First Trimester
Interlaboratory Comparison Program

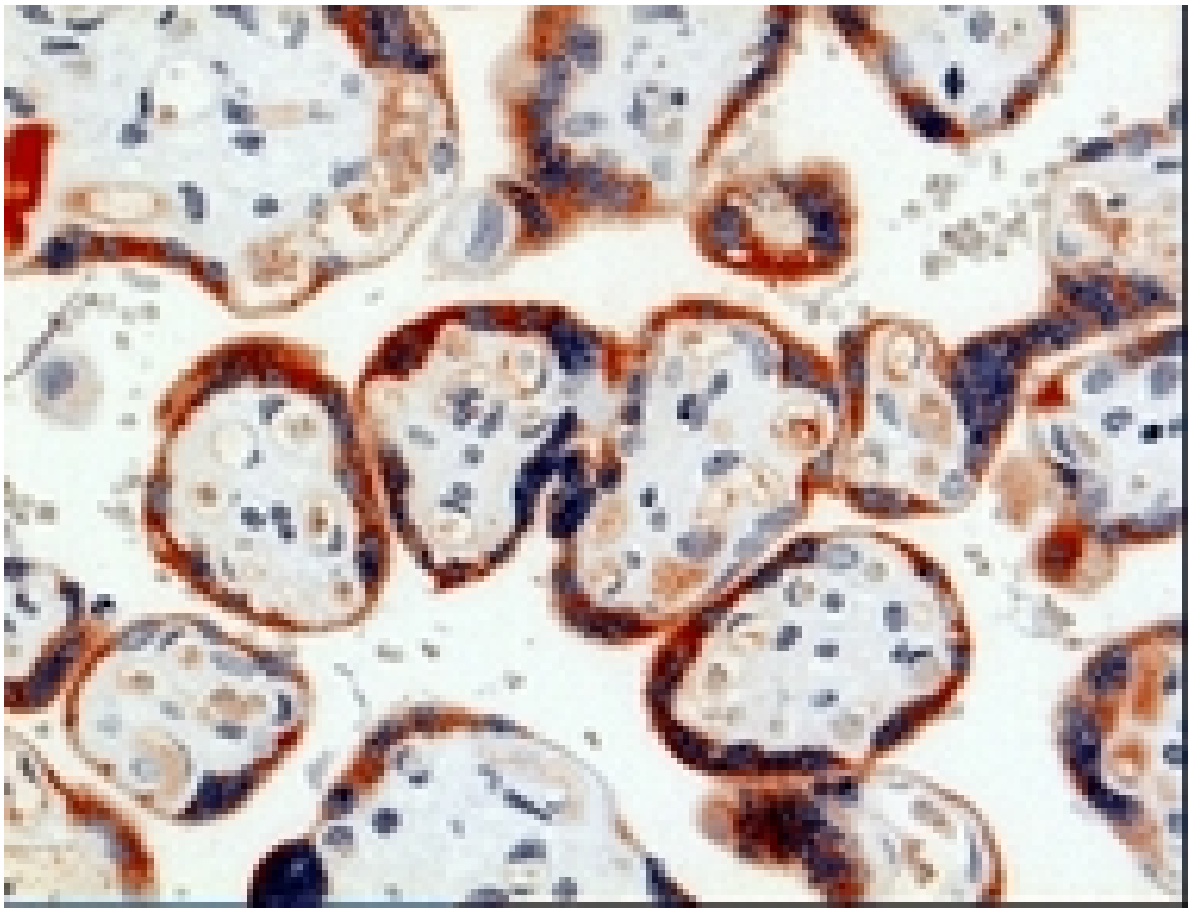

Figure: Immunohistochemistry staining with human PAPP-A antibody in human placenta. Staining (red-brown ring) is restricted to the syncytiotrophoblast of the chorionic villi (www.antibodyship.dk).

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

Distribution 2011 FT-B

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 then 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 at 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 (and interpretations) from participants using the free beta sample set are listed, 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 a more accurate risk. Almost all participants in the ICP report decimal age. Currently, the lab(s) that report integer maternal ages are listed separately. In the future, such results will be listed along with decimal ages but will not be included in the calculations.

NT MoM Reporting: For most challenges, the ICP only provides a target NT MoM. 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. For example, >1:10 indicates a risk that is higher than 1:10.

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 limited 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 would be excluded from 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: Many 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). The ICP converts these results 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 28 participants, 17 report results in ng/mL (or equivalent) and 11 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 more variable (400,501,395,368,369 respectively (*i.e.,* multiply mIU/mL by the factor to obtain ng/mL). Separate analyses are performed for each group. The CVs for the participants reporting in mIU/mL are higher (18% to 20%) than for those reporting in ng/mL (6% to 9%). This difference may be because all participants reporting in ng/mL are using a single manufacturer's reagents, while multiple manufacturer's reagents are included in participants reporting in mIU/mL.

MoM. Medians should, at least in part, normalize for the differences observed for PAPP-A values. 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.

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 13%). If systematic between-kit differences exist, they appear 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 high.

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 also high.

Down Syndrome Risk (All specimens):

The consensus trimmed risks for the five specimens in this distribution, ranked from highest to lowest are 1:160 (FT-09), 1:180 (FT-06), 1:400 (FT-07), 1:1390 (FT-10), and 1:1780 (FT-08). The CVs of the log risk for these ranked risks are 7%, 10%, 8%, 5%, and 10%, respectively. The CVs are relatively similar across the range of risks, which is in contrast to previous distributions where CVs decreased almost monotonically as risks decrease. This may be reflective of the narrow range of risks in this distribution.

Free beta results: Down syndrome risks for free beta subunit users were only included for the first specimen. For the remaining four specimens, the values are listed for inspection.

Calculation of gestational age and NT MoM Exercise (FT-10 and FT-10fb):

Participants were asked to calculate an NT MoM value, given a CRL of 59 mm (about 12.3 weeks' gestation) and an NT value of 1.6 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 1.54 mm for a CRL of 56 mm, which results in a NT MoM value of 1.04 (1.6/1.54). The trimmed mean NT MoM value is 1.04; identical to the expected value. One participant reported an outlier value 1.14.

In contrast to earlier exercises, <u>all</u> of the reported NT MoM values show 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 59 mm, all participants reported a gestational age of either 12.3 or 12.4 weeks (one lab reported an outlying value of 11.3 weeks). These small differences for the majority of results 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 participant).

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-06	A	Beckman DxI	562	2.43	200	200	1.00
	B	Beckman DxI	598	2.08	190	180	0.95
	C	Beckman DxI	567	1.53	226	171	0.76
	D	Beckman DxI	596	1.62	335	280	0.84
	E	Beckman Access	538				
FT-07	A	Beckman DxI	259	1.27	400	330	0.83
	B	Beckman DxI	269	0.99	760	460	0.61
	C	Beckman DxI	261	0.80	637	400	0.63
	D	Beckman DxI	273	0.80	992	523	0.53
	E	Beckman Access	287				
FT-08	A	Beckman DxI	219	0.86	880	770	0.88
	B	Beckman DxI	220	1.02	1800	770	0.43
	C	Beckman DxI	205	0.83	7900	3110	0.39
	D	Beckman DxI	246	0.98	5150	4400	0.85
	E	Beckman Access	239				
FT-09	A	Beckman DxI	378	1.77	150	120	0.80
	B	Beckman DxI	364	1.34	700	210	0.30
	C	Beckman DxI	322	0.97	392	341	0.61
	D	Beckman DxI	336	0.99	315	184	0.58
	E	Beckman Access	403				
FT-10	A	Beckman DxI	312	1.19	950	900	0.95
	B	Beckman DxI	314	1.27	3500	880	0.25
	C	Beckman DxI	294	0.97	5510	2860	0.52
	D	Beckman DxI	327	1.09	3180	3080	0.97
	E	Beckman Access	363				

[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).

Interpretative Questions: Racial adjustment of First Trimester Analyte Values

Of the first trimester markers, PAPP-A is the one most frequently adjusted for race, either by applying an adjustment factor or by generating race-specific median values.

Four of 27 (15%) centers using first trimester hCG also reported using a racial adjustment. No laboratories using free beta hCG reported using a racial adjustment, and no laboratories reported racial adjustments for the ultrasound markers nasal bone or NT.

Conclusions: PAPP-A is the first trimester marker most frequently adjusted for race.

First trimester PAPP-A adjustment

Of the 13 centers reporting adjustment of PAPP-A for race, all agreed that PAPP-A levels were higher in Black versus Caucasian women. Nine were applying an adjustment factor to the multiple of the median. This factor ranged from a low of 1.07 to a high of 1.57. Four laboratories reported using a separate set of median values for Black women. Two of the four did not report the difference in the median values for the racial groups, while one said the median was 10% higher for Blacks and the another participant reported that it was 34% higher.

Conclusions: PAPP-A is considered to be higher in Blacks than in Caucasians, but the actual adjustment factor is very broad; ranging from 7% to 57%.

First trimester total hCG adjustment

Of the five centers reporting adjustment of first trimester hCG for race, three agreed that levels were higher in Blacks compared to Caucasian women. One participant used an adjustment factor of 1.12. Two centers reporting using separate median values that were 10% or 21% higher for black women. One center reported using separate median values but no factor was provided and the calculated MoM values for FT-07 were the same regardless of race. The last of the five centers provided no data.

Conclusions: There is little consensus among participants whether hCG needs to be adjusted for race. Among those that do adjust, Blacks are considered to be higher. The effect size is similar to that found for second trimester hCG measurements.

Median maternal weight by race

Nineteen centers reported median (or average) maternal weights for both Black and Caucasian women. Four of these laboratories reported the same median weight, regardless of race. Another four laboratories reported that the two groups had median weights within 5% of each other. The remaining laboratories found larger differences, with Blacks weighing more. The median reported weight in Black women was 169 lbs. (range 140 to 182 lbs.) and in Caucasian women 156 lbs. (range 140 to 163 lbs).

Interaction between race and maternal weight leading to confounding

In an attempt to clarify the racial adjustment factor applied to Black women, laboratories reported the MoM value for sample FT-07 if the women were Black rather than Caucasian. The ratio of the reported MoM values can be used as an estimate of the adjustment factor used by laboratories. However; this 'calculated' adjustment factor is confounded by the maternal weight of 195 lbs. given for the clinical history for specimen FT-07.

As an example of the this confounding, consider the data reported by one ICP participant who applied both a racial adjustment factor for PAPP-A, and at the same time utilized an in-house race-specific maternal weight equation. In their population, the median weights for Caucasian and Black women are 144 lbs. and 162 lbs., respectively. The reported PAPP-A MoM value for FT-07 is 0.86, if the patient were white. However, if the patient were Black, the MoM is reduced to 0.74. This data yields an apparent racial correction factor for of 1.16 (0.86/0.74). However, the laboratory-specific racial adjustment actually used was 1.07. The discrepancy between the computed 1.16 and the actual factor of 1.07 is attributable to the adjustment for the woman's weight of 196 lbs. This weight is more extreme for a Caucasian woman than for a Black woman, yielding different maternal weight factors. The overall adjustment for both women is as follows:

Adjusted MoM = unadjusted MoM / race adjustment factor / wt adjustment factor

Caucasians	0.86	=	0.69	÷	1.00	÷	0.80
Blacks	0.74	=	0.69	÷	1.07	÷	0.86

This analysis is possible because this lab has generated a maternal weight adjustment equation that is race specific. If instead, this lab used the weight adjustment equation derived from the Caucasian population the adjustment factor for the MoM value would be much higher than expected even given the weight of 195 lbs. In effect, if the black woman in FT-07 were of average weight (162 lbs.) she would have her MoM adjusted upward because the average weight of the white population is 144 lbs. The adjustment factor for women of average weight with a race specific weight equation should be about 1.0. **Only three participants use race-specific weight adjustment, and therefore the reported MoM values for white versus Black women for sample FT-07 for most laboratories will be confounded by maternal weight.** This problem may account for the discrepancy between what some labs say their adjustment factor is for PAPP-A, as compared to the factor calculated using the ratio of the reported MoM values for the two groups.

Conclusions: Most participants report that Black women weigh more than Caucasian women, but few have implemented race-specific weight adjustments. Maternal weight can confound the attempt to determine the race adjustment factor for Black women.

First trimester race adjustments: the literature

Four studies are available that report on the effect of race on PAPP-A concentrations. Spencer *et al*.[1,2] found that PAPP-A levels were 57% and 55% times higher in Afro-Caribbean women, respectively, than in Caucasian women. Krantz DA, *et al.*[3] found that Black women had PAPP-A levels 35% higher than in Caucasians. Kagan KO, *et al*.[4] found that blacks of Afro-Caribbean origin had PAPP-A values that were 57% higher than in Caucasian women. Several participants have implemented their racial adjustments based on these publications. However, the underlying populations (mostly Afro-Caribbean) may not reflect the effect in African Americans. In addition, the analyses may be confounded by the lack of race specific maternal weight adjustments. There are few, if any, published data on the racial effect on first trimester total hCG levels.

Conclusions: First trimester PAPP-A is clearly and consistently elevated in the first trimester in Black versus Caucasian women. However, the estimate of the extent of elevation varies widely.

Gap in knowledge: Given the wide range in practice, estimates of effect size, and lack of directly relevant information, a large US-based study should be performed to answer these questions and help standardize screening practice.

References

1. Spencer K, *et al*. The influence of ethnic origin on first trimester biochemical markers of chromosomal abnormalities. *Prenat Diagn,* 2000;20(6):491-4.
2. Spencer K, *et al*. Ethnicity and the need for correction of biochemical and ultrasound markers of chromosomal abnormalities in the first trimester: a study of Oriental, Asian, and Afro-Caribbean populations. *Prenat Diagn,* 2005;25(5):365-9.
3. Krantz DA, *et al.* Maternal weight and ethnic adjustment within a first trimester Down syndrome and trisomy 18 screening program. *Prenat Diagn,* 2005;25(8) 635-40.
4. Kagan KO, *et al.* First trimester screening for trisomy 21 by free beta hCG and PAPP-A: impact of maternal and pregnancy characteristics *Ultrasound Obstet Gynecol,* 2008;31(5):493-502.

Interpretive question: Integrated screening

Twenty-one participants reported integrated risks using first trimester marker results (FT-10) in combination with the second trimester quadruple markers results (CAP FP survey, sample 12). Seven participants (including several manufacturers) do not perform integrated testing and one other did not respond. All of the 21 participants report integrated risks using second trimester quadruple markers; a subset also report integrated risks using a second trimester triple test. Only the integrated quadruple risk is 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. Also, adjusted risks are rounded to two significant figures. Two participants using the free beta subunit are included in the summary statistics (last two entries in Table 2). Two participants had their full integrated results identified as outliers.

The trimmed geometric mean risk for the second trimester quadruple test for ICP participants is 1:250. This is quite close to the consensus risk of 1:260 for FP-12 reported by the CAP FP-B participant summary report. The trimmed mean PAPP-A MoM value for specimen FT-01 is 0.68. The crossover point for PAPP-A is 0.6 MoM. The crossover is defined as the MoM value where the heights of the two distributions are equal. At this point, the likelihood ratio (LR) is expected to be 1.0 indicating that the serum integrated risk should be similar to the quadruple risk. The geometric mean for the serum integrated risk for ICP participants are both 1:250, consistent with expectation. The trimmed mean NT MoM value calculated by ICP participants for FT-10 is 1.04. The crossover point for the affected and unaffected distributions of NT MoM is approximately 1.45 MoM. The expectation is, therefore, that the risk for the full integrated test would be significantly reduced and the lowered geometric mean risk of 1:790 versus 1:250 for the serum integrated test meets expectation (LR of 0.32). The variability of the distribution of integrated risks is higher than that found for either second trimester or first trimester Down syndrome risks. The reason for this is not clear, and will be investigated in the next distribution. An important question is whether the variability is mainly attributed to assay differences (*e.g.,* MoM results), or to the implementation of the algorithm (computation of risks).

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	110	110	85	90	310	310
2	480	480	365	370	1500	1500
2	1300	1 300	1400	1400	8600	8600
2	220	220	120	120	570	570
2	209	210	120	120	440	440
2	136	140	121	120	434	430
2	75	80	71	70	240	240
2	380	380	410	410	2900	2900
2	140	140	110	110	530	530
2	222	220	551	550	1820	1820
2	240	240	260	260	1200	1200
2	120	120	85	90	290	290
2	100	100	75	80	380	380
2	310	310	150	150	540	540
2	148	150	85	90	180	180
3	190	100	290	210	1200	900
2	110	110	80	80	220	220
3	1500	1100	3700	2700	31000	23000
3	480	400	700	500	4200	3100
3	1400	1000[1]	2800[1]	2100[1]	20000[1]	15000[1]
3	840	600[1]	1400[1]	1000[1]	6400[1]	4700[1]
Trimmed		**None**		**None**		**15000,23000**
Geo Mean risk		**250**		**250**		**790**
CV(log risk)		**16%**		**21%**		**17%**
Mean-2SD		**40**		**26**		**90**
Low		80		70		180
High		1300		2700		8600
Mean+2SD		**1300**		**2000**		**7000**

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

George J. Knight, Ph.D. Glenn E. Palomaki, Ph.D. (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