th>
<th>DR @ SPR = 1%</th>
<th>DR @ SPR = 3%</th>
<th>DR @ SPR = 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double test (using free-β hCG)</td>
<td>46%</td>
<td>63%</td>
<td>71%</td>
</tr>
<tr>
<td>Triple test (using free-β hCG)</td>
<td>56%</td>
<td>70%</td>
<td>77%</td>
</tr>
<tr>
<td>Quadruple test (using free-β hCG)</td>
<td>66%</td>
<td>79%</td>
<td>84%</td>
</tr>
<tr>
<td>Combined test</td>
<td>66%</td>
<td>78%</td>
<td>83%</td>
</tr>
<tr>
<td>Serum integrated test</td>
<td>77%</td>
<td>86%</td>
<td>90%</td>
</tr>
<tr>
<td>Integrated test</td>
<td>84%</td>
<td>91%</td>
<td>93%</td>
</tr>
</tbody>
</table>

**Abbreviations:** SURUSS, serum urine and ultrasound screening study; DR, detection rate; SPR, screen positive rate.
screening because they wanted to know the result earlier than would be possible under the NHS scheme. This trend is reversing, as first trimester screening is now provided in the Sheffield NHS screening program. Since patients wanted earlier screening, it was clear that first trimester screening had to be made available with all of the consequent changes to the antenatal care package that this introduction process entailed.

Clearly, women did not choose the first trimester test on the basis of effectiveness, cost-effectiveness, and cost-benefit/hazard ratios. Rather, an ultrasound test is far more personally interactive and, in addition to allowing the first view of the baby, gives an immediate answer as to whether the baby has any major problems. It is, therefore, entirely understandable that first trimester testing was popular with patients.

A further question that could be raised is why use the obviously less effective first trimester combined test, when the integrated test appears to be far more effective? Here, there were good practical reasons why the integrated test was not introduced. For instance, the integrated test uses information collected in both trimesters of pregnancy and requires a wait until all of that data has been collected before calculating the risk estimate. This means there is the need to ensure that women attend on more than one occasion for the screening test to be performed, increasing the risk of dropout when appointments are missed. Furthermore, there is a delay during which information that may allow an early diagnostic test to be carried out is withheld. This delay has been criticized as being ethically unacceptable.7

An alternative method that was suggested to avoid the ethical objections to the integrated test was the contingent screen. In this test, the first trimester screening results were revealed and those at very low risk were excluded from the next stage while those at very high risk were offered a diagnostic test. Those in the middle were offered a second stage test in the second trimester and, depending on the result, were offered a diagnostic test. A trial of the acceptability and effectiveness of a contingent screen found that 16.7% of women booked too late to be offered the first trimester stage of the process, but the majority of women entering the screening process completed it, thereby proving that fears of high dropout rates were unfounded. Most women were happy to be offered contingent screening.8

International perspective
Internationally, Down syndrome screening practices vary. In the UK, standards have been imposed on all laboratories carrying out Down syndrome screening by the National Health Service Fetal Anomaly Screening Programme (NHS FASP) which specify that “all Down’s syndrome screening programmes must meet a target detection rate of greater than 75%, for a false positive rate of less than 3%, by 2007”. The FASP model of best practice9 recommends that this should be achieved using the combined test, the integrated test, or the serum integrated test for women who present in the first trimester, and the quadruple test for women who present in the second trimester. The FASP annual report for 2006–2008 shows that in 2007, the predominant screening strategy was the triple test with a significant number of centers using combined testing and a smaller proportion providing quadruple testing. The most recent report10 indicates that 113 of the 152 Primary Care Trusts (the fund-holding units of the NHS) in England will have changed to the combined screen by April 2010 and the remainder do not yet have firm plans to change. This does not, however, mean that these units have abandoned the triple screen since they need to have an alternative for patients who book too late for first trimester testing.

In the United States (US), it was only in 2007 that the American College of Gynecologists recommended that all women should be eligible for screening regardless of maternal age, whereas previously only women over the age of 35 were automatically offered genetic counseling and amniocentesis or chorionic villus sampling.12 A recent survey of US obstetricians showed that 95% now offered Down syndrome screening to all patients, with 70% offering first trimester screening and 86% offering the quadruple screen.13

In Australia, the combined test is recommended for the first trimester and the quadruple test for the second trimester.14 Similarly, in Canada the Society of Obstetricians and Gynecologists recommend this protocol.15

Is the triple test relevant?
It is clear that the triple test is now becoming increasingly irrelevant as a clinical test because other variants on the test have been mandated by national quality standards aiming to reduce the number of diagnostic tests required and to further decrease the iatrogenic risk to unaffected pregnancies.

The future
Ethical challenges
While of decreasing clinical relevance, the triple test still has great importance for the future. Antenatal screening for Down syndrome was one of the first examples of mass population testing to prevent a genetic disorder. The test was introduced as an extension of earlier neural tube defect
(NTD) screening programs which were designed to detect conditions that caused significant disability. Consequently, Down syndrome screening was introduced as a clinical service with research into test improvement being carried out on routinely collected data. This meant that the research was generally carried out without any review by a research ethics committee. In 2003, a survey of research ethics committees in the UK found that if ethical approval had been sought to allow research into Down syndrome screening, it is likely it would have been refused. In general, these ethics committees felt that screening for a condition that caused some learning deficiencies and minor reduction in life span was no more acceptable than screening to prevent children suffering the “socially embarrassing physical characteristics of red hair and freckles”.

Ethics can be a difficult area in which to research because there is no ‘correct’ answer to any question. Surveys of individuals have revealed there is more reluctance to consider termination of a fetus because it has Down syndrome than if it were affected by spina bifida or hemophilia. Surveys of physicians have demonstrated greater reluctance to terminate hemophilia-than Downs syndrome-affected fetuses and opinions of Anglophone and Francophone physicians differ. A survey of Lutheran pastors demonstrated that only 23% considered Down syndrome to be a sufficiently serious condition to warrant termination. The situation is further complicated by the fact that pregnant women regard the prospect of a Down syndrome-affected birth to be more burdensome than a procedure-related miscarriage and by studies that show that women’s views about screening are affected by available resources, their own feelings about having a child with Down syndrome, their moral beliefs, family and social influences, perceptions of their own health, and any difficulty in becoming pregnant.

Opinions also vary between countries: 33% of respondents in a survey in Russia indicated that they favored compulsory termination of pregnancy if testing identified a genetic disorder in the fetus and Russians were more in favor of prenatal screening, selective termination, and genetic manipulation to improve a child’s intelligence or reduce the probability of homosexuality. Thus, 13 years after this method for screening was developed, no routine application for aneuploidy detection has yet been described. This should be contrasted with the triple test which was first described in 1988. The first routine screening program in the UK NHS was introduced in 1990 and its effectiveness during its first full year of operation (Feb 1990–1991) was reported in 1993. Therefore, it only took two years of research indicating that antenatal serum screening for Down syndrome could be effective before the technique was.

**Alternative screening tests**

**Fetal cells in maternal blood**

Many years before the introduction of Down syndrome screening, it was discovered that fetal cells could be identified in the maternal circulation and used to identify fetal gender. After the introduction of Down syndrome screening, a great amount of effort was made to develop extraction methods that would allow fetal cells to be purified from maternal blood, which would allow noninvasive prenatal diagnosis. As a result, dozens of research teams worked on projects to extract erythroblasts, leukocytes, trophoblasts, etc. Unfortunately, while fetal cells have been successfully extracted in a research setting, this process has not been introduced into routine practice.

**Cell-free DNA in maternal serum**

In 1997, the presence of cell-free fetal DNA in maternal serum was identified. This was thought to offer another prospect for noninvasive prenatal testing. Thirteen years later, methods for antenatal rhesus typing and fetal gender have been described. A routine application for RhD typing has been tested and proven to be effective. Experimental methods for using this technology for Down syndrome testing have been reported and include: single nucleotide polymorphism allelic ratios; circulating placental messenger ribonucleic acid (mRNA) analysis; and epigenetic-genetic chromosome-dosage. It is worth noting that these experimental methods are mostly all described by the same research group, however. Thus, 13 years after this method for screening was developed, no routine application for aneuploidy detection has yet been described. This should be contrasted with the triple test which was first described in 1988. The first routine screening program in the UK NHS was introduced in 1990 and its effectiveness during its first full year of operation (Feb 1990–1991) was reported in 1993. Therefore, it only took two years of research indicating that antenatal serum screening for Down syndrome could be effective before the technique was.
introduced as a routine screening test. Cell-free DNA was identified 13 years ago and, although several techniques to allow it to be used for diagnosis have been described, none have been taken into routine use. We cannot, therefore, predict whether serum DNA will ever become a routine test: To become accepted, it has to supplant already established tests and the technical difficulties associated with amplification-based testing (which makes sample purity/ lack of contamination vital) must be overcome. Finally, there are ethical implications of DNA and RNA testing which must be considered.26

Conclusions

The triple test was first described in 19883 and rapidly entered routine use as an antenatal screen for Down syndrome.35 As the years have progressed, it has been superseded by newer variations (eg, first trimester combined testing and the quadruple test), which have been recommended in national guidelines, not because the triple test has been unreliable, rather because the other tests have proven more effective in terms of greater detection with lower screen positive rates.

The triple test remains relevant because it is the foundation upon which current antenatal screening tests for Down syndrome are rooted. It is also important because of the ethical dilemmas it creates. Where does the limit of acceptability for screening lie? Most countries currently allow their citizens to have reproductive autonomy, but the introduction of Down syndrome screening has lowered the barriers to other forms of genetic screening tests. It becomes possible that reproductive autonomy may be threatened by economic factors that may favor compulsory screening. Thus, Down syndrome screening will continue to be controversial for many years to come.

Disclosure

The author reports no conflicts of interest in this work.

References