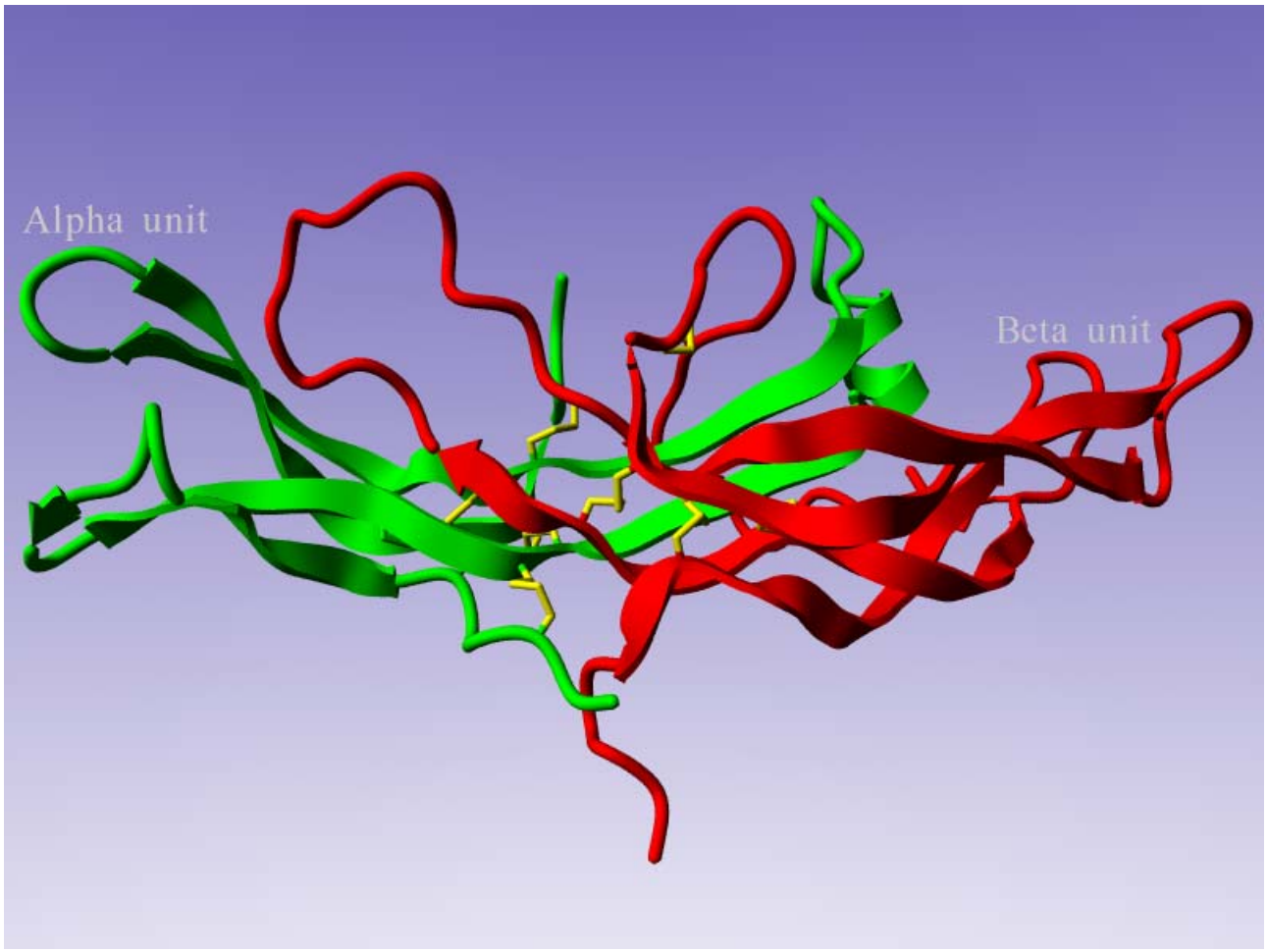
First Trimester
Interlaboratory Comparison Program

Distribution 2006 FT-C



hCG: the alpha chain in green and the beta chain in red. Disulfide rings are colored as yellow sticks. (http://gris.ulb.ac.be/intro.html)

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

DISCUSSION

Data Listing and Analysis

Reading the Data Listing: The five page data listing contains a summary of reported results for all participants; one page summarizing each of the five specimens. Your lab ID is listed at the beginning of the row with your results. Missing data (blanks) are likely due to manufacturers who do not screen or laboratories that are not yet offering clinical services. Missing data may also result because some laboratories do not measure 'total or intact hCG' but some other marker. 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, and after a logarithmic transformation of risk. In the future, we will implement more sophisticated trimming algorithms.

Manufactured Specimens: Samples **FT-11 and FT-13** are artificial samples composed of normal human serum spiked with recombinant hCG, recombinant DIA and a purified native PAPP-A concentrate derived from patient sera. Sample **FT-15** is a 1:1 dilution of **FT-11** with normal human serum (NHS). The PAPP-A results are well below the expected values and quite variable. It appears that the PAPP-A material used for the spike is degrading. However, the PAPP-A values for **FT-15** are overall approximately 50% of that for **FT-11**, as would be expected based on the 1:1 dilution. A comparison of the results for **FT-11** and **FT-15** is shown in Table 1. All three meet the expectation that sample **FT-15** is about half of sample **FT-11**.

Table 1. Comparison of results from FT-11 and FT-15 (expected FT-15 / FT-11 is 0.5)

PAPP-A			hCG			DIA		
FT-11	**FT-15**	**15/11**	**FT-11**	**FT-15**	**15/11**	**FT-11**	**FT-15**	**15/11**
1.6	0.8	0.50	201.0	151.4	0.75	392.6	246.3	0.63
0.3	0.1	0.33	233.3	101.1	0.43	416.0	190.0	0.46
0.3	0.1	0.33	208.0	101.0	0.49	393.3	203.5	0.52
0.3	0.2	0.67	194.0	98.3	0.51	189.1	148.6	0.79
0.8	0.5	0.63	162.1	97.5	0.60	200.1	138.2	0.69
0.4	0.2	0.50	192.6	95.8	0.50			
1.0	0.4	0.40	214.4	103.3	0.48		**Outliers**	none
0.9	0.4	0.44	168.8	86.3	0.51		**Number**	5
0.8	0.4	0.50	254.4	124.8	0.49		**Mean**	0.62
0.7	0.3	0.43	240.0	123.8	0.52		**SD**	0.13
0.9	0.4	0.44	228.2	109.6	0.48			
0.9	0.4	0.44	318.9	118.0	0.37			
0.9	0.4	0.44	215.8	107.7	0.50			
0.3	0.2	0.67	194.1	95.7	0.49			
0.2	0.1	0.50	200.4	96.7	0.48			
0.2	0.1	0.50						
0.2	0.1	0.50		**Outliers**	0.75			
0.2	0.1	0.50		**Number**	14			
				Mean	0.49			
	Outliers	none		**SD**	0.05			
	Number	18						
	Mean	0.49						
	SD	0.09						

Long Term Control (12 week pool): **Specimen FT-12** in this distribution was the same 12 week pool distributed as **FT-03** in the FT-A distribution, and **FT-07** in the FT-B distribution. The results on this sample allow an evaluation of kit consistency over time and the stability of the pool. We have attached a summary of these results for your review (Table 2). The agreement in the all lab trimmed mean (ALTM) was, in general, good for all three analytes, as was the within-kit consistency in the kit values over time. In particular, the ALTM for PAPP-A was excellent, and the between-kit agreement in absolute values was also very good. The table also contains the summary means (standard deviations and coefficients of variation) for kits where more than two users reported results. Overall, these results indicate that patient pools are a reliable and stable source for proficiency testing samples.

Sample **FT-14** was the same 12 week pool as **FT-12** but was spiked with 1.7 mIU/ml of PAPP-A to evaluate recovery of PAPP-A spiked into patient material. The 12 week pool has all of the compounds found in pregnancy serum, and therefore the matrix should be identical to that found in patients. Over- or under-recovery of the amount of the PAPP-A spike would suggest that the PAPP-A material used is problematic. The 1.7 mIU/ml spike should have should have yielded an approximate value of 4.4 mIU/ml (2.7 + 1.7). However, the ALTM for **FT-14** of 2.90 is only 0.23 mIU/ml higher than the average 2.67 mIU/ml calculated for **FT-12**. Also, results for any given lab are only slightly higher, or even unchanged, for **FT-14** as compared to **FT-12**. This failure to recover the added PAPP-A may have occurred because the PAPP-A used for the spike was unstable.

Variance of Mass Unit Results: The PAPP-A analyte values for the two specimens that consist exclusively or primarily of patient pools (**FT-12, FT-14**) yield a CV of 12%. Some of this variability appears to result from systematic differences between the DSL, DPC, and PE Delfia methods, but the between-kit differences in values are small. In contrast, the CVs for the artificial samples (**FT-11, FT-13, FT-15**) for PAPP-A are much larger (57%, 40%, and 56%, respectively). This indicates that the PAPP-A spike does not accurately reflect what is obtained with real patient specimens, where the between-kit differences are small. In contrast, the CVs for hCG in all five specimens are similar (10% to 13%), suggesting that recombinant hCG reflects the specificity observed for patient samples.

Variance of MoM Results: The CVs for PAPP-A MoM values for the manufactured specimens (**FT-11, FT-13, FT-15**) of 57%, 41%, and 63%, are higher than for patient pools (**FT-12, FT-14**) at 13% and 23%, respectively. This reflects the relatively high CVs for the PAPP-A values. The CVs for the hCG MoM values in patient pools versus artificial samples are similar and relatively tight (range 10% to 15%).

Table 2. Comparison of 12 week patient pool distributed three times in 2006 (Distributions A, B and C)

	PAPP-A				hCG					DIA				
	Kit	FT-03	FT-07	FT-12	CV	Kit	FT-03	FT-07	FT-12	CV	Kit	FT-03	FT-07	FT-12	CV
	DPC	2.7	2.9	3.0	**5.3%**	Abbott		97.4	106.3	**6.2%**	DSL	296	252	310	**10.6%**
	DPC	2.6	2.5	2.7	**3.8%**	Bayer	94.2	86.6	93.0	**4.5%**	DSL	307	297	324	**4.4%**
	DPC	2.7	3.1	3.1	**7.8%**	Bayer	84.6	85.6	80.8	**3.0%**	DSL	265	212.8	300	**16.9%**
	DPC		2.6	3.0	**10.1%**	Beck	96.3	107.1	105.3	**5.6%**	DSL		278	292	**3.5%**
	DSL	2.8	3.2	3.2	**7.5%**	Beck	98.7	81.2	79.9	**12.1%**	DSL		291		
	DSL	2.9	3.0	2.6	**7.3%**	Beck	92.2	93.1	97.4	**2.9%**	DSL	271	291	271	**4.1%**
	DSL	1.6	1.8	1.0	**28.4%**	Beck	82.7	100.1	102.2	**11.3%**					
	DSL	2.5	2.7	2.7	**4.4%**	Beck		89.2	85.1	**3.3%**					
	DSL	2.6	2.8	2.3	**9.8%**	DPC	65.7	82.9	81.0	**12.3%**					
	DSL		3.2	2.9	**7.0%**	DPC	83.2	85.3	75.1	**6.6%**					
	DSL	4.0	2.3	2.1	**37.3%**	DPC	88.7	86.5	85.7	**1.8%**					
	DSL	2.3	2.7	2.3	**9.5%**	DPC	70.8	67.1	71.3	**3.3%**					
	DSL		3.1	2.7	**9.8%**	DPC	74.4	73.5	82.1	**6.2%**					
	DSL		2.4	2.3	**3.0%**	DSL	93.8	69	88.4	**15.6%**					
	DSL	2.5	2.8	2.2	**12.0%**	PE		81.7	69.5	**11.4%**					
	PE	2.7	3.0	2.6	**7.5%**	PE	78.3	103.6	77.1	**17.3%**					
	PE		2.7	2.9	**5.1%**										
	PE	2.8	2.6	2.7	**3.7%**										
	PE	2.8	2.8	2.7	**2.1%**										
ALTM		**2.66**	**2.80**	**2.67**	**7.5%**		**84.9**	**86.9**	**88.1**	**6.2%**		**285**	**270**	**299**	**4.4%**
Number		**14**	**19**	**19**			**13**	**16**	**16**			**4**	**6**	**5**	
Median	DPC	**2.67**	**2.75**	**3.00**	**6.8%**	Beck	**92.5**	**94.1**	**94.0**	**7.1%**					
Median	DSL	**2.55**	**2.80**	**2.45**	**10.8%**	DPC	**76.6**	**79.1**	**79.0**	**6.0%**					
Median	PE	**2.80**	**2.75**	**2.70**	**4.6%**										

Dimeric inhibin-A: DIA measurements were reported by five participants (Table 3). All reported using the same method (Di-01 or DSL). The following table provides the reported DIA values and MoM levels for all five samples. Included also is the likelihood ratio for DIA in the context of the other two markers. In general, the laboratories report similar DIA MoM levels and changes in Down syndrome risk.

Table 3. Dimeric inhibin-A measurements for FT-C, 2006

Sample Number	Laboratory	Value	MoM	DS Risk (1:n)	DIA LR[1]
FT-11	A	392.6			
	B	416.0	1.69	10	1.00
	C	393.3	1.74	5	1.00
	D	189.1	1.08	5	1.25
	E	246.3			
FT-12	A	309.8			
	B	300.0	0.95	120	1.22
	C	323.6	1.32	17	1.06
	D	292.1	1.35	55	1.08
	E	271.3			
FT-13	A	261.7			
	B	259.0	1.12	150	2.24
	C	255.5	1.06	647	2.34
	D	217.8	1.11	385	2.33
	E	193.7			
FT-14	A	290.0			
	B	293.0	0.97	820	1.78
	C	276.8	1.26	115	1.44
	D	230.0	1.33	1110	1.42
	E	232.3			
FT-15	A	200.1			
	B	190.0	0.66	10	1.00
	C	203.5	0.92	20	2.22
	D	148.6	0.86	4	1.00
	E	138.2			

[1] For each participant, the increase/decrease in risk from the combination of NT, PAPP-A and hCG, divided by the risk that includes DIA measurements.

Interpretative Question: Using sonographer-specific medians

1. **Does your laboratory provide clinical results for Down syndrome screening that include NT measurements**? Of the 20 participating laboratories, four responded that they do not provide clinical results. Of the 16 clinical laboratories, 14 answered one or more of the interpretative questions. The following analyses focus on those 14 participants.

2. **Sonographer FST**
   a. **What is the median NT at a CRL of 60 mm (about 12w+4d) for this sonographer?** The average sonographer-specific NT computed by the 14 laboratories was 0.99 mm (SD of 0.045 mm, CV of 4.6%, range 0.95 mm to 1.04 mm). Given that all laboratories were provided with the same 150 observations to derive medians for this sonographer (and the consensus equation provided in the FT-B distribution), the high consistency of results is expected.
   b. **What is the NT MoM for sonographer FST?** The average MoM reported was 2.59 (SD of 0.10, CV of 3.9%, range 2.27 to 2.60 MoM, with 2 outliers at 1.00 and 3.55 MoM). It is not evident how the two laboratories with outlying NT MoM values obtained their results. To compute the NT MoM, the reported NT value of 2.45 mm would be divided by the reported median value. Eight of the laboratories did this correctly. Four other laboratories differed from this expected MoM by 3% to 10%; a difference that cannot be explained by rounding. The two outliers differed by 35% and 61% from the expected MoM (i.e., 2.45 mm/computed median). These six laboratories should re-examine their computations and determine the source of the discrepancies. The target NT MoM for sonographer FST was selected to match the 2.23 MoM provided in the clinical history for FT-12, which is close to the average reported NT MoM of 2.59.
   c. **What is the Down syndrome risk for FT-12?** The consensus estimate was 1:25 (95% prediction limits 1:6 to 1:110, log CV of 24%). The wider range of risks is expected, as each laboratory's biochemistry results are included in this computation. Two laboratories did not report Down syndrome risk and were, therefore, not included in this analysis.

3. **Sonographer GNG**
   a. **What is the median NT at a CRL of 60 mm (about 12w+4d) for this sonographer?** To answer this question, each laboratory needed to compute a new set of sonographer-specific median levels using the provided 150 observations. The average sonographer-specific NT computed by all 14 laboratories was 1.61 mm (SD of 0.040 mm, CV of 2.5%, range 1.54 mm to 1.75 mm). Those providing answers that differ from the consensus by more than a few hundredths (e.g., below 1.58 mm or above 1.64 mm) should check their computations. Three reported values below 1.58 mm and three above 1.64 mm. On average, however, this demonstrates that 14 of 16 clinical laboratories using NT measurements can now generate a reliable set of sonographer-specific NT medians given 150 observations.
   b. **What is the NT MoM for sonographer GNG?** Given that sonographer GNG reports systematically higher NT values than sonographer FST, it would be expected that the NT value of 2.45 mm would result in a much lower NT MoM. The average NT MoM reported was 1.55 (SD of 0.028, CV of 2.5%, range 1.48 to 1.58 MoM). The same two laboratories that reported outlying values for sonographer FST also reported outlying values for sonographer GNG (1.00 and 2.07 MoM). Nine laboratories reported a MoM that was equal to 2.45 mm divided by their calculated NT median. Three differed by 3% to 13% from the expected NT MoM and the two outlying results differed by about 35%.

c. **What is the Down syndrome risk for FT-12?** The consensus estimate was 1:1000 (95% prediction limits 1:610 to 1:1800), log CV of 3.9%). One outlying risk (1:91) was reported by one of the laboratories with an outlying NT MoM level. Among the 11 laboratories who reported both Down syndrome risks (and did not have any outlying values), the distribution of change in risks (risk for sonographer GNG/risk for sonographer FMF) was bimodal. One group of seven laboratories found the risk to be 24-fold lower (SD 4.7, CV 20%). The other group of four laboratories found the risk to be 98-fold lower (SD 8.4, CV 9%). As of now, it is not clear what has caused this difference, but we will attempt to identify the reason in future exercises.

Discussion

The set of NT values for sonographer FST and GNG were generated to yield gestational age-specific median values similar to the consensus median values reported in two large published studies, calculated using actual data from experienced well trained sonographers. As such, the calculated median values for these two sonographers correspond to what might be encountered by laboratories in interpreting NT measurements from different centers or different sonographers. The ALTM median value for sonographer FST for a CRL of 60 mm (12w +4d) is 0.99 mm, and using the provided NT measurement of 2.54 mm yields a consensus MoM of 2.59. In contrast, the ALTM median value for sonographer GNG for this same gestational age is 1.61 mm. Sonographer GNG is obtaining much higher NT measurements, on average, than sonographer FST. Consequently, the resulting NT MoM for sonographer GNG is much lower than for sonographer FST (i.e. 1.55). These systematic differences in the measurement of NT account for the much lower risk calculated for sonographer GNG versus sonographer FST — 1:25 versus 1:1000. This exercise, which has been designed to reflect differences that can be observed in actual practice, demonstrates the pitfalls in using a single NT median equation to interpret all NT measurements. A strong argument can be made that center and sonographer-specific median values should be developed and routinely monitored by the laboratory to ensure that NT measurements achieve their potential as a Down syndrome screening marker.

Important Note

For planning purposes, participants can expect that in 2007, it will be considered routine for the laboratory to be able to 1) use existing sonographer-specific medians for sonographers FST and GNG, 2) compute new sets of sonographer-specific medians based on smaller, and 'dirtier' datasets, and 3) compute the NT MoM based on sonographer-specific medians values given an NT value (in mm) with a gestational age provided as a CRL in mm, or a gestational age in weeks or days rather than being given an NT MoM in the case history. For those laboratories that have not yet been able to successfully perform such calculations, we encourage you to gain this ability prior to the next distribution in 2007. You should find an attached spreadsheet that may be helpful. Please contact the ICP coordinators if you have questions.

George J. Knight, Ph.D.	Glenn E. Palomaki, B.S.	(207) 657-7888
Jacob A. Canick, Ph.D.	Geralyn M. Messerlian, Ph.D.	(401) 453-7654

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