Chapter 5 First trimester ultrasound markers for trisomy 18

5.1 Introduction

Hundreds of published reports are available documenting the association between various ultrasound measurements and aneuploidy in the late first trimester of pregnancy. By far the most widely studied marker is the measurement of the width of the translucent space between the fetal spine and skin. This is referred to as the nuchal translucency, or NT, and its measurement is now widespread as part of routine prenatal care. Other ultrasound markers are considerably less well developed, with few actually having been introduced into routine practice outside of a few high risk centers. Many of the reports on these newer ultrasound markers (*e.g.,* the presence or absence of the nasal bone, or the shape of ductus venous blood flow on Doppler imaging), were based on cohorts that included women referred because of elevated fetal NT measurements or because higher NT measurements contributed to a high Down syndrome risk.

Many of the first trimester ultrasound studies are subject to important biases which can make interpretation of results difficult or even impossible. One of these is trimester of ascertainment. When this source of bias is present, it causes the detection rate to be overestimated, especially in demonstration studies. This is especially important, because few observational studies of NT have been published.

Another important bias that mainly affects the detection rate for NT measurements or other ultrasound markers is the proportion of ultrasound referral patients in a cohort. Several research groups active in this area are located in 'high risk' centers, where it may be difficult to distinguish between the women who have never been tested, and those who have been referred to that high risk center because an increased NT was identified in a primary care setting. For example, assume the median NT (expressed in MoM) was 2.0 for Down syndrome pregnancies with a corresponding 1.0 MoM in unaffected pregnancies. Also, assume that 2.0 MoM is about the 95th centile in unaffected pregnancies, and that women above this level would be routinely offered diagnostic testing. In a population undergoing primary ultrasound screening, the median MoM in the Down syndrome pregnancies would be expected to be 2.0 and 5% of women would have NT measurements above 2.0. However, if half of the women seen at a hypothetical high risk center were actually referrals (most having values above 2.0 MoM), then these observations will substantially increase the observed median NT MoM

in the Down syndrome pregnancies that are identified. This would result in an overestimate of both the observed and modeled detection rates.

A third potential bias occurs when the cohort of women studied is preselected by a factor that is correlated with the test of interest. For example, later in this chapter there will be data showing a clear correlation between measurements of NT and absence of the fetal nasal bone. As NT increases, it becomes more likely that the nasal bone will be absent. This effect occurs in both affected and unaffected pregnancies. Thus, a study of women with elevated NT measurements will overestimate the performance of nasal bone visualization. This problem does not occur, if the two factors are unrelated. For example, there is no correlation between NT and maternal age. Thus, a cohort of unscreened women over age 35 will be expected to have the same distribution of NT measurements in affected and unaffected pregnancies as a cohort of unscreened women under age 35.

Separate PubMed literature searches were performed for three selected first trimester ultrasound markers (NT, nasal bone, and ductus venosus). The aim for NT was to create distribution parameters suitable for multivariate modeling. For nasal bone and ductus venosus, the aim was to create univariate likelihood ratios for these categorical tests and determine whether they are independent of NT measurements. If not, it would be more difficult to include them in modeling, as the necessary between-marker correlations for trisomy 18 would be based on limited data. For other less commonly reported markers, a short informal summary was prepared, but parameters were not included, and these markers will not be part of future modeling. Methods for combining results from multiple studies are the same as those used in earlier chapters.

5.2 Nuchal translucency (NT) measurements

Introduction:

In a sagittal view of the fetus, a translucent space between the spine and skin can be seen in all fetuses in the late first trimester. A third echogenic line indicating the amnion should also be visualized, as care must be taken to ensure that the distance from the spine to the skin is measured, rather than the spine to the amnion. The NT measurement is different from the second trimester finding of nuchal skin fold thickness, in which the actual thickness of the skin is measured (Benacerraf *et al.*, 1985). In the original description of Down syndrome (Down, 1866), he describes the skin as being "deficient in elasticity, giving the appearance of being too large for the body". This is likely the observation of skin fold thickening. Nuchal translucency may be a precursor of this finding.

One of the first groups to describe the specific finding of nuchal translucency (NT), rather than the more classic findings of cystic hygroma (fluid filled sacs) or hydrops (generalized fluid accumulation) (Szabo and Gellen, 1990), found "accumulation of subcutaneous fluid in various amounts in the nuchal region" in all seven cases of Down syndrome identified through CVS, but in only one of 105 matched control pregnancies. This was later confirmed by several groups (Hewitt, 1993; Savoldelli *et al.*, 1993; Nicolaides *et al.*, 1992a). However, there were initial problems with some groups replicating the findings (Bewley *et al.*, 1995; Kornman *et al.*, 1996; Haddow *et al.*, 1998) leading to the development of specific training programs (Fetal Medicine Foundation in mid-1990s and the Nuchal Translucency Quality Review Program in 2005) that included formal coursework, testing, submission of sonographic images, and credentialing (www.fetalmedicine.com and www.ntqr.org). An innovative aspect in both of these existing programs is the use of external quality assessment to help identify sonographers who may not be performing as expected (Palomaki *et al.*, 2008; Cuckle, 2010).

In brief, a proper NT study requires that the sonographer obtain a true sagittal section with the image magnified to include only the head and thorax (Figure 5.2-1). The head should be in the neutral position, and the spine, skin and amnion should be visualized. The calipers are then placed in a specific manner (referred to as 'on-to-on') to obtain at least 3 measurements, with the largest of the three used for interpretation. Some groups have suggested using the median of the three observations. For long-term assessment of performance, individual NT results are collected and epidemiological monitoring is

carried out to determine the weekly increase in NT measurements (expected 20 to 25% per week), the median NT MoM (expected 1.00) and logarithmic standard deviation (between about 0.09 and 0.13). Usually, selected images are also reviewed for sonographer adherence to protocols.

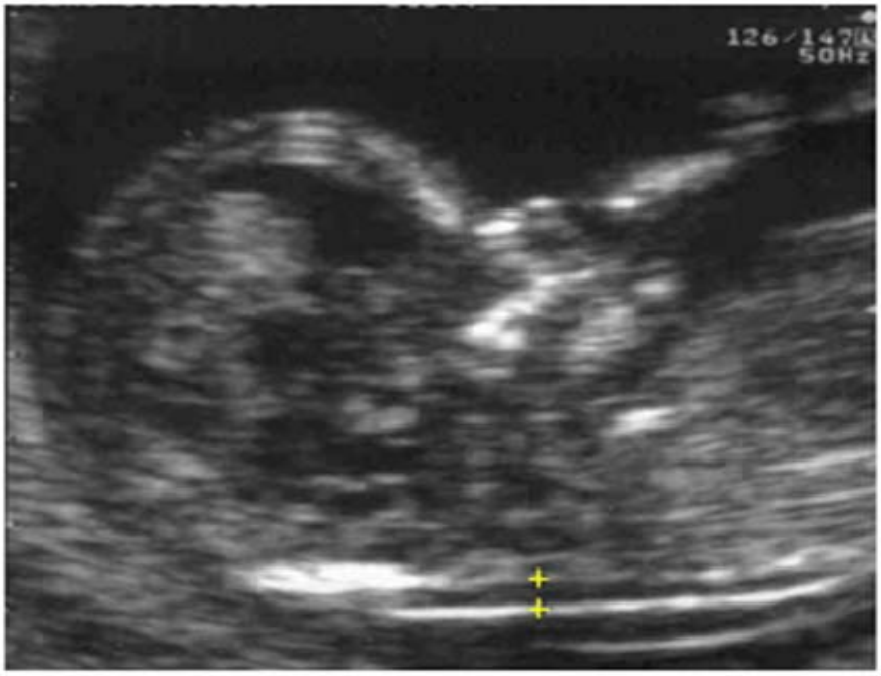

Figure 5.2-1. Sonographic image of a late first trimester fetal head and thorax showing the correct measurement of nuchal translucency. The calipers (+) are placed on the fetal spine and skin in the 'on to on' position. This indicates that the lower edge of the horizontal line in the '+' is on the lower edge of the fetal spine, and the upper edge of the horizontal line for the lower marker is on the upper edge of the fetal skin. Note the amnion visible below these two structures.

Literature search:

A literature search using the search phrase '[(nuchal translucency OR nuchal edema OR cystic hygroma) AND (Down syndrome or trisomy 18 or aneuploidy) AND first trimester]' was performed on PubMed; 682 references were identified. After reviewing titles, over 100 papers were examined for relevance, and reference lists were searched for additional publications. Case reports, studies that did not include any fetus affected with trisomy 18, reviews and publications in foreign languages were excluded. The usual protocol would then be to identify observational studies or those in which NT was not used in offering diagnostic testing. This helps avoid ascertainment bias. Unfortunately, only a few very early studies (usually referring to cystic hygroma) satisfied these criteria (Gembruch *et al.*, 1988; Bronshtein *et al.*, 1989; Cullen *et al.*, 1990; Droste *et al.*, 1991; MacLeod and McHugo, 1991; Shulman *et al.*, 1992; van Zalen-Sprock *et al.*, 1992; Ville *et al.*, 1992; Hewitt, 1993; Nadel *et al.*, 1993; Savoldelli *et al.*, 1993; Trauffer *et al.*, 1994; Podobnik *et al.*, 1995; Nicolaides *et al.*, 1992a; Cullen *et al.*, 1995), and they did not contain sufficient observations to allow appropriate population parameters to be derived. Early studies also suffered from non-standard measurement of NT and did not account for the increase in NT measurements by gestational age.

Many of the studies that included large numbers of affected pregnancies appear to be strongly influenced by biases, especially referral bias. For example, consider a series of papers from Kings College (Nicolaides *et al.*, 1994; Pandya *et al.*, 1994; Sherod *et al.*, 1997; Nicolaides *et al.*, 1992a; Pandya *et al.*, 1995). This group first reported initial experiences and then issued updated reports derived from their expanding database. In earlier publications (Pandya *et al.*, 1994; Pandya *et al.*, 1995) they declared that a high proportion of their samples were referred due to elevated NT from surrounding practices (23% and 30%, respectively). Later, the proportion was not reported, but it seems likely that referrals continued. A high rate of referrals can also be inferred from the very high rate of trisomy 18 reported (1:18 and 1:20, respectively) that would be unlikely, even if the women were of advanced maternal age. Studies with clear referral bias are unsuitable for creating population parameters, as high NT measurements are overrepresented and identifying which cases might or might not be referred is not possible.

In this situation, a potential alternative to a formal summary of the literature might be to identify a single study that had limited biases. If these biases could be accounted for by analysis, then the resulting parameters may be suitable for future modeling and assignment of risks. The study would need to be relatively large, occur at a time when

the technique of NT measurements was well described, come from a site that had demonstrated compliance with the techniques, and would be subject to few, if any, other biases. The only bias that can be readily accounted for is trimester of ascertainment. One study demonstrated how this bias could be accounted for when screening for Down syndrome (Nicolaides *et al.*, 1998). A random subset of observations with positive NT measurements was removed from the dataset, and the log means and standard deviations were recalculated. The proportion of samples removed was equivalent to the fetal loss expected from the late first trimester to term. I have also applied that methodology to another Down syndrome/NT dataset with trimester of ascertainment bias (Spencer *et al.*, 2003). The resulting 'adjusted' distribution parameters (Palomaki *et al.*, 2007) agree well with another parameter set derived from a general screening study (Wald *et al.*, 2003; Wald, 2006).

Figure 5.2-2 shows a hypothetical distribution of 100 NT observations derived from a log Gaussian distribution with a median NT MoM of 2.0, and a logarithmic standard deviation of 0.12 (these are actually representative values for Down syndrome pregnancies; the distribution for trisomy 18 is, as yet, considered unknown). Further, assume that these observations are followed to term without diagnostic testing or selective terminations. In Chapter 2, it was found that 72% of trisomy 18 pregnancies do not survive to term. The 72 of 100 observations that will not survive are shown in the second panel of Figure 5.2-2 as open circles. The filled circles represent the 28 of 100 observations that will survive to term (panel 3). Although fewer observations are present at term, the distribution of NT measurements is essentially unchanged and either would provide a more unbiased estimate for the NT distribution.

However, it is not possible to know about all 100 cases in the first trimester, and it is also now routine for diagnostic testing to occur and for identified trisomy 18 pregnancies to be selectively terminated. In Figure 5.2-3, the same 100 trisomy 18 pregnancies are observed in the late first trimester (panel 1). However, in this scenario, all will have NT measurements, with an NT above 2.0 MoM resulting in diagnostic testing and likely termination of the pregnancy. Since the median of the distribution is 2.0, half will be above, and half below, this value. In panel 2, we see that all of the affected pregnancies above 2.0 MoM have been identified via NT testing, regardless of whether or not they would have survived to term. However, among the cases with NT levels below 2.0 MoM, only the filled circles (those 28% destined to go to term) will be identified at birth. The trisomy 18 pregnancies that are spontaneously lost (open circles below 2.0 MoM in panel 2 of Figure 5.2-2) will not be identified. Panel 3 shows the resulting distribution of NT

measurements among those with lower levels that go to term, and those with higher levels that were identified in the late first trimester. This is the 'trimester of ascertainment bias' and its presence results in the distribution of NT measurements being skewed towards higher values and a smaller standard deviation (notice the distribution of panel 3 in the Figure 5.2-3 appears 'tighter' than the corresponding distribution in Figure 5.2-2).

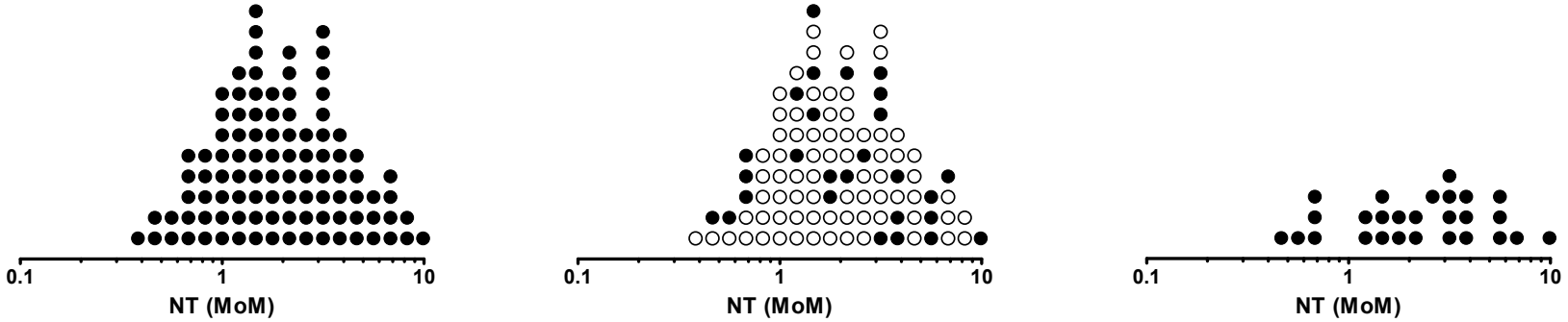

Figure 5.2-2. Hypothetical distribution of NT measurements in trisomy 18 pregnancies in the absence of screening and selective termination. In the left figure is the distribution of NT measurements in the first trimester, centered at 2.0 MoM. The middle figure shows those pregnancies that will survive to term (filled) versus those that will be spontaneously lost (filled). The right figure shows the distribution of NT measurements at term. The distribution, although based on smaller numbers, is equivalent to the left figure for mean and SD.

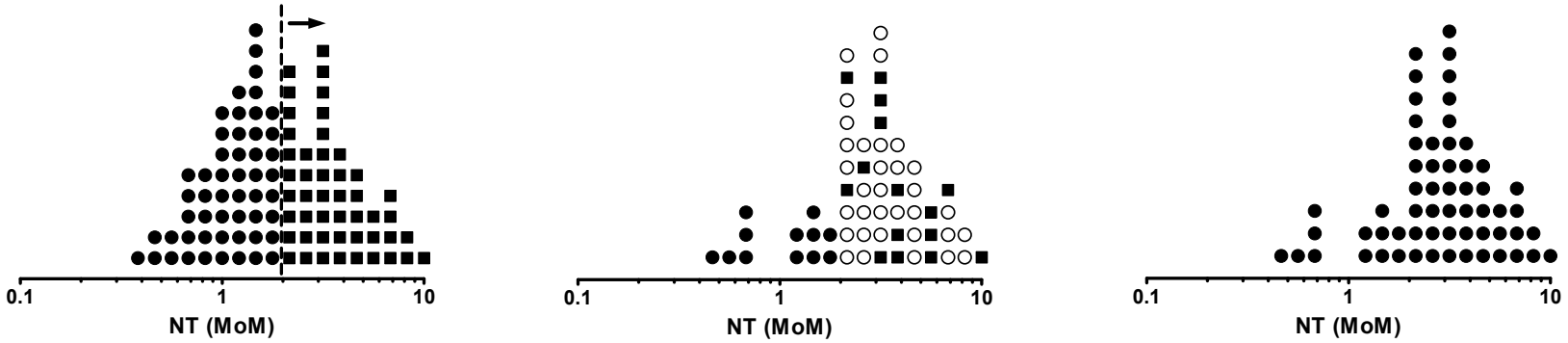

Figure 5.2-3. Hypothetical distribution of NT measurements in trisomy 18 pregnancies in the presence of screening and selective termination. In the left figure is the distribution of NT measurements in the first trimester. The half that are screen positive (filled squares) will be diagnosed in the first trimester. The middle figure shows the expected losses for those below 2.0 MoM (as in the upper figure), and what would have occurred would those with the higher levels not have been terminated. The right figure shows the distribution of NT measurements at term, now centered at 3.0 rather than 2.0 MoM. This latter distribution suffers from trimester of ascertainment bias. Methods described in the text can help overcome this problem.

Results

With this as background, one study (Tul *et al.*, 1999) appears suitable to apply a methodology for adjusting for trimester of ascertainment bias, in order to estimate a more unbiased distribution of NT measurements in late first trimester trisomy 18 fetuses. It comes from a site that adheres to the Fetal Medicine Foundation methodology for NT measurements (thus reliable measurements), is reasonably large (50 trisomy 18 fetuses) and is subject only to the trimester of ascertainment bias. Figure 5.2-4 shows the probability plot of the 50 cases from that study, as they were reported on the left-hand side of the page.

In order to 'unbias' the data, it is necessary to selectively remove "about half of those terminated in the first trimester and one-third of those terminated in the second trimester"(Nicolaides *et al.*, 1998). Without direct access to the individual pregnancy outcome, this is not possible, so a statistical approach will be taken instead. Few terminations would have likely occurred among screen negative women, and the majority of women with screen positive results would have diagnostic testing. Since this cohort of women was screened using maternal age and NT measurements (the aim of the study was to find the distribution of serum markers in trisomy 18), the NT measurements will dominate the Down syndrome risk calculation. For example, in a 20 year old and a 40 year old, NT measurements above 1.9 and 1.2 MoM, respectively would be considered screen positive (term risk greater than or equal a 38 year old woman). Thus, affected fetuses with NT MoM levels below 1.2 would have likely been screen negative and have been diagnosed at term, while those above 1.9 would likely have been diagnosed in the late first or early second trimester. Together, an estimated 72% (of these early detected cases would have been spontaneously lost (Table 2.4-1, (Morris and Savva, 2008)). In order to 'unbias' this selection of trisomy 18 cases, each case falling above 1.9 MoM will have a 72% random chance of being removed from the analysis. For the few samples falling between 1.2 and 1.9 MoM, a linear extrapolation from 72% to 0% removal was applied. None of those falling below 1.2 MoM would be removed, because the assumption is that they were identified at term.

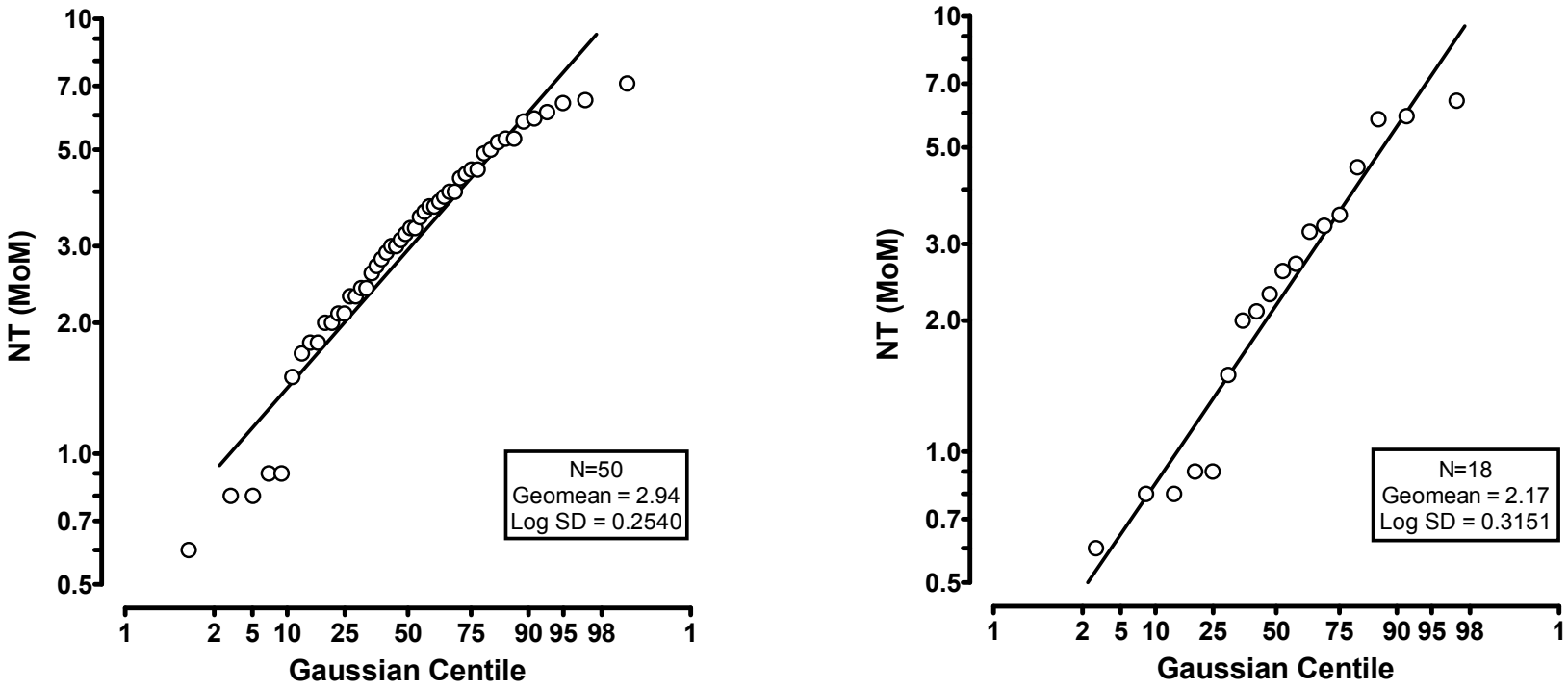

Figure 5.2-4. The effect of accounting for bias of ascertainment on the NT distribution parameters for trisomy 18. On the left side is the dataset as published (Tul *et al.*, 1999). On the right side, an unbiasing technique has been applied (see text for a full description). In that technique, a high proportion (72%) of elevated NT measurements have been removed (likely to have been diagnosed in the first trimester, but not likely to go to term), while those with lower measurements have been retained (likely to have gone to term). In this adjusted dataset, the median is considerably lower (2.17 versus 2.94), while the standard deviation is higher (0.3151 versus 0.2540).

This technique was then performed 10 times, with the overall median MoM and log standard deviation being 2.17 (log mean 0.3373 or 2.16) and 0.3151, respectively. Figure 5.2-4 shows (on the right) one of the 10 sample datasets from that exercise. Although the number of observations is reduced considerably (in the 10 simulations, N ranged from 14 to 22). The slope of the straight line indicates the summary standard deviation, and goes through the summary log mean of 0.3373 (2.16 MoM). These data fit a straight line far better than in the original plot (left figure in 5.2-4). The distribution's median value is lowered, from about 2.94 to 2.17, while the logarithmic standard deviation increases from 0.2540 to 0.3151.

Table 5.2-1 examines the effect of these changes in NT parameters when interpreting NT measurements for trisomy 18 risk. These are univariate detection and false positive rates and do not include maternal age or other markers. Distribution parameters for unaffected pregnancies (log SD 0.1105 at 11 weeks) are from a recent update (Bestwick *et al.*, 2010). When comparing the two parameter sets, the detection rate for NT measurements alone at a 1% false positive rate is 80% (Tul *et al.*, 1999), versus 60% after the unbiasing procedures. Similar differences are seen at higher false positive rates. When detection rates are held constant (bottom of Table 5.2-1), the differences are also clearly visible, with false positive rates of 0.1% and 6%, respectively, at a detection rate of 70%.

At least two publications (Wald *et al.*, 2003; Hyett *et al.*, 1996) have reported that both unaffected and affected pregnancies (usually Down syndrome) have a higher rate of spontaneous fetal loss as the NT measurement increases. In SURUSS, for example (Wald *et al.*, 2003) women with NT MoMs above the 95th centile, or about 1.5 MoM, had twice the rate of fetal loss than those with lower levels. This indicates a potential for the trisomy 18 cases identified via high NT measurements to be lost at even a higher rate than the 72% used in modeling. However, this is unlikely to have any important effect on the revised parameters computed here, because 44 of the 56 cases of trisomy 21 in the dataset (Tul *et al.*, 1999) were at or above 1.5 MoM. There are too few data to suggest that one can further stratify levels above the 95th centile into a dose-response effect. Thus, this refinement to the model is unnecessary.

Table 5.2-1. Performance of NT measurement in identifying trisomy 18 in the late first trimester, according to one dataset before and after an unbiasing protocol was applied

False positive rate	Detection rate (Original)	Detection rate (Adjusted)	Change (%)
1%	80%	60%	-20%
2%	83%	64%	-19%
3%	85%	66%	-19%
4%	86%	68%	-18%
5%	87%	69%	-18%

Detection rate	False positive rate (Original)	False positive rate (Adjusted)	Change (%)
60%	<0.1%	1.0%	0.9%
65%	<0.1%	2.5%	2.4%
70%	0.1%	6.0%	5.9%
75%	0.4%	13.0%	12.6%
80%	1.1%	25.6%	24.5%

If the biases act as expected on the distribution of NT measurements in trisomy 18 fetuses, then a dataset with both trimester of ascertainment bias (as described earlier) and referral bias should deviate even more from a Gaussian distribution and provide even higher estimates of test performance. Referral bias occurs when a tertiary care site receives patients that have already been tested, and are being referred because of an abnormal measurement. In this setting, it might be primary care practices performing NT measurements, but referring those with elevated measurements to a 'high risk' practice for confirmation and diagnostic testing.

Figure 5.2-5 shows a paired set of probability plots, the left being a repeat of the uncorrected dataset with only ascertainment bias (Tul *et al.*, 1999). The right is a larger dataset of 106 cases of trisomy 18 from a high risk center published two years earlier (Sherod *et al.*, 1997), that acknowledged high NT referrals to that center in earlier, overlapping publications (Pandya *et al.*, 1994; Pandya *et al.*, 1995). The median (3.20 MoM), geometric mean (2.82 MoM) and logarithmic standard deviation (0.2528) have been derived from the Figure and were not reported.

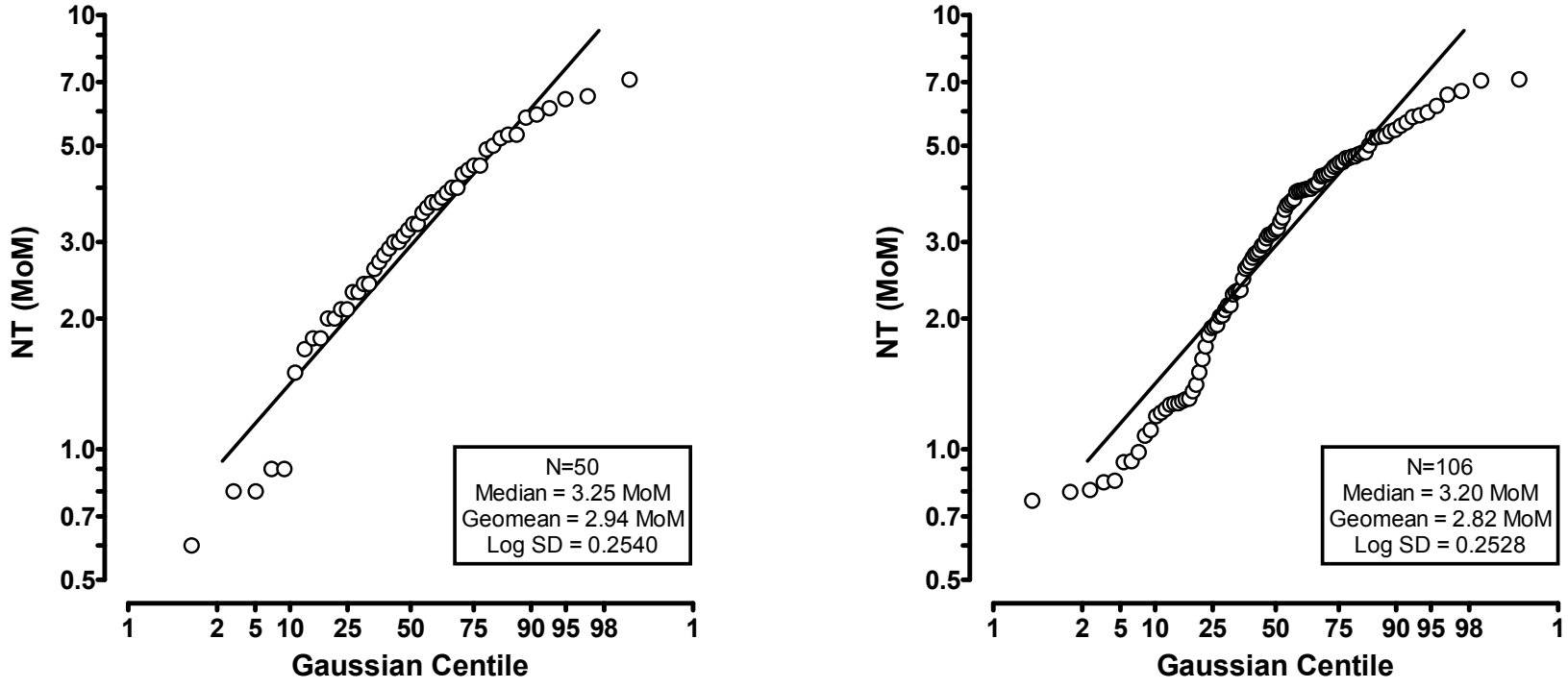

Figure 5.2-5. The effect on the NT distribution parameters for trisomy 18 of including a referral population with increased NT measurements. On the left side are data shown earlier with suspected bias of ascertainment (Tul *et al.*, 1999). On the right side, are data from a study suspected to have both ascertainment and referral biases (Sherod *et al.*, 1997). In this larger dataset, the deviation from the straight line (representing a log Gaussian distribution) is more pronounced, even though the geometric mean and log SD are similar to that found in the smaller dataset on the left.

Conclusions

At least two important biases, trimester of ascertainment and referrals, cause problems in determining the distribution parameters for NT measurements in trisomy 18 pregnancies. Although these biases exist when testing for Down syndrome as well, the problem is exacerbated when dealing with trisomy 18 because the fetal losses during pregnancy are so much higher than for Down syndrome (78% versus 32%). It was possible to identify one relatively large dataset that appeared to have only the bias of ascertainment. A previously described statistical method was used in an attempt to account for the impact of the bias on the reported distribution. The adjusted parameters indicate that the separation between NT measurements in trisomy 18 and unaffected pregnancies is less than published estimates, and revisions in the logarithmic standard deviation are also needed. This is especially important given the recent publication of a 'mixture model' designed to better fit the distribution of NT measurements in affected pregnancies (Wright *et al.*, 2008). These authors rely on the 'non-Gaussian' appearance of the data (like that shown in Figure 5.2-5 in the right scatterplot) as the justification for more complex modeling. They do not appear to have considered that the deviations might be due to well described biases. This effort appears to be a futile modeling exercise aimed at fitting biased data that does not represent the distribution of NT measurements in a general pregnancy population. Others in this research group in London have 'confirmed' the improvement of the mixture model on an independent dataset (Kagan *et al.*, 2008a). However, that new dataset was also subject to the same two biases, so the finding of a 'good fit' is both expected, and irrelevant for trisomy 18 testing in the general pregnancy population.

In the next Chapter (6), results from the review of the first trimester serum markers for trisomy 18 (Chapter 4) will be combined with these NT results to model how maternal age in combination with, PAPP-A, free hCG and NT measurements might perform.

5.3 Nasal bone

Definition:

Among the findings first described by Langdon Down in 1886 (Down, 1866) is one that reads: “The nose is small”. As with excess skin at the back of the neck leading to measurement of nuchal thickness in the second trimester, and nuchal translucency in the first trimester, this statement suggests that the nasal bone is early pregnancy might be smaller, or even absent in the presence of Down syndrome. Radiographic post-mortem studies 15 to 40 weeks’ gestation provided evidence that the nasal bone is absent in 25% of Down syndrome fetuses. This finding was rare among normal fetuses (Stempfle *et al.*, 1999).

The first study to report on nasal bone (NB) measurements (Cicero *et al.*, 2001) noted that the methodology for identifying the presence/absence of the NB was not straightforward.

> For examination of the fetal nose, a mid-sagittal view of the fetus was obtained, with the beam of the ultrasound transducer being parallel to the nasal bone. In this position, the skin of the nose produces an echogenic line, which can be misinterpreted as the nasal bone. To avoid this mistake, the ultrasound transducer was gently tilted from side to side to ensure that the nasal bone was seen separate from the nasal skin.

That same author (Cicero *et al.*, 2006) elaborated on the methodology five years later.

> For examination of the nasal bone, the image was magnified so that the head and upper thorax only were included in the screen and a midsagittal view of the fetal profile was obtained. The ultrasound transducer was parallel to the direction of the nose and the probe was gently tilted from one side to the other of the fetal nose. When these criteria were satisfied, 3 distinct lines were seen at the level of the fetal nose. The first 2, which are proximal to the forehead, are horizontal and parallel to each other, resembling an “equal sign”. The top line represents the skin and the bottom one, which is thicker and more echogenic than the overlying skin, represents the nasal bone. A third line, almost in continuity with the skin, but at a higher level, represents the tip of the nose. The nasal bone is considered to be present if it is more echogenic than the overlying skin and absent if it is either not visible or its echogenicity is the same or less than that of the skin.

Figure 5.3-1 shows an appropriate view of the fetal head, with the clearly visible "equal sign". From this image, the nasal bone is clearly "present" (lower line in the equals sign).

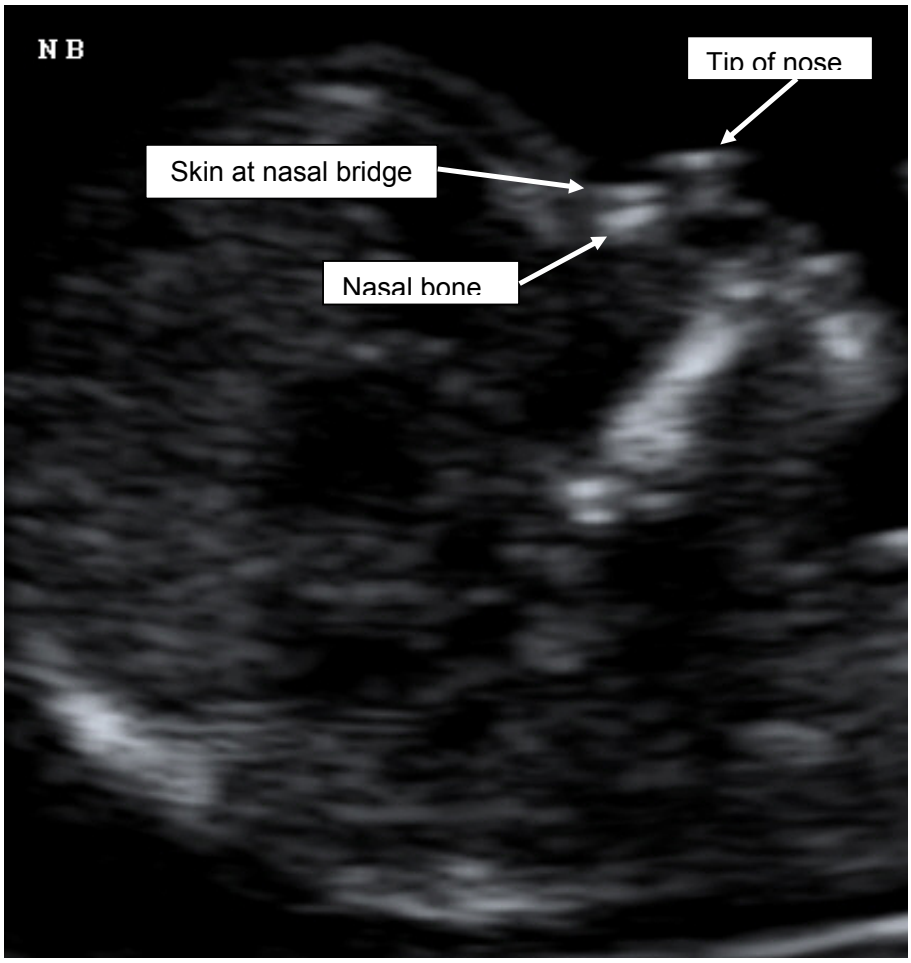

Figure 5.3-1. An ultrasound image showing a fetal head in the late first trimester, along with the identification of the nasal bone. In this midsagittal view, the "equal sign" is clearly visible, made up of the nasal bone (lower line) and the skin on the nose (upper line). The tip of the nose is also visible. These three landmarks indicate an appropriate view for detecting the presence (or absence) of the nasal bone (image courtesy of Dr. Alan Donnenfeld).

Literature search

A literature search was performed using the terms "nasal bone", "Down syndrome" and "Trisomy 18" and 21 references were identified. After abstract review, potentially relevant articles were obtained in full and the reference lists searched for other relevant

articles. In order to be included, sufficient data needed to be present to describe the clinical scenario (prior to CVS after a positive screening test) and the proportion with an absent nasal bone. If data were available for Down syndrome and/or trisomy 13 pregnancies, these were included in the summaries. In four studies, trisomy 18 pregnancies were included, but the proportion of cases with an absent NB was not reported (Kozlowski *et al.*, 2006; Peralta *et al.*, 2005; Prefumo *et al.*, 2005; Weingertner *et al.*, 2006). A total of 12 publications provided NB information in at least one first trimester fetus with trisomy 18 (Cicero *et al.*, 2001; Otano *et al.*, 2002; Cicero *et al.*, 2003; Orlandi *et al.*, 2003; Viora *et al.*, 2003; Wong *et al.*, 2003; Zoppi *et al.*, 2003; Cicero *et al.*, 2004; Malone *et al.*, 2004; Cicero *et al.*, 2006; Kagan *et al.*, 2009a; Sepulveda *et al.*, 2010). Figure 5.3-2 shows the number of reported trisomy 18 (and trisomy 13) cases by year having nasal bone measurements, with the first report appearing in 2001. One group in the UK has published extensively on nasal bone, and their reports include overlapping datasets (Cicero *et al.*, 2001; Cicero *et al.*, 2003; Cicero *et al.*, 2004; Cicero *et al.*, 2006; Kagan *et al.*, 2009a). These publications are connected by a thin dotted line. Circles and squares are used to indicate trisomy 18 and trisomy 13 cases, respectively. Filled symbols indicate that most, if not all, of the study participants were screen positive, usually by a combination of nuchal translucency and maternal age.

Results

Table 5.3-1 shows relevant data on trisomy 18 and nasal bone from the 12 included studies. All were subject to important potential biases that have been described earlier. Data for trisomy 13 and Down syndrome are also listed for comparison. All studies restricted gestational age to 11 to 14 weeks' gestation, but few observations were actually later than 13 completed weeks'. The four studies presenting overlapping data are listed at the bottom of the table and are not included in the analysis. In the remaining eight studies, nasal bone measurements were obtained in between 77% and 99% of fetuses examined. The summary estimate was 96.2% (95% CI 88.8% to 98.8%) but there was considerable heterogeneity ($Q=1020$, $p<0.001$, $I^2=99\%$). Although the FASTER study estimate of 77% was the lowest (Malone *et al.*, 2004) another study also reported a relatively low success rate of 83% (Wong *et al.*, 2003).

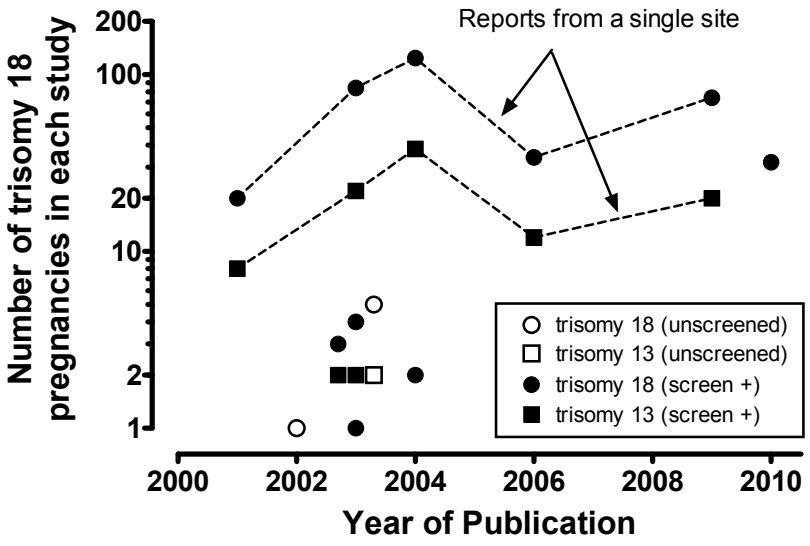

Figure 5.3-2. A summary of trisomy 18 (and trisomy 13) observations available to determine associations with absent nasal bone. Circles and squares are used to indicate trisomy 18 and trisomy 13 cases, respectively. Filled symbols indicated that most, if not all, of the study participants were screen positive, usually by a combination of nuchal translucency and maternal age. One group in the UK has published extensively on this topic, and their reports tend to include overlapping data. These publications are connected by a thin dotted line

The nasal bone was absent in between 0.6 and 4.4% of unaffected fetuses. The summary estimate was 0.9% (95% CI 0.5 – 1.5%), but heterogeneity was again high ($Q=95$, $p<0.001$, $I^2=96\%$). This heterogeneity is also seen in the multiple reports from the UK group; where one study reports a rate of about 2.6% (Cicero *et al.*, 2003; Cicero *et al.*, 2004; Kagan *et al.*, 2009a) while the remaining two have rates of about 0.5% (Cicero *et al.*, 2006; Cicero *et al.*, 2001).

The proportions of Down syndrome, trisomy 18 and trisomy 13 fetuses with absent nasal bone are all shown in Table 5.3-1, to ensure that the rates in trisomy 18 and 13 are put in the context of the larger experience in Down syndrome screening. The highest detection rate (69%) is found for Down syndrome in 401 reported cases. The detection rates are significantly lower for trisomy 18 (56% in 194 cases) and trisomy 13 (33% in 44 cases). Taken as a whole, the likelihood ratios associated with absent NB in the three major trisomies are 76, 62 and 37, respectively.

The majority of studies were, however, performed in settings that included women who were already screen positive, usually by a combination of maternal age and nuchal translucency. This would not bias the information about NB, unless the absence of the NB were correlated with NT measurements in unaffected pregnancies, affected pregnancies, or both. There are too few reported pregnancies affected with trisomy 18 or 13 to examine the NB/NT relationship, but sufficient numbers do exist for Down syndrome pregnancies. Some early reports (Cicero *et al.*, 2001; Orlandi *et al.*, 2003) stated that the two markers were completely independent. With experience, however, a clear relationship was found (Zoppi *et al.*, 2003; Cicero *et al.*, 2004; Cicero *et al.*, 2006). As NT levels increase, the proportion of Down syndrome (and control) pregnancies with absent NT also increases. Unfortunately, there are no sufficiently large datasets that performed NB measurements as a screening test in all women that could document the association across the range of NT values. One study included mainly older women (Orlandi *et al.*, 2003) in their cohort of over 1000 pregnancies, but only 15 cases of Down syndrome and 3 cases of trisomy 18 were studied. In a much larger study of women screen positive using maternal age and NT (Cicero *et al.*, 2004), the proportion of Down syndrome fetuses with an absent NB showed a modest increase from about 60% near the 95th centile of NT to nearly 80% when the NT were above 5.5 mm. However, few of these women actually had NT measurements in the normal range, as all were screen positive. Among the unaffected pregnancies, the change was more dramatic. The false positive rate (absent NB) was about 2% at the 95th centile of NT, but 15% for the few unaffected pregnancies with an NT above 5.5 mm. Due to the much larger change in the false positive rate in unaffected pregnancies, the likelihood ratio for NB becomes lower as the NT increases, with a reported range of 37 (in women near the 95th centile) to 5 (in women with NT >5.5 mm). Whether there is an association between NB and NT in trisomy 18 pregnancies has not been reported, but it would be reasonable to assume that it could be similar to that found for Down syndrome.

Identifying the nasal bone (or its absence) is considered by experienced sonographers to be more difficult than performing NT measurements (Cicero *et al.*, 2003; Senat *et al.*, 2003). One study (Senat *et al.*, 2003) compared three operators' measurements using 657 prerecorded loops. They found inter-operator kappa values of about 0.3 and intra-operator values of about 0.4. Both would be considered poor to fair. The operators agreed on only about 80% of the assessments (present, absent, uncertain). Another (Bekker *et al.*, 2004) compared the results of a 2D and 3D ultrasound and found fair to good correlations as well as a clear gestational age relationship with the length of the nasal bone. Adding to the difficulty in interpreting NB testing is a potential racial effect

(Cicero *et al.*, 2003; Prefumo *et al.*, 2005). In one of these reports (Cicero *et al.*, 2003), for example, 10.4% of unaffected fetuses of Afro-Caribbean ancestry had an absent NB, compared to only 2.5% in their Caucasian counterparts (2.5%).

One study examined the effect of NB measurements on the biochemistry results (Cicero *et al.*, 2005). They reported no significant difference in PAPP-A or free β hCG measurements in trisomy 18 pregnancies with absent (N=19) or present (N=15) NB. The point estimates do appear to be quite different between the groups, but the differences are not statistically significant. This may be due to the small sample size. Even if the differences were to be statistically significant, there is likely to be little impact on overall test performance, as one marker (free β hCG) is in the direction of lower risk, while the other PAPP-A) is in the direction of increased risk.

Summary

Specialized training is needed for reliable ultrasound detection (or measurement) of NB. Oversight is also necessary and would require more than monitoring of the rate of absent NB. The proper interpretation of NB would require adjustment for gestational age, the associated NT measurement, and race/ethnicity. The reported test performance is likely to be overestimated, due to trimester of ascertainment bias, and by the inclusion of referral patients in the population. Without adjusting for these biases, the univariate likelihood ratios associated with absent NB are 76, 62 and 37, for Down syndrome, trisomy 18 and trisomy 13, respectively. Demonstration projects performed in an unbiased low risk population, after appropriate training and oversight, have been recommended (Rosen *et al.*, 2007).

Table 5.3-1. Summary of published studies reporting nasal bone (NB) measurements in unaffected and common trisomic pregnancies between 11 and 14 weeks' gestation

Author	Country	Study type[1]	NB Success	NB Absent			
				Unaffected (%)	Down S (%)	T18 (%)	T13 (%)
(Otano *et al.*, 2002)	Argentina	MR cohort	NR	1/ 175 (0.6)	3/ 5 (60)	1/ 1 (100)	0
(Orlandi *et al.*, 2003)	Italy	MR cohort	94%	10/ 1,000 (1.0)	10/ 15 (67)	2/ 3 (67)	0/ 2 (0)
(Viora *et al.*, 2003)	Italy	HR cohort	92%	24/ 1,709 (1.4)	6/ 10 (60)	1/ 4 (25)	1/ 2 (50)
(Wong *et al.*, 2003)	Hong Kong	HR cohort	83%	1/ 114 (0.9)	0	0/ 1 (0)	0
(Zoppi *et al.*, 2003)	Italy	MR cohort	99%	34/ 5,525 (0.6)	19/ 27 (70)	19 /27 (70)	0/ 2 (0)
(Cicero *et al.*, 2004)	UK	HR cohort	99%	129/ 5,223 (2.5)	229/333 (69)	68/124 (55)	13/38 (34)
(Malone *et al.*, 2004)	US	LR cohort	77%	21/ 4,790 (0.4)	0/ 11 (0)	1/ 2 (0)	0
(Sepulveda *et al.*, 2010)	Chile	Case only	NR	0	0	17/ 32 (53)	0
All (random effects model)			**96.2%**	220/18,536 **(0.9)**	266/401 **(68)**	109/194 **(56)**	14/44 **(33)**
95% CI			(88.8 – 98.8)	(0.5 – 1.5)	(63 – 72)	(49 – 63)	(21 – 48)
Studies with data included in Cicero S, 2004 (included above)							
(Cicero *et al.*, 2001)	UK	HR cohort	100%	3/ 603 (0.5)	43/ 59 (73)	11/ 20 (55)	2/ 8 (24)
(Cicero *et al.*, 2003)	UK	HR cohort	100%	93/ 3,358 (2.8)	161/242 (67)	48/ 84 (57)	7/22 (32)
(Cicero *et al.*, 2006)	UK	HR cohort	98%	113/20,165 (0.6)	98/142 (69)	22/ 40 (55)	3/12 (25)
(Kagan *et al.*, 2009a)	UK	HR cohort	99%	513/19614 (2.6)	73/122 (60)	19 /36 (53)	9/20 (45)

[1] HR = high risk, indicating that all, or nearly all women were screen positive, usually by NT and maternal age, MR = moderate risk, indicating that at most, a small portion were screen positive, but many were age 35 or older, LR = low risk, indicating close to a general pregnancy population.

5.4 Ductus Venosus

Introduction

Heart defects are commonplace in both Down syndrome and trisomy 18 fetuses (Chapter 1). There also appears to be a linkage between certain types of heart defects and increased nuchal translucency (NT) measurements (Hyett *et al.*, 1997). This led to observations of abnormal ductus venosus flow in chromosomally abnormal fetuses with increased NT measurements (Matias *et al.*, 1998; Kiserud, 1997). The ductus venosus is present only in the fetus and allows up to 80% of the umbilical vein blood flow to go directly into the fetal circulation. Using color Doppler ultrasound imaging techniques, it is possible to identify and quantify abnormalities in the flow. Figure 5.4-1 is a stylized diagram of the three main ductus venosus waveforms observed in the late first trimester.

There are several methods for identifying abnormal ductus venosus flow. The simplest is to classify the waveforms as in Figure 5.4-1: positive, absent or reversed. Some researchers measure the height of the A-wave (in cm/sec). Still others compute a pulsatility index for veins (PIV) for the ductus venosus. Both of these latter methods aim to quantify how abnormal the waveform might be.

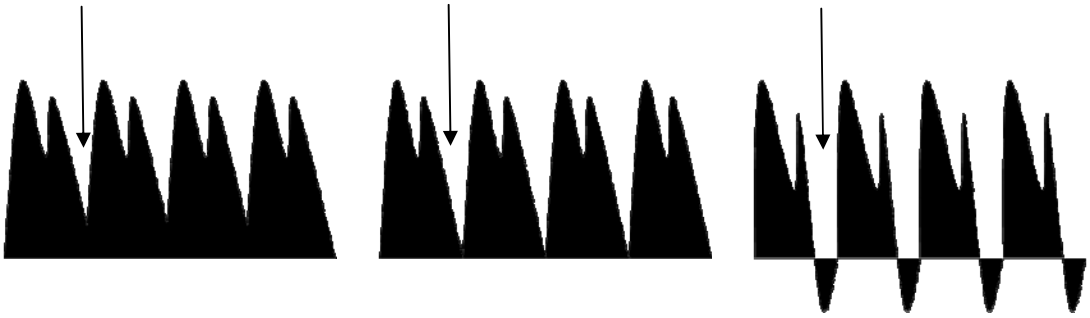

Figure 5.4-1. Three ductus venosus Doppler waveforms demonstrating differences in the A-wave. In these three figures, time is on the x-axis, and venous blood flow is on the vertical axis. The baseline indicates no blood is flowing. In the left figure, the arrow points to a tracing showing a positive A-wave, in the middle figure, the A-wave is absent, and in the right figure, the A-wave is reversed (negative), indicating venous blood flow going to the placenta. The positive A-wave is considered normal, while the absent or reversed A-wave is considered abnormal (figures courtesy of Dr. Tony Borrell).

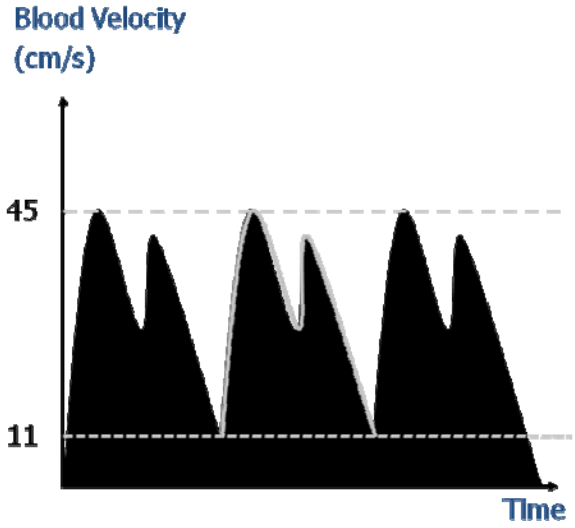

Figure 5.4-2. A ductus venosus waveform showing the components needed to compute the pulsatility index for veins (PIV). The A-wave height is 11 cm/sec (lower horizontal dashed line) while the maximum S-wave height is 45 cm/s (upper horizontal dashed line). The average height of the waveform integrated over time is 31 cm/sec. The PIV is computed as (S-A)/average. In this example, the PIV is 1.10 (45-11)/31 (Figure adapted from one provided by Dr. Tony Borrell).

Figure 5.4-2 shows how the PIV is calculated. For positive A-wave findings, the PIV will be relatively small (*e.g.,* 1.10). If the A-wave is reversed, the A-wave value will be negative, resulting in larger PIV values. For example, the negative A-wave on the right side of Figure 5.4-2, might have an S-wave of 45, an A-wave of -15, with a time averaged height of 21. This would result in a PIV of 2.76 and be considered abnormally high.

As with measurement of NT and NB, there is expert guidance in the performance of reliable ductus venous measurements. Specifically, the fetus should be quiescent with a sagittal view of the entire trunk. Pulsed Doppler is applied with a minimal insonation angle. The sweep speed should be sufficiently high to accurately capture the waveform. When performing measurements (A-wave height or PIV), some groups suggest performing the study three times and taking the mean value for interpretation. Due to concerns about fetal safety, the time spent in the Doppler imaging mode is usually restricted to 5 or 10 minutes, or less.

Literature search

A literature search using the terms “ductus venosus”, “Down syndrome” and “trisomy 18” was performed, and 8 references were identified. After abstract review, potentially relevant articles were obtained and the reference lists searched for other relevant articles. In order to be included, sufficient data needed to be present to describe the clinical scenario (prior to CVS after a positive screening test) and the proportion with an abnormal ductus venosus (DV) blood flow. If data were available for trisomy 18 and/or trisomy 13, these and any associated Down syndrome pregnancies were included in the summaries. One study (Prefumo *et al.*, 2005) measured DV blood flow in 12 trisomy 18 and 5 trisomy 13 pregnancies, but did not report any of those results and was excluded. A total of 11 publications satisfied inclusion criteria (Huisman and Bilardo, 1997; Montenegro *et al.*, 1997; Matias *et al.*, 1998; Antolin *et al.*, 2001; Bilardo *et al.*, 2001; Mavrides *et al.*, 2002; Murta *et al.*, 2002; Borrell *et al.*, 2003; Toyama *et al.*, 2004; Maiz *et al.*, 2009; Zoppi *et al.*, 2002). In two of these studies, results were not reported separately for each aneuploidy (Mavrides *et al.*, 2002; Borrell *et al.*, 2003) but are included in the data listings, given that DV blood flow is considered to be equally effective for identifying all three aneuploidies.

Figure 5.4-3 shows the number of reported trisomy 18 (and trisomy 13) cases by year, with the first report appearing in 1997. Circles and squares are used to indicate trisomy 18 and trisomy 13 cases, respectively. Filled symbols indicate that most, if not all, of the included study participants were screen positive, usually by a combination of nuchal translucency and maternal age. Open symbols indicate populations with a lower prevalence of aneuploidy and fewer participants subject to preliminary screening tests.

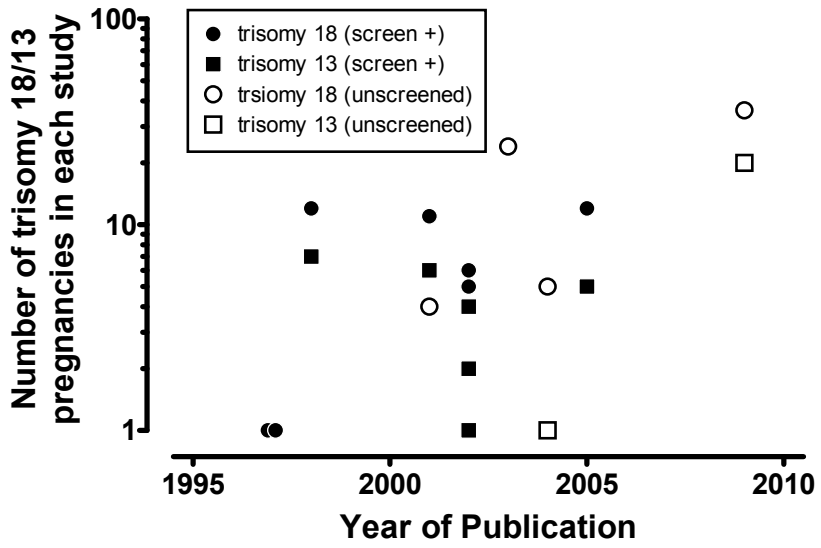

Figure 5.4-3. A summary of trisomy 18 (and trisomy 13) observations available to determine associations with abnormal ductus venosus flow. Filled symbols indicated that most, if not all, of the study participants were already screen positive for Down syndrome, usually by a combination of nuchal translucency and maternal age.

Results

Table 5.4-1 shows relevant DV data for pregnancies affected with trisomy 18 from the 11 studies. Many of these studies suffer from important potential biases that have been described earlier. Data for Down syndrome and trisomy 13 are also listed, when reported. All studies restricted gestational age to 11 to 14 weeks' gestation, but a few observations were later than 13 completed weeks'. There are 12 datasets, as one study reported results for both A-wave height and DV-PIV (Borrell *et al.*, 2003). The upper half of the table includes studies of the A-wave only (using categories of positive versus absent or reversed, or the equivalent group of the height of the A-wave). The lower half of the table shows study results for those using the DV-PIV.

Table 5.4-1 shows the reported false positive rates, ranging from 1.9% to 17%. The overall false positive rates in the two groups of studies, and overall, are 4.0%, 8.4% and 5.8%, respectively. All of these estimates are associated with high heterogeneity ($Q>60$, $p<0.001$ and $I^2>90\%$ for each). Two studies (Bilardo *et al.*, 2001; Zoppi *et al.*, 2002) had unusually high false positive rates of 13% and 17%, respectively, and both reported results as PIV. However, these studies were not associated with higher detection rates,

compared to the studies with lower false positive rates. This may indicate a difference in technique or cut-off level that was not apparent in the reports. With these two studies removed, the range of false positive rates is 1.9 to 8.4%, with all but two between 3.1 and 6.6%. Summary detection rates for Down syndrome, trisomy 18 and trisomy 13 are shown in the last three columns. Individual rates generally fall between 60 and 90%. Overall detection rates for these three aneuploidies are 74%, 67% and 56%, respectively, but due to the relatively small sample sizes (94, 74 and 47 cases, respectively), these rates are not significantly different (Q=3.8, p=0.15). Using the summary false positive rate of 8.4%, the likelihood ratios for the three aneuploidies are 12, 11 and 10, respectively. Were a more conservative false positive rate of 5% to be used (the median of the rates after excluding the two high estimates), the likelihood ratios would increase to 15, 13 and 11, respectively.

In order for DV measurements to be useful in routine clinical practice, it is important that the test results be reproducible both within and between sonographers. Three studies examined this aspect of testing (Mavrides *et al.*, 2002; Prefumo *et al.*, 2001; Borrell *et al.*, 2007). The Italian group (Prefumo *et al.*, 2001) collected four repeated measurements in 22 fetuses for the PIV, S-wave, A-wave, and time averaged maximum velocity, with corresponding coefficients of variation (CVs) of 10%, 13% 22% and 13%, respectively. Between-observer differences for an average result for the same four measurements resulted in CVs of 9%, 14%, 27% and 15%, respectively. All sonographers provided the sample normal/abnormal classifications for all fetuses examined. At about the same time, a group in England (Mavrides *et al.*, 2001) found good with-in observer performance for PIV (9%), but less good performance for S-wave, A-wave and S/A ratios (19%, 29% and 25%, respectively). The CVs between observers are also less good, ranging between 12 and 47%. They concluded that only the PIV measurement could be considered reproducible. More recently, a group from Catalonia (Borrell *et al.*, 2007) examined reproducibility of the A-wave and the PIV and confirmed that PIV is the most reliable DV measure.

Summary

A ductus venosus blood flow test can be reliably measured and is associated with an LR of about 12 (range 10 to 15). This finding appears to be consistent in both high risk and lower risk settings, and varies only slightly for the three aneuploidies studied. Some unusually high false positive rates for DV-PIV measurements were reported by some groups. However, other studies provide some evidence that the DV-PIV may be more reliable and more predictive than measuring the A-wave height or categorizing the direction of flow. Such testing could potentially be used in routine testing, but its overall complexity, required equipment, extensive training and continuing external oversight makes the test more appropriate for high risk referral centers.

Table 5.4-1. Summary of published studies reporting Ductus Venosus (DV) measurements in unaffected and common trisomic pregnancies between 11 and 14 weeks' gestation.

Author	Country	Study type[1]	DV Success	Abnormal Ductus Venosus Flow			
				Unaffected (%)	Down S (%)	T18 (%)	T13 (%)
Ductus Venosus (A wave absent or negative)							
(Huisman and Bilardo, 1997)	Netherlands	Case study	NR	NR	0	1/ 1 (100)	0
(Montenegro *et al.*, 1997)	Portugal	HR Cohort	NR	NR	4/ 4 (100)	1/ 1 (100)	0
(Matias *et al.*, 1998)	Portugal/UK	HR Cohort	100%	13/ 423 (3.1)	35/ 38 (92)	12/12 (100)	5/ 7 (72)
(Mavrides *et al.*, 2002)	UK	HR Cohort	98%	10/ 204 (4.9)	30[2]	5[2]	4[2]
(Murta *et al.*, 2002)	Brazil	HR Cohort	99.7%	6/ 321 (1.9)	17/ 18 (94)	1/ 1 (100)	2/ 2 (100)
(Borrell *et al.*, 2003)	Catalonia	MR Cohort	NR	162/ 3,248 (5.0)	28/ 48 (58)	7/12[3] (58)	7/12[3] (58)
(Toyama *et al.*, 2004)	Brazil	LR Cohort	NR	79/ 1,195 (6.6)	5/ 7 (71)	3/ 5 (60)	1/ 1 (100)
(Maiz *et al.*, 2009)	UK	LR Cohort	NR	622/19,614 (3.2)	81/122 (66)	21/36 (58)	11/20 (55)
(excludes T18/13 combined data from Mavrides 2002)				920/25,502 (4.0) (2.7 – 5.8)	217/284 (75) (62 - 85)	46/68 (62) (46 - 75)	26/42 (61) (43 - 76)
Ductus Venosus (pulsatility index for veins or PIV)							
(Antolin *et al.*, 2001)	Spain	LR Cohort	100%	68/ 1,351 (5.0)	6/ 7 (86)	4/ 4 (100)	0
(Bilardo *et al.*, 2001)	Netherlands	HR Cohort	86%	22/ 130 (17)	12/ 19 (63)	9/11 (92)	1/ 6 (17)
(Zoppi *et al.*, 2002)	Italy	HR Cohort	98%	38/ 292 (13)	14/ 20 (70)	6/ 7 (86)	1/ 1 (100)
(Borrell *et al.*, 2003)	Catalonia	MR Cohort	NR	162/ 3,249 (5.0)	36/ 48 (75)	7/12 (58)	7/12 (58)
				290/ 5,012 (8.4) (5.5 – 12.7)	68/ 94 (72) (56 - 84)	26/34 (74) (55 – 87)	9/19 (47) (24 - 71)
All[1]				**5.8%** (2.8 – 11.7)	**74%** (64 – 82)	**67%** (54 – 77)	**56%** (40 – 70)

1 HR = high risk (prevalence of Down syndrome of about 1:20 or higher), MR = moderate risk (prevalence 1:20 to 1:100), LR = low risk (prevalence <1:100 and includes few, if any, screen positive women.

2 Individual rates for the three aneuploidies were not reported, but the overall rate was 58%.

3 14 of 24 trisomy 18/13 had abnormal flow patterns, and the rates for the three aneuploidies were said to be 'nearly equal'.

5.5 Other ultrasound markers in the first trimester

Fetal heart rate, crown rump length, and tricuspid regurgitation have also been reported to be associated with trisomy 18 fetuses in the late first trimester of pregnancy. Studies are subject to the many of the same biases that were discussed earlier (especially trimester of ascertainment, referral, and correlation). No attempt has been made to provide an unbiased estimate of performance. Instead, the effect is described for each marker and a general introduction to the literature is provided. Several other markers were mentioned in the literature, but these had either low predictive ability, few affected fetuses examined, or both (*e.g.*, reversed umbilical vein blood flow, gestational sac volume, choroid plexus cysts, and specific heart defects). These will not be reviewed.

- **Fetal heart rate**: The fetal heart rate is known to vary by gestational age in the late first trimester. In a report from a series of 25,000 high risk pregnancies seen in a referral center, the fetal heart rate (FHR) declined from about 168 beats per minute (bpm) at a CRL of 40 mm to 155 bpm at 80 mm (FHR = 181.41 – 0.324*CRL_in_mm, r= -0.44) (Liao *et al.*, 2000). The group then compared the proportions of Down syndrome (N=554), trisomy 18 (N=219) and trisomy 13 (N=95) fetuses with bpm below the 5th, above the 50th and above the 95th centiles of the unaffected distribution. That distribution had a constant standard deviation of 6.6 bpm. The distribution of FHRs in Down syndrome fetuses was only slightly higher than that found in unaffected fetuses (5.2% < 5th centile, 54% > 50th centile, and 9.7% > 95th centile). This results in less than a two-fold increase in risk of Down syndrome, should the FHR be above the 95th centile. However, the FHR in trisomy 18 fetuses were somewhat lower than expected (19.7%, 39.7%, and 4.5%, respectively). Thus, an FHR below the 5th centile would result in a four-fold increase in risk for a trisomy 18 pregnancy. The distribution was more abnormal for trisomy 13, with those pregnancies having higher FHRs (2%, 90%, and 64%, respectively). Thus, an FHR above the 95th centile would represent a nearly 13-fold increase in risk for trisomy 13.

  One more recent study examined FHR measurements in the context of other ultrasound and serum markers (Kagan *et al.*, 2008b). When the FHR was included with the combined test (maternal age, NT, PAPP-A and free β hCG) at a 5% false positive rate, the increases in detection for Down syndrome, trisomy 18 and trisomy 13 were 0%, -2% and 5%, respectively. This confirms that measuring FHR might be

beneficial only for identifying trisomy 13. Others have also reported FHR in aneuploid pregnancies, although with much smaller numbers. Two studies found higher odds ratios (10-fold and 4-fold) for Down syndrome and elevated FHR (Hyett *et al.*, 1996; Martinez *et al.*, 1998). One of these studies (Hyett *et al.*, 1996), confirmed that trisomy 18 fetuses had lower, and trisomy 13 higher, FHR.

Given that FHR measurements are of real value only for detecting trisomy 13, it is unlikely that they would be routinely measured as part of a screening program.

- **Crown rump length (CRL):** First trimester growth delay was reported in five trisomy 18 fetuses (Lynch and Berkowitz, 1989). All were five days earlier than by LMP dating, and four were more than seven days early. Later, a study found that Down syndrome and trisomy 13 fetuses were not growth retarded, but among trisomy 18 fetuses, 34% were at or below the 5th centile for CRL (Kuhn *et al.*, 1995). A much larger trisomy 18 case-only report (Sherod *et al.*, 1997) found a 1.1 standard deviation reduction in CRL, with 20% below the 5th centile. Another case-only report identified trisomy 18 as having the greatest growth delay, with an OR of 14 for a 20% or more reduction in the expected CRL (Bahado-Singh *et al.*, 1997). Finally, a UK group (Falcon *et al.*, 2005) reported the greatest mean difference in CRL measurement (-12.3 mm) for trisomy 18 fetuses. Smaller CRL differences of -5.9 and -0.8 were reported for trisomy 13 and Down syndrome fetuses.

  All of these studies relied on women reporting reliable menstrual dates, which limited the usefulness of CRL in predicting aneuploidy. However, there is a consistent finding that trisomy 18 fetuses are the most severely growth retarded, followed by trisomy 13. Down syndrome does not appear to be associated with growth retardation at this time in pregnancy.

- **Tricuspid regurgitation (TR):** In a small proportion of fetuses, some blood leaks back into the right atrium during the contraction of the right ventricle. This leakage is called tricuspid regurgitation, as it occurs through the tricuspid valve. During fetal echocardiography, a pulsed wave Doppler study of the tricuspid valve can be used to diagnosis TR. The association between heart defects and chromosome abnormalities is well documented. A UK group used a strict definition of TR (regurgitation lasting at least half of the systole and having a peak of 80 cm/sec or greater) to survey a high risk population (Faiola *et al.*, 2005). Overall, a TR measurement was attainable in 97% of the fetuses and was present in 8.5% of

unaffected pregnancies. Among Down syndrome (N=83) and trisomy 18/trisomy 13 (N=51) fetuses, the corresponding detection rates were 65% and 53%, respectively. There was a positive correlation with elevated NT measurements. That same group updated their work a year later in another dataset (Falcon *et al.*, 2006), with 12 obstetricians performing the ultrasound exams who were trained in fetal echocardiographs. The false positive rate was lower (4.4%), the Down syndrome detect rate was similar (67% of 114 cases), and the detection rate for trisomy 18 fetuses alone was provided (33% of 42 cases). The agreement between the observers was high (kappa = 0.9). They again found a positive correlation between TR and NT measurements.

Another group in the UK reported the results of tricuspid regurgitation in a more routine setting (Kagan *et al.*, 2009b). A lower false positive rate was found (0.9%), and this may be due to the lower proportion of women with elevated NT measurements in this non-referral, but high risk setting (about one Down syndrome per 108). However, detection rates for Down syndrome (56% of 122) were somewhat lower. Trisomy 18 (33% of 36) and trisomy 13 (30% of 20) detection rates were similar. Adding TR to combined testing in a contingent protocol was found to be of little use for either trisomy 18 or trisomy 13, but did improve Down syndrome detection.

Tricuspid regurgitation appears to be associated with all three chromosomal abnormalities, but the strongest association is with Down syndrome. Studies were performed in high risk settings with experienced fetal echocardiologists and may not be suitable for routine practice, where sonographers are not experienced in Doppler heart studies.