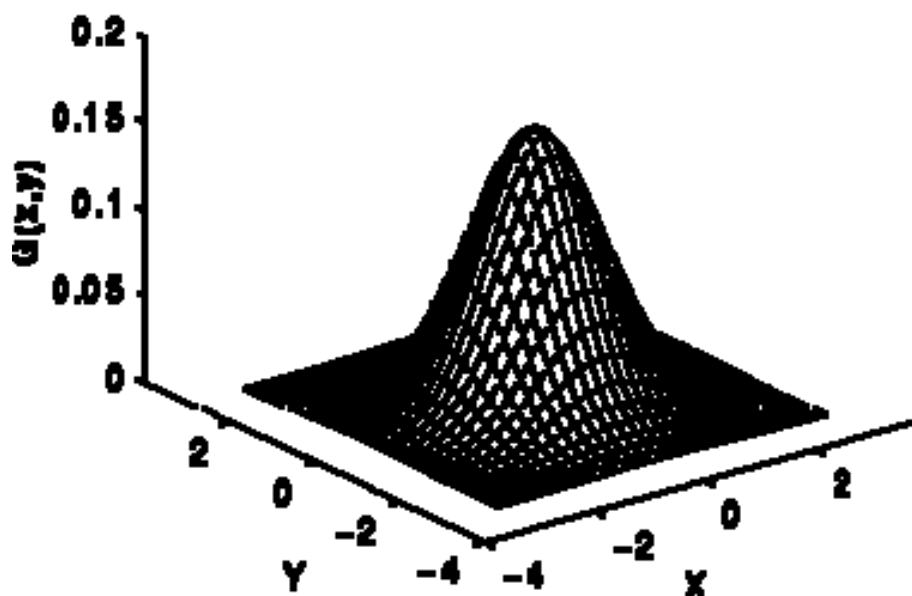


First Trimester Interlaboratory Comparison Program

Distribution 2009 FT-A



A two-dimensional Gaussian distribution. This might represent the distribution of PAPP-A and hCG measurements in unaffected pregnancies. Notice the center is at 0,0 (or 1,1 without the logarithmic transformation). Also, it seems to have a low correlation (haystack) rather than a high correlation (ridge).

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

INTRODUCTION

Explanation of Data Listing and Analysis

Specimen Options: The ICP now offers two choices for specimens. The first specimen set is designed for those labs using hCG in their screening marker combination (**hCG sample set**). The second set is designed for those labs using free beta subunit in their screening marker combination (**free beta sample set**). The specimens 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) and PAPP-A concentrate (and inhibin), or 3) normal human serum spiked with recombinant hCG or recombinant free beta subunit (but not both) and PAPP-A concentrate.

The PAPP-A target concentrations in both sets are the same for all five specimens, allowing for a unified evaluation for both specimen sets. However, participants should not test or report hCG measurements made in the free beta set, nor free beta measurements in the hCG set. This can result in spurious results for the following reasons:

- *In hCG specimen sets:* specimens spiked with recombinant hCG can yield a very high non-physiologic level of measured free beta subunit (the recombinant hCG preparation has a significant amount of free beta present).
- *In free beta specimen sets:* some free beta specimens are spiked with recombinant free beta subunit and others are made by spiking normal human serum with recombinant free beta subunit and PAPP-A concentrate. However, no recombinant hCG will have been added to either type of sample, so measurement of these specimens for hCG is not appropriate.

A limited number of additional samples sets are available upon request (free of charge) so that laboratories considering switching from hCG or vice versa can receive both the hCG and the free beta sample sets.

Reading the Data Listing: The five page data listing (attached) contains a summary of reported results for all participants, with each page summarizing one specimen. Your lab ID is listed at the beginning of the row with your results. Missing data (blanks) are likely due to participants who are manufacturers rather than screening labs, or to laboratories that are not yet offering screening services. Outliers for gestational age (or maternal age) are identified as those outside +/- 0.2 weeks (or 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 trimmed standard deviations, after accounting for rounding. A log transformation is used for the analysis of Down syndrome risks.

Conversion of reported risks to first trimester risks. Most all laboratories report first trimester risks, but some laboratories report second trimester or term risks. 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 by a single statistic, 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 (43% 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 to first trimester risks by multiplying by 0.57.

Down syndrome risks from participants using the free beta sample set are listed in the data sheets, but are not included in the calculation of the consensus. When sufficient numbers are available, a separate analysis can be performed.

Maternal Age Reporting: Maternal age can be reported either as a decimal or as completed years (integer). Although the difference in risk is small for most ages, use of decimal age rather than completed years can be more significant for a woman whose age falls close to a whole year (e.g., 34.1 versus 34.9 years). Each women would be called 34 completed years even though they are almost one year different in age. Laboratories commonly calculate risk using a maternal age equation rather than a table of risks, and it is straightforward to use the more accurate risk to obtain better precision. Almost all labs in the ICP report decimal age. Currently, lab(s) that report integer maternal ages are listed separately on the data summary results, but 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 provides a target NT MoM for most challenges. Participants need to generate the MoM values provided in the histories by trial and error, usually by entering various combinations of CRL / NT / GA values. Approximate CRL values (in mm) and GA values (in weeks and days) are provided as an aid for this process. Participants are asked to report the MoM value that they actually obtained to serve as a check on how reliably they could reproduce the targeted MoM value. Almost all participants report NT MoM values that closely match the targeted value. If participants are having difficulty generating a reliable MoM, we can provide assistance.

The ICP also includes at least one challenge that provides a patient CRL and NT value (in mm), along with a set of NT and CRL values from the submitting 'hypothetical' sonographer (identified by initials) who has provided those measurements. Participants are asked to use the set of sonographer-specific NT / CRL values to generate an appropriate set of NT medians for converting the NT values (in mm) to MoM. That NT MoM is then used along with the chemistry results to calculate the patient-specific Down syndrome risk. We also provide an Excel spreadsheet that can be used to calculate the CRL / NT median equation with accompanying quality assurance parameters (e.g. slope and log standard deviation).

Labs that do not use the MoM for interpretation of NT 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 table for each of the five specimens lists the analyte and MoM values for the free beta measurements for those laboratories using the free beta specimen set. A median is reported, but a comprehensive analysis is not performed due to small numbers. However, each of these participants can review their own results by inspection of the data listing. Currently, all participants receive the free beta sample set report in term risks and these are listed in the "**Report**" column under the "**Down S Risk**" heading. It is important to remember that these term risks would not be comparable to the risks listed in the "1st trim" column, even if converted to 1st trimester risks, because of the use of free beta subunit rather than hCG in the calculation of risk.

RESULTS

FT-01 and FT-01fb: Participants were asked to calculate an NT MoM value, given a CRL of 75 mm (~ 13.5 weeks' gestation) and an NT value of 1.6 mm submitted by sonographer FST. Participants were provided with a set of 150 NT/CRL measurements for FST and may have already calculated a sonographer-specific median equation (sent again in this distribution for those who may need to recompute the median equation). However, participants may or may not have used those medians to calculate their MoM value, depending on existing laboratory protocols. The expectation is that the resulting MoM values reported by laboratories that use sonographer-specific medians should be similar, while those using a single fixed set of NT medians might be quite different. We calculated the median equation for sonographer FST to be: median NT = $10^{(-0.373+0.00613 \cdot \text{CRL})}$ using the EXCEL calculator supplied to participants. This equation yields an expected median NT value of 1.22 mm for a CRL of 75 mm, which calculates to a MoM value of 1.31 (1.6/1.22). The consensus NT MoM (calculated as the trimmed mean value) value is 1.31, in agreement with the expected value. Three results were significant outliers (0.78, 0.87 and 0.89 MoM). Two of these labs indicate that they do not use sonographer specific medians but instead use a single set of medians, which could account for the discrepant results. The third lab indicates that they use a combination of Sonographer specific and center-specific medians. It is not clear whether this lab used existing medians or attempted to calculate a median equation using the supplied FST measurements, either using in-house software or the calculator supplied by us. In either case it is recommended that they determine the source of the discrepant MoM values. We can assist this lab if needed. Overall, most laboratories can derive a common median equation given sonographer NT/CRL values, and can use those results to provide clinical interpretations.

A CRL was provided for this sample, requiring each lab to calculate gestational age. Assigned gestational ages for FT-01 ranged from 13.1 to 13.7 weeks. As pointed out previously, differences reflect the 'CRL to decimal weeks' equation selected by laboratories. The Supplemental Question in this distribution report addresses this issue and reviews the equations in common use.

The CVs of the PAPP-A values and MoM were very low (8% and 9%, respectively), while the corresponding hCG values and MoM were somewhat higher (12% and 13%, respectively). The low CVs for PAPP-A are noteworthy because typically CVs are in the mid to high teens. This likely reflects the relatively high PAPP-A values which tend to have lower CVs. The CV of log risk was low (8%). This specimen had consensus risk of 1:1400 and all considered it screen negative.

FT-02 and FT-02fb: The CVs for the consensus PAPP-A values and PAPP-A MoMs are higher than the consensus hCG values and MoMs, which are typical for PAPP-A results with values below approximately 3.0 mIU/ml. The consensus risk was very low (1:6700), and all labs reported it to be screen negative

FT-03 and FT-03fb: This specimen is a pool of sera from 12 week pregnancies, and should, therefore, more accurately reflect actual between-lab and between-kit differences than manufactured samples. The CVs for both PAPP-A and hCG values are similar (11% and 14%, respectively). These results are reasonable considering that results have not been stratified by kit manufacturer. The trimmed mean consensus risk was 1:5000, and all labs considered it to be screen negative.

FT-04 and FT-04fb: This specimen was targeted to have a low PAPP-A and a very high hCG value. The trimmed mean consensus values of 0.48 mIU/mL and 365 IU/mL, respectively, were consistent with these targets. The CV of PAPP-A was relatively high (22%), likely because the value was low. The CV of log risk (48%) was significantly higher than is typical for other samples. This reflects the fact that the PAPP-A consensus MoM of 0.32 is very low and the consensus hCG MoM of 3.85 is very high, falling at the extreme of the population distribution of values. Small differences

in MoM values can yield relatively large differences in the likelihood ratios that are used in the risk calculation. Also, most reported risks are in single digits, which also magnifies between-lab differences. The consensus risk was 1:4 and all labs called the specimen screen positive.

FT-05 and FT-05fb: The CV for PAPP-A values was 14%, which is similar to the CV of 11% for specimen FT-03, both of which are patient pools. The risk for this 19.8 year old woman was low (1:550), and all labs called her screen negative.

Dimeric inhibin-A (DIA)

First trimester DIA measurements were reported by four participants (Table 1). Three labs report using Beckman assays, while one uses the DSL ELISA. Table 1 also lists the reported DIA values and MoM levels for each of the five samples. DIA values show good between-method and between-lab agreement. MoM levels also show good between-lab agreement. Included in the table are the DIA likelihood ratios (LR) in the context of the other markers, and again are in reasonable agreement. All four labs report the same first trimester risk with and without DIA for FT-04, yielding a LR of 1.0. This is probably a consequence of the very high risk for the first trimester markers. The risks may be truncated, or the addition of DIA might not impact the already high risk to a measureable extent.

Table 1. Dimeric Inhibin-A results for FT-A 2008

Sample No.	Lab	Method	Value ¹	MoM	DS Risk (1:n)	DIA LR ²
FT-01	A	DSL Elisa	853	3.21	700	2.43
	B	Beckman Dxl	855	3.89	670	4.46
	C	Beckman Dxl	875	3.94	112	6.85
	D	Beckman Access	965	3.94	211	3.70
FT-02	A	DSL Elisa	173	0.69	<10000	-
	B	Beckman Dxl	160	0.63	10000	0.39
	C	Beckman Dxl	166	0.64	22600	0.34
	D	Beckman Access	192	0.63	33900	0.36
FT-03	A	DSL Elisa	277	0.93	4400	0.77
	B	Beckman Dxl	276	1.14	2390	0.66
	C	Beckman Dxl	255	0.94	10700	0.51
	D	Beckman Access	342	1.17	16400	0.63
FT-04	A	DSL Elisa	939	2.47	10	1.00
	B	Beckman Dxl	811	2.69	4	1.00
	C	Beckman Dxl	823	2.28	4	1.00
	D	Beckman Access	930	2.54	7	1.00
FT-05	A	DSL Elisa	306	1.60	380	0.89
	B	Beckman Dxl	302	1.52	257	0.65
	C	Beckman Dxl	299	1.50	870	0.62
	D	Beckman Access	351	1.47	1110	0.75

¹ Rounded value

² For each participant, the DIA 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 likelihood ratio 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).

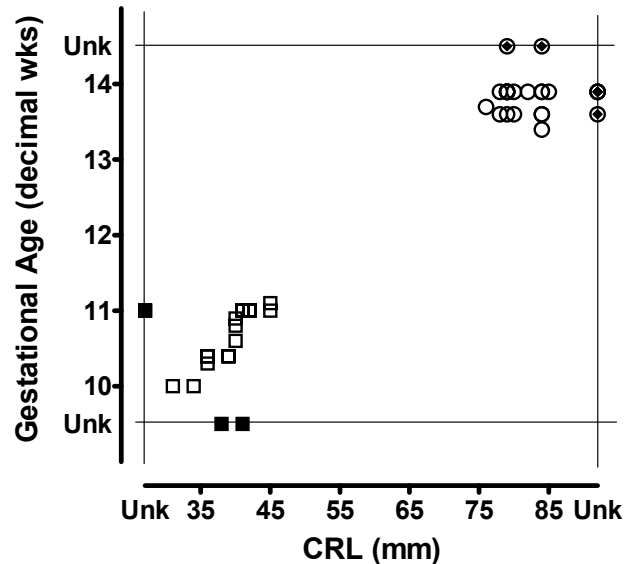
Supplemental Questions: CRL and gestational age

Q1. Do you routinely receive NT results as part of first trimester screening. Of the 27 participating laboratories, 24 reported that they did receive NT measurements. The remaining three are likely to be manufacturers. The subsequent analysis will be restricted to these 24 participants.

Q2. Estimate the percent of NT results having an associated CRL measurement. Overall, 21 reported 100%, 2 reported 99% and the last reported 95%. Clearly, the vast majority of samples with an NT measurement also have an associated CRL measurement, but at some laboratories, this appears not to be mandatory.

Q3. The gestational age range of acceptable NT measurements is defined by CRL, GA, our GA using reported CRL or other. Nine laboratories (38%) use the CRL measurement, one laboratory (4%) uses the GA provided by the physician, 13 (54%) use a GA computed within the laboratory using the reported CRL measurement, and one reports that the sonographer only sends samples in the appropriate range. A fixed CRL limit (response 1), or a fixed GA limit based on a single conversion equation (response 3) are essentially equivalent. Thus, 92% of laboratories used fixed limits that are directly, or indirectly, based on the reported CRL.

Q4/5. The earliest/latest our laboratory accepts an NT measurement is (response in CRL mm or GA in weeks). Most labs reported cut-offs in both CRL and GA, but some reported only in mm or decimal weeks. The figure shows graphically, the responses from the 24 participants. Each open symbol corresponds to a matched set of GA/CRL-mm for the lower (squares) or upper (circles) boundaries of acceptable measurements. The filled symbols show the results for participants that reported only the cut-off in mm or in decimal weeks. There is more variability in the lower cut-off level (range 31-45) when compared to the higher one (range 76-85mm), but a clear relationship between the CRL and GA is observed. Such a relationship is not seen for the upper cut-off level. For example, there are 6 laboratories reporting a cut-off level equivalent to 13.9 weeks, but the corresponding CRL measurements vary from 78 to 85 mm.



Q6/7. Provide the corresponding gestational age your laboratory assigns for CRL measurements between 35 and 95 mm. This figure shows the conversion of CRL to GA (in days) for each of the laboratories, stratified by the reference that they reportedly use. Ten reported using Hadlock et al., [Radiology 1992;182:501-5](Figure 2a), three used Robinson & Fleming [Br J Obstet Gynaecol 1975;82:702-10] (Figure 2b), and two used Daya et al. [Am J Obstet 1993;168:908-8] (Figure 2b). The remaining 10 laboratories reported various other methods/references, including two unknown methods (Figure 2c). Among the participants reportedly using Hadlock et al., one center seems to be lower than expected at higher CRL measurements. Figure 2d shows all data together, retaining the line styles from the earlier figures. The choice of the conversion method is of less importance than ensuring that a consistent equation is used when generating NT median values that are center or sonographer specific.

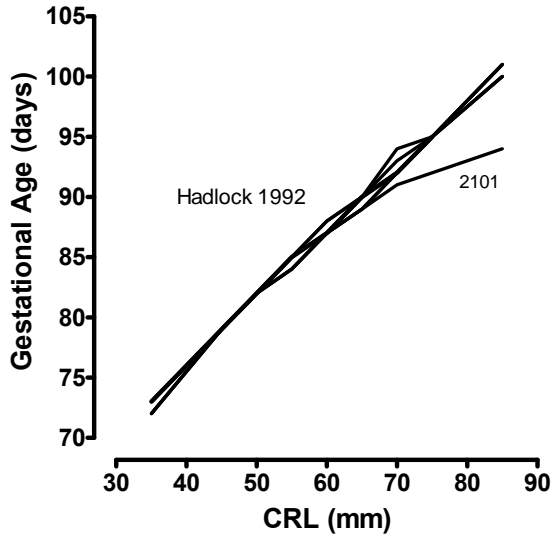


Figure 2a

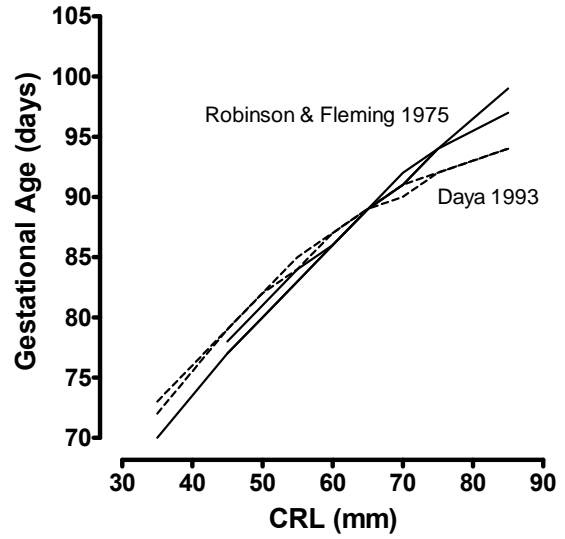


Figure 2b

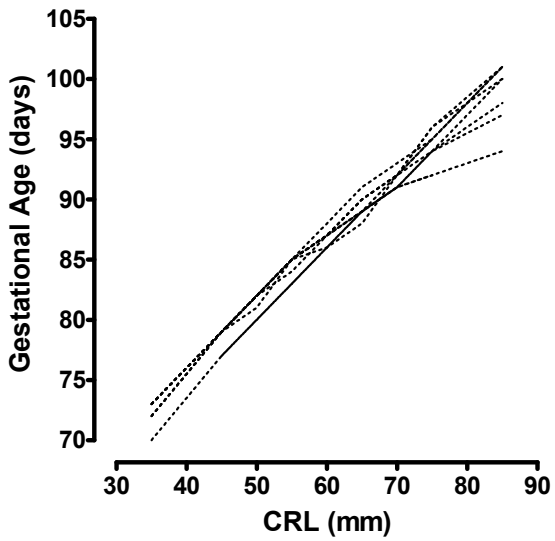


Figure 2c

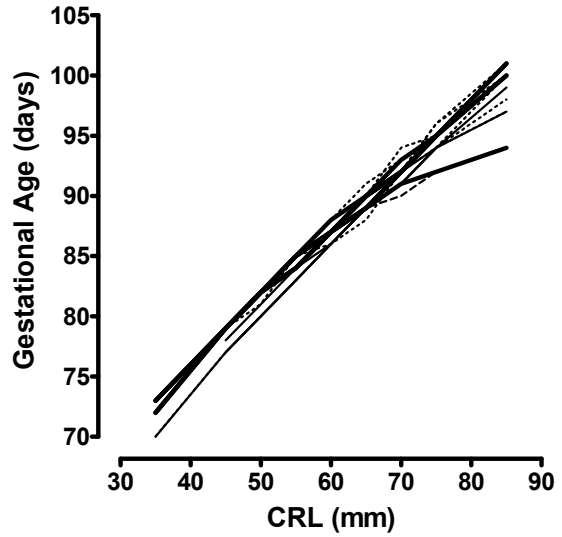


Figure 2d

Interpretive Questions – Integrated Screening for Down Syndrome.

Q8. Does your laboratory perform integrated risk interpretations? Among the 21 participants, 16 reported offering integrated screening as part of a formal program; the five others do not. Of the 16, all use the quad test as their standard second trimester test for the serum integrated test, but only four use it in their full integrated test. These latter four use the triple test instead. In this exercise, the second trimester dimeric inhibin-A (DIA) value for the CAP FP-04 sample is very high (~2.5 MoM). Consequently, the labs using the triple test calculate much higher risks for the full integrated test (at least by a factor of 10) compared to those using the quad test. The full integrated test results can therefore not be compared for these two groups. For this reason, only the risks for the quad and serum integrated tests are displayed in the table for these labs. All four labs using the triple test report term risks and these have been converted to second trimester in the table with the corresponding term risk in parentheses. Risks for these four labs along with the LRs for the “**Q/SI**” are displayed in the table for informational purposes and are not included in the summary statistics.

Q9/10. Using the Down syndrome risks from FP-04 (CAP FP-A 2009 Survey) report integrated risks using FT-03 results (after appropriate modifications to the draw date). Analyses are restricted to the 12 participants reporting full integrated risks using the quad test (Table 3). Columns 1, 2, and 3 display the individual and consensus risks for the quad (Q), serum integrated (SI), and the full integrated (FI) tests. The next three columns examine how the individual laboratories modified the risk based on the addition of markers. The “**Q/SI**” column examines the impact of adding first trimester PAPP-A to the quadruple test, the “**Q/FI**” examines adding both first trimester PAPP-A and NT measurements to the quadruple test, and “**FI/SI**” looks at the impact of adding NT to the serum integrated test.

- (Q)** The trimmed mean second trimester quad markers risk (column 1) was **1:35** with no outlying risks.
- (Q/SI)** The median reported serum integrated risk (column 2) was **1:89**, with one outlier of 1:550. Given the consensus PAPP-A MoM value of 1.01, the risk would be expected to be reduced when included with the quad test. The consensus likelihood ratio of **0.37**, obtained by dividing the quad test risk by the serum integrated risk (**Q/SI**) is consistent with this expectation (column 4). The lab with the risk outlier of 1:550 also had an outlier LR of 0.07, attributable to this risk outlier. The risks for both the quad and serum integrated tests for the four labs who reported term risks (converted to second trimester risks in the table) are notably lower than those labs reporting second trimester risks. However, their LRs are reasonably consistent with the consensus LR of 0.37, indication that some systematic factor is responsible for the lower risks for both the quad and SIT tests.
- (Q/FI)** The median consensus risk for the FI test is **1:91** (column 3) with one outlier of 1:1080. Given the consensus risk of **1:35** (column 1) for the quad test the expectation is that the LR should be about **0.38**. The consensus LR is **0.39**, consistent with expectation. One lab was a LR outlier (0.03).
- (FI/SI)** The consensus NT MoM of **1.40** for FT-03 is close to the cross-over point (about 1.45 MoM) of the overlapping distributions of unaffected and affected pregnancies. This should yield a likelihood ratio close to 1.0. Since NT is not correlated with serum markers, the addition should result in little change in the risk. The median FI risk of **1:91** is virtually identical to the SIT risk of **1:89**. The trimmed mean LR for the quad risk divided by the full integrated risk is **1.15**, close to expectation.

Q11. Do you use the same parameter sets for both the quadruple and integrated test? Sixteen participants responded. Eleven (69%) reported that they do use the same parameter set (the preferred methodology); four reported that they did not (25%) and one did not know (6%).

Table 3. Comparison of quadruple (triple) Down syndrome risks to various integrated risks

Down syndrome risk (1:n) ¹			Likelihood ratio ²		
Quadruple (Q) (FP-04)	Serum Integrated (SI) (FP-04 & FT-03)	Full Integrated (FI) (FP-04 & FT-03)	Q/SI	Q/FI	FI/SI
30	210	200	0.14	0.15	0.95
39	40	38	0.98	1.03	0.95
36	550	1080	0.07	0.03	1.96
25	80	120	0.31	0.21	1.50
63	85	167	0.74	0.38	1.96
37	92	113	0.40	0.33	1.23
38	87	73	0.44	0.52	0.84
25	NR ³	30		0.83	
30	100	103	0.30	0.29	1.03
35	110	95	0.32	0.37	0.86
60	110	110	0.55	0.55	1.00
20	70	65	0.29	0.31	0.93
35	89	91	← Trim Mean		
		outlier	0.07	0.03	.None
		Trimmed LR	.039	0.39	1.15
		log SD of LR	0.24	0.25	0.14
		Mean - 2sd	0.14	0.13	0.62
		low	0.14	0.15	0.84
		high	0.98	1.03	1.96
		Mean + 2sd	1.13	1.18	1.00

Results for Labs Using the Triple Test

67 (90)	178 (240)	0.38
67 (90)	318 (450)	0.21
126 (170)	318 (450)	0.41
89 (120)	252 (340)	0.35

¹ Term risks converted to second trimester risks using the factor 0.74. Corresponding term risks listed in parentheses adjacent to second trimester risks

² Derived by dividing the associated Down syndrome risks

³ Not Reported

George J. Knight, Ph.D.
Jacob A. Canick, Ph.D.

Glenn E. Palomaki, B.S.
Geraldyn M. Messerlian, Ph.D.

(207) 894-6610
(401) 453-7650

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