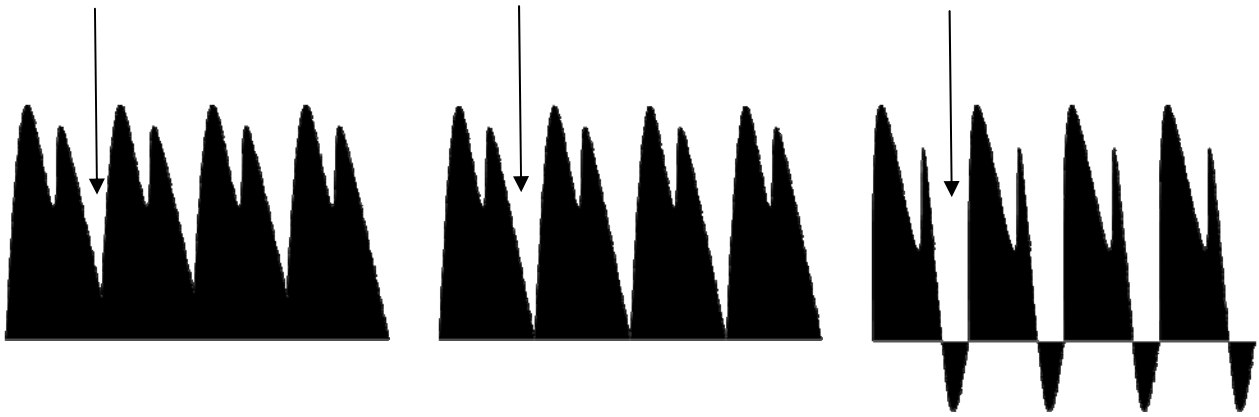
First Trimester
Interlaboratory Comparison Program



Three ductus venosus Doppler waveforms demonstrating differences in the A-wave. 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). Down syndrome fetuses are more likely to have abnormal A-waveforms.

Sponsored by:
Department of Pathology and Laboratory Medicine
Women & Infants Hospital
Providence, RI

Distribution 2011 FT-A

INTRODUCTION

Explanation of Data Listing and Analysis

Specimen Options: The ICP offers two choices of specimens for analysis. One specimen set is designed for those participants using intact/total hCG in their screening marker combination (**hCG sample set**). The other set is designed for those participants using the free beta subunit of hCG in their screening marker combination (**free beta sample set**). The specimens may consist of: 1) unmodified patient pools, 2) patient pools diluted with normal human serum and spiked with recombinant hCG or recombinant free beta subunit (but not both), recombinant inhibin and a PAPP-A concentrate, or 3) normal human serum spiked with recombinant hCG or recombinant free beta subunit (but not both), recombinant inhibin-A and a PAPP-A concentrate. A limited number of additional samples sets are available upon request. This allows participants switching from hCG to free beta (or vice versa) to test both sample sets.

Reading the Data Listing: The five page data listing (in a separate pdf file) contains a summary of reported results for all participants, with each page summarizing one specimen. Your laboratory identification number (ID) is listed at the beginning of the row with your results. Missing data (blanks) are likely due to participants who are manufacturers, or to participants that are not yet offering screening services. Outliers for decimal gestational age (or decimal maternal age) are identified as those outside ±0.2 weeks (or ±0.2 years) of the correct answer. For the assay results (in mass units or MoM) and Down syndrome risks, outliers are defined as being outside of ±2 standard deviations, after accounting for rounding. A revised SD is them computed. A logarithmic transformation is used for the analysis of Down syndrome risks.

Conversion of Reported Down Syndrome Risks to First Trimester Risks: Most participants report risk relevant for the first trimester, but some report risk for the second trimester or term. If the reported risks are not first trimester, these risks are displayed in the column labeled “Report” under the “Down S Risk (1:n)” heading. To allow all risks to be evaluated together, second trimester risks are converted to first trimester risks using the factor 0.74. This accounts for fetal loss between the first and second trimesters (46% from first trimester to term and 23% from second trimester to term). For example, if the second trimester risk is 1:1000, the first trimester risk is 1:[1000 x 0.74], or 1:740. Term risks are converted by multiplying by 0.54 in a similar manner.

Down syndrome risks from participants using the free beta sample set are listed in the data sheets, and may be included in the calculation of summary statistics if the free beta MoM levels are similar to those for hCG. Otherwise, the risks are listed in the “Report” column but not included in the analysis. When sufficient numbers are available, a separate analysis will be performed.

Maternal Age Reporting: Maternal age can be reported either as a decimal or as completed years (integer). Although the difference in Down syndrome risk is small for most ages, use of decimal age can be important for women age 35 and older, especially one whose age falls close to a whole year (*e.g.*, 36.1 versus 36.9 years). Each of these women would be considered to be 36 completed years, even though they are almost one year apart. Participants commonly calculate risk using a maternal age equation rather than a table of risks, and it is straightforward to use the more accurate age to obtain better accuracy. Almost all participants in the ICP report decimal age. Currently, the lab(s) that report integer maternal ages are listed separately on the data summary results. In the future, such results will be listed along with decimal ages but will not be included in the calculations.

NT MoM Reporting: The ICP only provides a target NT MoM for most challenges. Participants need to generate these NT MoM values by trial and error, usually by entering various combinations

of CRL/NT/GA combinations. Approximate CRL values (in mm) and GA values (in weeks and days) are provided as an aid. Participants are asked to report the MoM value that they actually obtained to serve as a check on how reliably they can reproduce the targeted MoM value. Almost all participants report NT MoM values that closely match the targeted value. If participants are having difficulty generating these target NT MoM levels, we can provide assistance.

The ICP also includes at least one challenge each distribution that provides only a patient CRL and NT value (in mm), along with a set of NT and CRL values from the submitting 'hypothetical' sonographer (identified by three initials) who provided those measurements. Participants can then use that set of NT/CRL values to generate sonographer-specific NT medians. Those medians can then be used to convert the reported patient NT value (in mm) to MoM. That NT MoM is then used along with maternal age and the chemistry results to calculate the patient-specific Down syndrome risk. We have provided an Excel spreadsheet that can be used to calculate the CRL/NT median equation with two accompanying quality assurance parameters (*e.g.* slope and log SD). Participants that do not use NT MoM for interpretation will only be evaluated for analyte values.

Greater Than and Less Than Risks: Risks that are reported as less than (<) or greater than (>) are displayed in the "Report" column under the "Down S Risk (1:n)" column. These risks are listed as the actual numeric risk in the "1st trim" column and may be included in the final calculation of the consensus risk.

Free Beta Subunit Results: The data listings include the analyte and MoM values for the free beta measurements for those participants using the free beta specimen set. A median is reported, but a comprehensive analysis is not currently performed, due to the small number of participants. However, each of these participants can review their own results by inspection of the data listing.

If the consensus MoM for free beta is similar to the hCG MoM, the reported risks for free beta users are included in the summary statistics (after converting risks to the first trimester). A close approximation in MoM values is possible for most manufactured specimens (but not all) because advantage is taken of the high correlation between hCG and free beta values (r values of ~0.8). Roughly, absolute free beta values in mg/mL are approximately 40% to 80% of the absolute values of hCG for patient specimens in IU/mL (*e.g.*, 100 IU/mL hCG will typically have a free beta value of 50-60 ng/mL). For some specimens, the relationship for manufactured specimens does not hold, and these are not included in the risk summary statistics. Even if the hCG and free beta MoM values are similar, the risks may differ because the referent parameters differ for the two analytes.

PAPP-A Values: Some participants are now using the Beckman DxI (or Access) assay for measuring PAPP-A. The Beckman assays are calibrated in ng/mL (rather than mIU/mL), which gives absolute values that are approximately 300-350 times larger. PAPP-A results in ng/mL are now listed separately from those in mIU/mL. Some participants reported PAPP-A results in ug/mL (ng/mL divided by 1000) or in ug/dL (ng/mL divided by 10). We converted these results to back to ng/mL as a way to avoid introducing further complexity in the report. Finally, most participants using the PerkinElmer assay report results in mIU/mL, but some report in mIU/L. These values have also been converted to mIU/mL by dividing by 1000 in the report.

Values in Boxes: The ICP uses two types of highlighted boxes in the individual data listings:

- Thick lined boxes identify values that are outliers as compared to the consensus (e.g., 25.0).
- Thin lined boxes are used to call attention to values that are significantly different from the consensus but are not considered outliers (e.g., 1.07). The most common reason is that they are now outside of ±2 trimmed SDs, but were initially within ±2 untrimmed SDs.

RESULTS

PAPP-A mass and MoM values (All specimens):

Values. Among the 30 participants, 20 report results in ng/mL (or equivalent) and 10 report in mIU/mL (or equivalent). Over time, the relationship between these units has ranged from 300 to 350. The conversion factors for the five current specimens are 324, 367, 352, 353, and 348, respectively (*i.e.,* multiply the factor by mIU/mL to obtain ng/mL). Separate analyses are performed for each group. The CVs for the participants reporting in mIU/mL are higher (20% to 29%) than for those reporting in ng/mL (5 to 8%). This difference may be because all participants reporting in ng/mL are using a single manufacturer's reagents, while multiple manufacture reagents are included in participants reporting in mIU/mL. A this time, there are insufficient numbers to perform method-specific analyses.

MoM. Medians should, at least in part, normalize for the differences observed for PAPP-A values. However, the CV of MoM values for PAPP-A in this distribution range from 12% to 22%. The CV of PAPP-A MoM values have been relatively high historically, as compared to corresponding values observed for hCG. This relationship is changing, however, as the newer, higher precision methods begin to dominate. This assumes that participants generate their own reliable median values. It will also be of interest to see whether differences in between-kit mass values are proportional over the range of values, (*e.g.,* differences in values attributable only to calibration differences). We suggest that ICP participants also review their MoM results in the context of other users of their method.

hCG mass and MoM values (All specimens):

Values: The all method CVs for the hCG values are typically low, and this distribution is no different (range 9% to 11%). Although systematic between-kit differences may exist, they are likely to be small.

MoM: The all-method CVs for MoM values range from 9% to 14% for the five specimens in this distribution. This is almost as precise as the mass values, indicating that participants have developed reliable kit-specific population-specific medians.

Free beta mass and MoM values (All specimens):

Values: The number of participants using free beta subunit measurements (all use PerkinElmer assays) is too few to allow a separate analysis (mean, SD, CV). However, from visual inspection of the data it is evident that the between-participant agreement is good.

MoM: As is true for free beta mass values, the number of participants using free beta subunit measurements is insufficient to allow a separate analysis (mean, SD, CV). However, the limited data suggests that the between-participant agreement is good.

Down Syndrome Risk (All specimens):

The consensus trimmed risks for the five specimens in this distribution, ranked from highest to lowest are 1:5 (FT-01), 1:57 (FT-02), 1:220 (FT-04), 1:370 (FT-05), and 1:1050 (FT-03). The CVs of the log risk for these ranked risks are 39%, 17%, 10%, 12%, and 8%, respectively. The CVs decrease almost monotonically as risks decrease. The high CV of 39% for FT-01 reflects the fact that the reported risks are very high, usually with only a single digit of precision.

Free beta results: Down syndrome risks for free beta subunit users are included in the summary statistics except for FT-03 where the consensus MoM for free beta users is 3.88 versus 0.78 for hCG. This results in a much higher Down syndrome risk for free beta users.

Calculation of gestational age and NT MoM Exercise (FT-13 and FT-13fb):

Participants were asked to calculate an NT MoM value, given a CRL of 56 mm (about 12.3 weeks' gestation) and an NT value of 1.3 mm. Results were submitted by sonographer identified as "GNG". Participants were provided with a set of 150 NT/CRL measurements for sonographer GNG (participants may have already calculated sonographer-specific medians for this sonographer in a previous exercise). However, participants may or may not have used those medians to calculate their MoM value, depending on their own laboratory protocols. The expectation is that the resulting MoM values reported by participants that use sonographer-specific medians should be similar, while those using a single fixed set of NT medians might be different. We calculated the median equation for sonographer GNG to be:

$$\text{median NT} = 10^{(-0.239+0.00724*\text{CRL})}$$

using the ICP Excel calculator provided to all subscribers. This equation yields an expected median NT value of 1.47 mm for a CRL of 56 mm, which results in a NT MoM value of 0.88 (1.3/1.47). The trimmed mean NT MoM value is 0.89; nearly identical to the expected value.

In contrast to previous exercises (*e.g.,* 2010 FT-01), the 2011 FT-01 NT challenge found that all of the reported NT MoM values showed good agreement. Those participants with outlying NT MoM values in the earlier distribution had indicated that either they did not use sonographer-specific medians, or they did not know the source of their NT medians. The use of a single set of medians is a likely explanation why they had outlying NT MoM results compared to the consensus. We specifically designed data for sonographer GNG to allow for agreement between all participants in NT MoM so that the focus would primarily be on the serum measurements. GNG's measurements correspond to the median values advocated by the Fetal Medicine Foundation. That group assumes that all credentialed sonographers would achieve the same NT measurements for any given woman. Although a laudable goal, it may not always be achieved in practice. The use of sonographer-specific medians could be viewed as an interim step in for practitioners who have not yet met this goal.

Using the CRL of 56 mm, participants reported a gestational age in the range of 11.9 to 12.3 weeks (consensus 12.1 weeks). These small differences are likely attributable to the use of different conversion equations (see the 2009 FT-A report that includes an analysis of the 'CRL to decimal weeks' equation reported by each laboratory).

Dimeric inhibin-A (DIA)

First trimester DIA measurements were reported by five participants (Table 1). DIA mass values show excellent between-laboratory agreement, but MoM levels were more variable. Included in Table 1 are the DIA likelihood ratios (LR), in the context of the other markers.

Table 1. Dimeric Inhibin-A results and comparison of DS risks including, and omitting DIA

Sample No.	Lab	Method	Value[1]	MoM	DS Risk (1:n) including DIA	DS Risk (1:n) omitting DIA[2]	DIA LR[3]
FT-01	A	Beckman DxI	344	1.29	10	10	1.00
	B	Beckman DxI	297	1.21	6	6	1.00
	C	Beckman DxI	295	0.91	25	10	0.40
	D	Beckman DxI	336	1.07	14	6	0.43
	E	Beckman Access	331	-	-	-	-
FT-02	A	Beckman DxI	225	1.10	58	43	0.74
	B	Beckman DxI	223	0.85	180	100	0.56
	C	Beckman DxI	245	0.76	183	105	0.57
	D	Beckman DxI	231	0.68	184	86	0.47
	E	Beckman Access	222				
FT-03	A	Beckman DxI	207	0.83	380	280	0.74
	B	Beckman DxI	190	0.96	1400	510	0.36
	C	Beckman DxI	196	0.83	5290	1810	0.34
	D	Beckman DxI	203	0.83	5510	2310	0.42
	E	Beckman Access	200				
FT-04	A	Beckman DxI	162	0.76	250	150	0.60
	B	Beckman DxI	156	0.59	660	300	0.45
	C	Beckman DxI	160	0.48	496	298	0.60
	D	Beckman DxI	159	0.47	516	230	0.45
	E	Beckman Access	162				
FT-05	A	Beckman DxI	325	1.50	85	100	1.18
	B	Beckman DxI	302	1.59	170	120	0.71
	C	Beckman DxI	302	1.14	648	548	0.85
	D	Beckman DxI	314	1.35	569	441	0.78
	E	Beckman Access	315				

1 Rounded value in ng/mL

2 DS risk reported for NT, PAPP-A, and hCG

3 For each participant, the DIA likelihood ratio (LR) is computed by dividing the reported risk for NT, PAPP-A and hCG by the risk that also includes DIA measurements. If blank, the LR cannot be reliably determined, usually because one (or both) of the risks are very high (*e.g.,* >1:10) or very low (*e.g.,* <1:10,000).

Interpretive Questions: Additional First Trimester Ultrasound Markers

Q1. Does your laboratory provide clinical results for Down syndrome screening?
A total of 27 participants responded "Yes". The following analyses are restricted to these participants.

Q2. Excluding NT, does your lab incorporate any ultrasound findings when calculating Down syndrome risk?
Five of the 27 participants reported using other measurements. One other participant indicated they use CRL. This is mainly used to calculate gestational age for interpreting chemistry and NT results. Since the question was intended to focus on other markers of Down syndrome, this result was excluded.

Q3. Which of the following US markers do you allow (use)?
Among the five participants reporting the use of an additional ultrasound marker (19%), all reported using the presence or absence (or hyperplasia) of the nasal bone (19%). Interestingly, measurements of tricuspid regurgitation and/or ductus venous flow were not reported by any of the participants.

Q4. If you use nasal bone findings in your DS calculation, enter the revised Down syndrome risks for specimen FT-04.

Table 2. Down syndrome risk with the presence or absence of the nasal bone.

	Combined Test	Nasal bone present		Nasal bone absent	
Lab	Risk (a)	Risk (b)	LR (a/b)	Risk (c)	LR (a/c)
B	1: 250	1: 760	0.33	1: 15	17
A	1: 300	1: 940	0.32	1: 17	18
D	1: 310	1: 950	0.33	1: 18	17
C	1: 434	1:1330	0.33	1: 23	19
E	1:1900	1:6000	0.32	1:110	17
Median	1: 310	1: 950	0.33	1: 18	17

Discussion:

The primary first trimester ultrasound measurements used in risk calculations for the combined or integrated test(s) are crown rump length (CRL) and nuchal translucency (NT). CRL is a used as a surrogate for gestational age, and is used to generate an equation for calculating gestational-age specific median values. Several other first trimester ultrasound measurements have been described that could be incorporated into the risk calculations for the 'Combined Test' [NT, PAPP-A, free beta subunit (or hCG)] to improve performance. The most common of these are absence/presence of the nasal bone, ductus venosus flow, and tricuspid regurgitation. A cartoon representing normal and abnormal ductus venosus flow is show on the title page of this report. The detection rate (DR) for these markers has been recently summarized as 60%, 66%, and 55% with false positive rates (FPR) of 2.5%, 3.0 %, and 1.0%, respectively (Nicolaides *et al.*, *Prenat Diagn* 2011;31:7-15). These results could be used to calculate likelihood ratios (DR/FPR) that would increase or to decrease the Down syndrome risk based on the 'Combined Test'.

For example, assume the first trimester DS risk (NT, PAPP-A, and hCG) is 1:1000. The likelihood ratio for absent nasal bone is 24 (60% DR/2.5% FPR). The revised risk with

absent nasal bone is then 24:1000 or 1:42. The likelihood ratio for present nasal bone is 0.41 ([100%-60%]/[100%-2.5%]). The revised risk with present nasal bone is then 0.41:1000 or 1:2400.

These calculations assume that the nasal bone marker is independent of NT and gestational age. In practice, nasal bone absence occurs more frequently with increasing NT (Rosen *et al., Obstet Gynecol* 2007;110:399-404). This correlation reduces the effectiveness of the taking both markers into account as done in the previous paragraph. A further finding is that the nasal bone is absent more frequently at earlier gestational ages in fetuses without Down syndrome (euploid), which would yield a higher false positive rate than later in gestation. Lastly, nasal bone hyperplasia is more common in women of African descent and Southern Asian descent. These covariates require the use of multivariate models to generate optimum DS risk, similar to those used for serum markers to account for race and maternal weight.

The five participants using nasal bone to modify DS risk reported an LR of about 0.33 when the nasal bone was present, compared to an LR of about 18 when it was absent. Participants were not asked about the source of their data for nasal bone calculations, but their reported likelihood ratios are similar to those provided in the earlier example.

Interpretive question: Integrated screening

Twenty-one participants reported integrated risks using first trimester marker results (FT-01) in combination with the second trimester quadruple markers results (CAP FP-02). Seven participants (including several manufacturers) did not report integrated risks. One participant did not respond. All participants now report integrated risks using the second trimester quadruple markers. Some participants also report a risk for the triple test, but that is not analyzed. Seventeen participants report risks in the second trimester; five report risks at term. The term risks are adjusted to the second trimester using a 0.74 survival coefficient (*e.g.*, a term risk of 1:1000 equals a second trimester risk of 1: 740). Table 2 lists these Down syndrome risks (1:n), along with the trimester of risk. The last column in each category contains the risk after adjustment to the second trimester, to allow for direct comparison. Participants using free beta subunit are included in the summary statistics. One full integrated result was identified as an outlier.

Table 2. Summary of Integrated Down Syndrome Risks

Trimester of Risk	Quad Risk (FP-02) Reported	Quad Risk (FP-02) Adjusted	Serum Integrated Risk Reported	Serum Integrated Risk Adjusted	Full Integrated Risk Reported	Full Integrated Risk Adjusted
2	520	**520**	150	**150**	2100	**2100**
2	560	**560**	43	**40**	610	**610**
2	7900	**7900**	6300	**6300**	20000	**20000**
2	1800	**1800**	920	**920**	18000	**18000**
2	2505	**2500**	130	**130**	1600	**1600**
2	445	**450**	178	**180**	2318	**2300**
2	1200	**1200**	370	**370**	5600	**5600**
2	5000	**5000**	2700	**2700**	5000	**5000**
2	690	**690**	85	**85**	990	**990**
2	546	**550**	232	**230**	1920	**1900**
2	1300	**1300**	390	**390**	7100	**7100**
2	1100	**1100**	300	**300**	4500	**4500**
2	840	**840**	140	**140**	2000	**2000**
2	2100	**2100**	190	**190**	3000	**3000**
2	775	**780**	30	**30**	270	**270**
3	1000	**700**	95	**70**	1100	**800**
2	440	**440**	40	**40**	330	**330**
3	6900[1]	**5100[1]**	9800[1]	**7300[1]**	20000[1]	**14800[1]**
3	7300	**5400**	10000	**7400**	100000	**74000**
3	1000	**700**	230	**200**	3300	**2400**
3	7300[1]	**5400[1]**	10000[1]	**7400[1]**	<20000[1]	**<14800[1]**
Trimmed		None		None		74000
Geo Mean		1400		400		2800
CV(log risk)		13%		31%		16%
Mean-2SD		200		10		230
Low		440		30		270
High		7900		7400		20000
Mean+2SD		8400		12000		33000

[1] Participants using free beta subunit rather than total/intact hCG measurements

The trimmed geometric mean risk for the second trimester quadruple test for ICP participants is 1:1400 (similar to the consensus risk of 1:1510 for FP-02 on CAP FP-A report). The risk's CV is relatively high compared to previous distributions. This may reflect the unusual pattern of marker MoMs for specimen FT-02 (consensus MoM values of 1.71, 0.99, 2.48.and 1.09 for AFP, uE3, hCG, and DIA, respectively). In particular, the hCG MoM is elevated while the inhibin MoM is close to the unaffected median. The levels of hCG and inhibin are the most correlated of the second trimester markers and these results would therefore be considered unusual. The trimmed mean PAPP-A MoM value for specimen FT-01 is 0.52. The crossover point (likelihood ratio of 1.0) for affected and unaffected PAPP-A distributions is approximately 0.6 MoM. Thus, the expectation is for an increase in the serum integrated risk compared to the quadruple risk. The geometric mean for the serum integrated risk of 1:400 for ICP participants versus 1:1400 is consistent with this expectation (LR of 3.5). Note, however, that the CV of 32% for the adjusted log risk is very high. This likely reflects the fact that the PAPP-A MoM of 0.52 falls at the lower extreme of the distribution of values, thereby introducing more volatility into the likelihood ratio used in the risk calculation. The trimmed mean NT MoM value

calculated by ICP participants for FT-01 is 0.88. The crossover point for the affected and unaffected distributions of NT MoM is approximately 1.45 MoM. The expectation is, therefore, that full integrated test risk would be significantly reduced, and the lowered geometric mean risk of 1:2800 versus 1:400 for the serum integrated test meets expectation (LR of 0.14).

George J. Knight, Ph.D.	Glenn E. Palomaki, B.S.	(207) 894-6610
Jacob A. Canick, Ph.D.	Geralyn M. Messerlian, Ph.D.	(401) 453-7650

Department of Pathology and Laboratory Medicine
Women & Infants Hospital
Providence, Rhode Island