MCH/CCS-04-

IMPROVED PRENATAL DOWN SYNDROME SCREENING: PAIRED TESTING

George J. Knight, Ph.D.
Prenatal Screening Laboratory
Foundation for Blood Research
Scarborough, ME 04074

Published: June, 2004

Final Report
Grant 5R40 MC 00195
Project Period: 02/01/01-01/31/04

Copies of this report may be obtained for a fee from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22161, Telephone 703-487-4650

Prepared for:

THE MATERNAL AND CHILD HEALTH RESEARCH PROGRAM
MATERNAL AND CHILD HEALTH BUREAU, HRSA, PHS, DHHS
PARKLAWN BUILDING
5600 FISHERS LANE
ROCKVILLE, MARYLAND 20857

Executive Summary of the Comprehensive Final Report

R40 MC 00195

IMPROVED PRENATAL DOWN SYNDROME SCREENING: PAIRED TESTING

Statement of the Problem: Down syndrome is the most common of the major chromosomal disorders that are compatible with life, having a prevalence of about 1:700 births in the general population. The condition creates ongoing medical and societal challenges for families with affected members. Down syndrome can be reliably diagnosed in pregnancy by chromosomal analysis of fetal cells obtained by amniocentesis, but these procedures are expensive and carry a risk of procedure-related fetal loss. For this reason, maternal serum screening tests have been developed to identify women whose pregnancy is at a high enough risk of Down syndrome to justify a diagnostic procedure. Currently, the most widely used screening test relies on second trimester measurement of three substances in maternal serum. This 'triple' test can detect 70 to 75 percent of Down syndrome cases by identifying 7 to 8 percent of the pregnancy population at high risk (screen positives). More than 2 million pregnant women are screened annually in the U.S. for Down syndrome. However, most screen positive women will not have a baby with Down syndrome. These false positive results can cause psychological distress, add expense to the health care system, and subject pregnancies to the risk of procedure-related loss. Improved screening tests that reduce the false positive rate would be an important contribution for patients, providers and the health care system in general.

This Maternal and Child Health Bureau funded study was a non-randomized intervention trial to evaluate the acceptability and performance of a new method of screening for Down syndrome. The 'integrated serum' test combines the biochemical measurements from both a first and a second trimester maternal serum sample with maternal age to calculate a single Down syndrome risk. The result is available for interpretation in the second trimester. The integrated serum test has the potential to reduce the false positive rate to below that of the triple test, while maintaining a high detection rate. The study population studied is drawn from over 11,000 pregnant Maine women.

Research Objectives: The overall goal of the current study is to establish that the integrated serum test can be successfully implemented in a variety of primary care settings through a centrally administered program with a resulting reduction in the false positive rate. The success of the integrated serum screening approach is assessed by determining the:

- proportion of women enrolled by providing a first trimester serum sample
- proportion of enrolled women providing a second trimester serum sample
- proportion of enrolled women receiving an integrated serum test result
- proportion of enrolled women with positive integrated serum test results

- proportion of women with positive screening results if the triple test had been used instead
- proportion of Down syndrome cases detected
- costs and benefits associated with the integrated serum test

Findings: During the 24 month enrollment phase, 11,159 women provided a first trimester sample (representing 61% of all pregnant women being screened by our institution). Of the enrolled women, 9,723 (87%) provided a second trimester blood specimen required for completing the integrated serum test. For 8,773 of these women, matching first and second trimester specimens were identified within the specified gestational age range and an integrated serum interpretation provided. The number of women enrolled and completing the process was higher than estimated in our proposal.

Among the 11,159 enrollees, 1,436 women (12%) did not complete the integrated serum test because their second trimester sample was not received. The most common reasons for this were miscarriage (40%), declined testing (31%), and opting for amniocentesis instead of further screening (17%). Among this latter group, the vast majority (87%) were 35 years of age or older, suggesting that these ultimately women wanted the reassurance associated with diagnostic testing. Among the 11,159 enrollees, another 950 women (9%) did not receive an integrated serum test report, because the first trimester sample was collected outside the acceptable gestational age of 8 to 13 weeks. Most of these (92%) were drawn too early. These 950 women were screened using the 'quadruple' test, the best maternal serum test possible using only a second trimester sample.

Among the 8,773 women screened using the integrated serum test, the false positive rate was 3.2 percent compared to 4.5 percent if the triple test were to have been used instead (a reduction of 29%). However, if the analysis is restricted to pregnancies dated by ultrasound, the rates are 2.7 percent compared with 4.5 percent, respectively (a reduction of 40 percent. This latter analysis confirms our prediction (that was based on ultrasound dated pregnancies) of integrated serum testing reducing the false positive rate by up to half. For every 10,000 ultrasound-dated women screened using the integrated serum test, 180 fewer women (270 vs 450) would be referred for diagnostic procedures(s). For this analysis, the screening cut-off level for the triple test was set to provide an estimated detection rate of 70 percent, equivalent to that expected for the integrated serum test. Preliminary data from our pregnancy follow-up indicated that the detection rate for the serum integrated test is 64 percent.

Two patient satisfaction surveys of 30 women each were administered at 6 and 18 months into the enrollment phase. The key findings among these 60 women include:

- all remembered having the integrated serum test
- almost all (98%) remembered having a prenatal test in their previous pregnancy
- three quarters indicated that they did not experience anxiety because they had to wait for the interpretation until the second trimester
- almost all (95%) would consider integrated serum testing in a future pregnancy.

A cost analysis compared the additional laboratory costs of offering integrated serum testing with savings resulting from the reduction in the number of diagnostic procedures (ultrasound examinations and amniocenteses). The additional costs of offering integrated serum screening compared to the triple test in a cohort of 10,000 women is $470,000. This is slightly more than the $234,000 savings related to reduced diagnostic costs. Additional considerations include the reduction in anxiety for 130 fewer screen positive women (270 versus 450) and half the number of procedure-related losses (1 versus 2) for integrated serum testing versus the triple test. In this analysis, the Down syndrome detection rate was again held constant.

Recommendations:

- The current study provides evidence that one form of integrated screening (the integrated serum test that is based solely on maternal serum markers) can be successfully introduced into routine practice in a distributed health care. Most women reported that they did not experience excess anxiety because they had to wait a month or more to get their test results. However, the process of matching samples to women and the need for close monitoring of the PAPP-A assay might make translation to other laboratories difficult. These issues need to be addressed prior to introducing the integrated serum testing into routine prenatal care.
- Women choosing the integrated serum test will be less likely to have a false positive screening result than the triple test at essentially the same Down syndrome detection rate. The expected reduction of 40 percent only occurs, however, when the pregnancies are dated by ultrasound. Integrated serum testing should not be performed if the pregnancy is dated by last menstrual period.
- A total of 9 percent of enrolled women did not receive an integrated serum test, usually because the first trimester sample was drawn too early. Routine first trimester ultrasound dating would allow these women to also receive the benefits of integrated serum testing.
- Another 12 percent of enrolled women did not receive an integrated serum test because a second trimester sample was not received. For at least some of these women, existing fetal demise could be identified by having a routine first trimester ultrasound, and these women would not be candidates for integrated serum testing.
- There is some evidence (difficulty of measuring low levels of PAPP-A and the wide variability in the distribution of values) that offering integrated screening prior to 10 weeks' gestation may be more difficult. Recommending screening at 10 weeks or later could increase the efficiency of the integrated serum test by reducing the chance that women would be inadvertently be enrolled too early for reliable screening.
- The recommendations above indicate the need for ultrasound-based dating to optimize the integrated serum test, preferably in the first trimester. The integrated serum test could be further improved, and the false positive rate further reduced, if the first trimester ultrasound measurements included a nuchal translucency (NT) measurement. First trimester Down syndrome screening using serum markers and NT measurements (combined test) is now being offered at selected high risk perinatal centers in the United States. However, NT measurements can only be performed by certified sonographers who participate in ongoing quality control monitoring. It has been suggested that obstetricians in primary practice could reliably

obtain this measurement with proper training and ongoing quality assurance measures. The Maternal and Child Health Bureau might consider funding a prospective trial of fully integrated testing for Down syndrome in routine practice – something that has not yet been tried in the United States.

List Of Products

Knight G.J. Results from the integrated serum test study: A U.S. Screening Project. presented at *Prenatal Screening for Down Syndrome: Introducing the Integrated Test into Medical Practice.* Brown University, Rhode Island. March 28-29. 2003

Knight G.J. Down syndrome Screening: What's New. Maine Medical Center. *Obstetrics and Gynecology Grand Rounds*. November 14, 2002.

Knight GJ. Integrated serum screening in Maine. *Down's Screening News*. February 2002. Leeds University, UK. Editor P.bloom@leeds.ac.uk.

Knight GJ, Palomaki GE, Haddow JE. Integrated serum screening for Down syndrome: An intervention trial involving 11,159 women. Manuscript in preparation

I. INTRODUCTION

A. Nature of the Research Problem

Fetal Down syndrome: Down syndrome is an important medical condition that creates ongoing medical and societal challenges for families with affected members. It is the most common of the major chromosome disorders that are compatible with life, having a prevalence of about 1 in 700 live births. Karyotyping the cells of affected individuals is a highly reliable diagnostic method, including cells obtained during pregnancy from chorion villi (in the first trimester) or amniotic fluid (in the early second trimester). Identifying Down syndrome prenatally has proven helpful to many families in decision-making about pregnancy management. However, diagnostic procedures for obtaining fetal cells carry some risk for pregnancy complications (e.g. fetal loss in 1 in 200 procedures), and it is not practical or cost effective to perform a diagnostic procedure on all pregnant women.

Down syndrome screening – From Maternal Age Alone to the "Triple Test": Various methods have been developed to identify women at sufficient risk for carrying a baby with Down syndrome to warrant offering an invasive test. The first of these (asking a woman her age), was introduced in the 1970s and continues to be used today. It takes advantage of the well-documented rise in risk for Down syndrome in women with increasing maternal age. Women age 35 or older are considered to be at sufficiently high risk for offering a diagnostic procedure to karyotype the fetus. Currently in the United States about 14 percent of all pregnant women are 35 years or older, and 50 percent of Down syndrome cases occur in this group. However, maternal age is a poor screening test because it requires a high percentage of all pregnant women (14%) to undergo a diagnostic procedure (amniocentesis) to detect half (50%) of Down syndrome pregnancies. Approximately 180 amniocenteses / karyotypes are performed to identify each pregnancy affected with Down syndrome. A significant advance occurred when it was discovered that substances in maternal serum are altered when the mother is carrying a baby with Down syndrome. The first of these was alpha-fetoprotein (AFP). In 1984, the discovery that lower second trimester maternal serum AFP levels are associated with a Down syndrome pregnancy made it possible, for the first time, to offer screening to pregnant women younger than age 35. This test was similar in performance to maternal age, with a detection rate 20 to 25 percent, with a 5 percent false positive rate. However, AFP and maternal age are independent markers and could be combined to improve overall screening performance. Additional maternal serum markers were subsequently discovered in the late 1980s. Currently, the most widely used screening test is the triple test [1], which combines measurements of three substances in maternal serum in the second trimester of pregnancy (AFP, unconjugated estriol - uE3, and human chorionic gonadotrophin - hCG). The triple test can detect up to 70 to 75 percent of Down syndrome cases by identifying 7 to 8 percent of the pregnancy population as screen positive, a significant improvement over previous screening tests. Currently, an estimated 2.5 million women are screened annually in the United States [2].

False Positives — The Down Side of Prenatal Screening: Although the triple test is a significant improvement over maternal age screening (and over the combination of maternal age and AFP measurements alone), it still is true that almost all women with positive screening results are false positives. Approximately 1 in 50 woman with a positive result will have a baby with Down syndrome (positive predictive value of 1 in 50). This means that almost all women referred for amniocentesis will not have a baby with Down syndrome. Because only about 1 in 50 positive results will be a true positive, the false positive rate is essentially equivalent to the screen positive rate. False positive screening results cause psychological distress, cost to the health care system for expensive diagnostic procedures, and the potential loss of unaffected fetuses attributable to second trimester amniocentesis [3-10]. In addition, false positive rates that are considered 'high' may lead both health care providers and patients to avoid screening because of the belief that it is too non-specific. To address this issue, some laboratories in the United States have added dimeric inhibin A (DIA) to the triple test as a fourth marker: the "quadruple" test. Quadruple (or quad) testing can increase the detection rate to as high as 80 percent while slightly reducing the screen positive rate [11]. However, it still requires approximately 40 amniocenteses to detect each case of Down syndrome. The problem of many false positives to detect each case of Down syndrome remains. The relatively small increase in detection with quad marker screening as compared to the triple test exemplifies the phenomenon that new markers provide only marginal gains in increasing the detection rate. However, that same new marker can decrease the false positive rate by one-third to one-half, if the detection rate were to be held constant (this can be accomplished by modifying the risk cut-off level defining a positive test result).

It can reasonably be argued that the time has come to shift the focus of prenatal screening for Down syndrome from obtaining marginal gains in detection to reducing the burden of false positive screening results.

The Integrated Serum Test - Current Study Aims: The current MCHB-funded study is a non-randomized intervention trial to evaluate a new method of screening for Down syndrome, the integrated serum test.[12] This new test is designed to significantly reduce the false positive rate (compared to the current standard of care, the triple test) while maintaining the detection rate. The integrated serum test combines the best first and second trimester maternal serum biochemical measurements to assign each pregnancy a Down syndrome risk for interpretation in the second trimester. The integrated serum test is projected to detect 70 to 75 percent of Down syndrome cases (same as the triple test), but with fewer false positives. Reducing the number of women with false positive test results while maintaining a high detection rate would be an important contribution for both patients and the health care system.

B. Purpose, scope and methods of the investigation

The Integrated Serum Test: The purpose of this investigation is to develop and validate a new approach to serum based prenatal screening for fetal Down syndrome — the integrated serum test — that will reduce the false positive rate

(compared to the current standard of care) while maintaining the same detection rate. An intervention trial is the vehicle used for this assessment. The integrated serum test combines both first and second trimester maternal serum biochemical measurements to assign each pregnancy a Down syndrome risk in the second trimester. Those risks are then used to identify women at sufficient risk to warrant offering second trimester amniocentesis and fetal karyotyping. Compared to the current standard of care (the triple test), the integrated serum test will maintain a high Down syndrome detection rate but will reduce the false positive rate (and by extension, the amniocentesis rate) by one-half. Thus, fewer women will experience anxiety and require diagnostic procedures, but the same number of cases of Down syndrome will be detected. The medical and financial costs of diagnostic testing will also be reduced proportionally. Use of easily collected serum will allow the integrated serum test to be routinely available to the general pregnancy population, regardless of geographic location. The study population will be women in Maine receiving first trimester prenatal care. After informed consent, a blood sample will be drawn in the first trimester (8 to 13 weeks' gestation) and sent to the laboratory for measurement of one component of the integrated serum test. A subsequent blood sample will be drawn in the second trimester (15 to 20 weeks' gestation) for measurement of the four additional components of the integrated serum test. After the computer matches the two sets of results for a given woman, the analytic results will then be analyzed and a single risk for Down syndrome reported. Women with high risks (screen positives) will be managed according to current medical practice.

C. Nature of the Findings (a brief general reference)

The integrated serum test was successfully introduced as part of routine prenatal care in both rural and urban setting in the state of Maine. Almost 8800 women were screened during a two-year period, representing an estimated 60% of those eligible for the test. Women who elected integrated serum testing whose pregnancy was dated by ultrasound were 40 percent less likely to have a false positive screening test as compared to the triple test. Patient surveys found that the test was well accepted by pregnant women, and did not cause additional anxiety because of the need to wait for the final results until the second trimester. The cost to the health care system will be slightly higher than second trimester triple testing, but the reduction in maternal anxiety, and potential loss of healthy normal fetuses are additional benefits. If a screening program chooses to implement the integrated serum test it is recommended that women only be offered the test if they have had an ultrasound confirmation of gestational age before the second trimester interpretation.

II REVIEW OF THE LITERATURE:

To put the integrated serum test in context, the following section reviews the current and evolving practice of prenatal screening for Down syndrome.

<u>Current Down Syndrome Screening Practice - The Second Trimester Triple Test.</u> Initially, AFP measurements were used to identify women at increased risk for fetal open neural tube (open NTD) defects[13]. Population based prenatal screening for

open NTD began in the United States in the late 1970s and remains essentially unchanged [14]. The current standard of care for prenatal screening in the United States is the triple test performed in the second trimester of pregnancy [1]. Currently, two screening protocols operate simultaneously using the same serum sample. In addition to screening for open NTD, nearly all laboratories also measure additional maternal serum analytes (e.g., uE3, and hCG) to screen for Down syndrome. This second protocol utilizes measurements of all three maternal serum analytes in combination with maternal age to calculate a patient-specific risk for fetal Down syndrome[1]. Typically, women with a risk cut-off greater than a 35-year-old woman (approximately 1:270 in the second trimester) are identified as being screen positive and offered diagnostic testing. The triple test can optimally detect 70 to 75 percent of Down syndrome pregnancies with 7 to 8 percent of all women having a screen positive test result [15,16]. About a third of laboratories in the United States (including our laboratory) use a second trimester screening cut-off of 1:190. This reduces the percentage of women with a positive screening test to approximately 5 to 6 percent, but is associated with a drop in the detection rate to 65 to 70 percent.

Evolving Practice of Prenatal Screening for Down Syndrome.

- *Adding Dimeric Inhibin-A (DIA) Measurements in the Second Trimester – The Quadruple Test:* In the United States, many screening programs for Down syndrome have added DIA to their existing triple test to create the four marker test quadruple (or quad) test [17,18]. Compared to the triple test, the quad test has higher detection by 8 to 10 percent, with about a 1 percentage point drop in the false positive rate[19]. Presently, the quad test is the best second trimester Down syndrome screening method available.
- *First trimester Combined Testing:* Attention is now being focused on moving Down syndrome screening from the second to the first trimester because of patient privacy issues and, if necessary, safer pregnancy termination. In the first trimester, two maternal serum markers have been identified. Pregnancy associated plasma protein-A (PAPP-A) and the free-β subunit of hCG (free β) [20]. First trimester ultrasound measurements of nuchal translucency (NT) are also useful.[21,22] These three results can be combined to produce a first trimester risk for Down syndrome. The results of this "combined" test [23] is used by the physician and patient to decide whether to have diagnostic testing (chorion villus sampling or amniocentesis). Two prospective intervention studies [24,25] have shown that about 80 to 85 percent of Down syndrome cases can be detected with about 5 percent of women having a positive test result (high risk). This screening performance is somewhat better than the second trimester quad test. However, it has several limitations. First, NT measurements can only be performed by specially trained sonographers who participate in ongoing proficiency testing. It is not a test that can be performed by primary care providers. Obtaining an NT measurement is also expensive compared to serum testing. Some third party payers will not pay for the testing except in high-risk patients. Finally, the first trimester combined test cannot detect neural tube defects. A second trimester serum AFP measurement will still be required [26]. As presently performed, combined testing is likely to remain a niche test.

- *The Integrated Test - Combining All First and Second Trimester Down Syndrome Markers:* A novel screening approach has been proposed that acknowledges the potential of both first and second trimester screening markers. Rather than choosing either first trimester screening or second trimester screening for Down syndrome, the 'integrated test' chooses the best of both [12]. The first trimester serum marker PAPP-A and the ultrasound marker NT are measured, but not acted upon until the results of the second trimester AFP, uE3, hCG and DIA measurements are also available. First trimester hCG (or free beta-subunit) measurements are omitted because this analyte is already measured in the second trimester, where its discriminatory ability is greater. Once both the serum and ultrasound measurements are available, a single risk estimate is calculated for interpretation in the second trimester. It has been projected that 94 percent of Down syndrome cases can be detected with 5 percent of women having a screen positive test result. This would make the integrated test the best Down syndrome screening test available. However, the integrated test is still dependent on reliable NT measurements which, as discussed earlier, are not currently suitable for wide scale screening.
- *The Integrated serum test – Combining First and Second Trimester Serum Markers:* Most of the benefits of the integrated test can be realized without the inclusion of NT measurements. A variant of integrated testing called the integrated serum test relies only on the serum PAPP-A measurement in the first trimester [12], in combination with the second trimester four-marker quad test (AFP, uE3, hCG, and DIA). This combination is projected to detect approximately 70 percent of Down syndrome cases with a 2.1 percent false positive rate. The integrated serum test would be a significant improvement over the current standard of care, the triple test. Table 1 compares the expected screening performance of the triple test with the performance of the proposed integrated serum test when applied to a hypothetical population of 100,000 women.

Table 1. A Comparison of Two Down Syndrome Screening Protocols in a Hypothetical Population of 100,000 Pregnant Women

	Current Triple Test	**Proposed Integrated Serum Test**
Number of Down syndrome pregnancies	154	154
Number of Down syndrome detected	116	116
Down syndrome detection rate	75%	75%
Number of unaffected pregnancies	99,846	99,846
Number of unaffected pregnancies positive	6,889	2,097
False positive rate	6.9%	2.1%
Amniocenteses per Down syndrome detected	59	18
Procedure related losses (@1:200)	34	10
Down syndrome detected:Unaffected fetus lost	3:1	12:1

The most important benefit of the lower false positive rate for the integrated serum test is the dramatic reduction in the amniocentesis/karyotype rate, while still maintaining the detection rate. Although amniocentesis and karotyping laboratories are widely available and extremely reliable, these procedures cost $1000 or more, and procedure related fetal losses do occur. A randomized trial estimated that 1 in 110 fetuses are lost due to second trimester amniocentesis [3]. In the United States, the procedure-related loss rate is often quoted to be 1 in 200. With the triple test, 59 amniocenteses are required to detect each case of Down syndrome. With the integrated serum test, this number is reduced to 18 amniocenteses per case detected. This is a major advance, especially when compared to the original method for screening (asking a woman her age), which requires 150 amniocenteses per case of Down syndrome detected. The only requirement for the integrated serum test beyond that of routine second trimester screening is a serum sample that can be easily collected at the time of the first prenatal visit. The integrated serum test provides an alternative screening method with excellent screening performance for those women without access to specialized centers that perform NT measurement.

III STUDY DESIGN AND METHODS

A. Study Design Overview: This proposal was a demonstration project to test the feasibility of combining first and second trimester maternal serum biochemical measurements into the integrated serum test. This integrated serum test is aimed at maintaining the high Down syndrome detection rate of about 70 to 75 percent, while at the same time significantly reducing the number of women with a screen positive test results. Performance is compared to the current standard of care in the United States, the triple test. The study population for the intervention trial is drawn from women in Maine receiving prenatal care in the first trimester. After informed consent, women are asked to provide a serum sample (or an aliquot of blood drawn for other testing) that was transported to the laboratory at FBR along with patient identification, demographic and other pregnancy-related information. The information is entered into our secure patient record system, and the sample is frozen for later PAPP-A testing. Once the PAPP-A measurements are obtained, they are entered into a separate database for later linkage with pregnancy information and the second trimester serum measurements (i.e., AFP, uE3, hCG and DIA). A receipt report is issued stating that the first trimester sample and information have been received. That report also reminds the physician that a subsequent blood sample is also needed in the second trimester to complete the integrated serum testing process. When the second serum sample is received, the four additional marker measurements are entered into the patient records system. A specialized computer algorithm matches first and second trimester assays results (the matching algorithm and it's performance will be described in more detail later). Results from all five maternal serum measurements are then combined with maternal age to determine the pregnancy-specific Down syndrome risk. A final computerized report is generated and sent to the providers. Positive results are communicated by phone and fax. Screen positive women are managed according

to protocols that are routinely used for second trimester women. Typically, these women are referred for perinatal evaluation using high-resolution ultrasound examination with possible amniocentesis, enabling the fetus to be karyotyped. The projected benefit to the women being screened by the integrated serum test is a significant decrease in the chance of being identified as screen positive (compared to the triple test) while high detection is maintained.

B. Study Population: The study population for the demonstration project is drawn from the estimated 10,000 to 11,000 women in Maine annually receiving prenatal care in the first trimester. Approximately 14,000 births occur annually in the state, with approximately 80 percent receiving prenatal care in the first trimester. The population of Maine is 97 percent Caucasian (White) with the remaining 3 percent being African American (Black), Asian, Hispanic, and Native American. Approximately one-half of the population has more than 12 years of education. Table 2 contains relevant information about racial/ethnic background and number of years of education for the state of Maine from an analysis in 1996 to 1997. The screening program is equally accessible to the entire population in the state.

Table 2. Racial/Ethnic Background and Number of Years of Education for the Study Population

Years of School	White	Asian	Native American	Black	Hispanic	Other or Unknown	Total
< 12	3,025	68	55	32	0	6	3,186
12	10,190	112	73	72	0	17	10,464
> 12	13,143	126	63	63	0	39	13,434
Total	26,358	306	191	167	0	62	27,084

C. Methods:

Creating a Steering Committee and Defining the Roles of the Study Staff: A Steering Committee consisting of the Medical Director, Director of Biometry (and his Data Coordinator), Study Coordinator, Prenatal Screening Laboratory Supervisor, the Director of Computer Services (and his Senior Programmer) was formed by the Principal Investigator. This Committee initially met every two weeks to assign responsibilities and tasks. Subsequently, the Committee met weekly to review progress, identify problems, and assign duties. The Study Coordinator is responsible for the collection and distribution of meeting notes.

Developing and Validating Physician and Patient Educational Materials: Under direction of the PI and the Steering Committee, the Study Coordinator developed patient and physician materials. The physician brochure includes a description of the study, its purpose, inclusion and exclusion criteria, limitations, and the benefit to the patient of participation. A two-sided laminated “quick reference guide” for physician staff and phlebotomists was also developed that included a reduced copy of the two part requisition form with annotations to assist staff in completing the form. Experienced prenatal care providers reviewed these materials, and minor changes

were made based on their recommendations. The quick reference guide was well received by office staff who found it to be particularly helpful. The patient material explained the study, the need for two blood samples, the need to delay reporting until the second trimester, the importance of sampling at specific times in gestation, and the benefits to the patient. The brochure was evaluated by the Fry Readability Index and was found to be at the 8th grade reading level. The draft brochure was given to 10 prenatal patients at a local clinic for review, and their suggestions led to minor changes.

Modification of Computer Software for Patient Management: The existing clinical patient record system software was modified to accommodate the integrated serum test. This effort focused on developing data processing procedures for log-in and for matching the first and second serum sample, and included the following items:

- a database for sample matching that included first trimester maternal serum sample log-in with a unique sample ID, patient demographics and documentation of informed consent.
- a sample receipt report to be sent to physicians after receipt of the first trimester sample. It records patient demographics, documents sample receipt, and reminds the office that a second trimester sample is needed to complete the integrated serum test protocol (IST).
- preprinted freezer labels with a unique sequential identifier that is used to match a patient's first trimester sample to her second trimester sample when it arrives.
- a computer program that examines all second trimester test requests to determine whether a likely first trimester sample match exists. Data linked after the computer match is visually verified.
- an overdue report lists all unmatched first trimester samples when gestational age is 18 weeks' or later. A reminder is then faxed to the physician. If no match is found by the 20th week, physicians are contacted and documentation of patient status is recorded in the research database.
- a PAPP-A assay run sheet is generated daily for retrieving matching first trimester samples from the freezer, setting up the assay, and for merging the results with the second trimester data.
- a Down syndrome risk algorithm, which combines maternal age and the results of the five serum assays.
- a revised patient reporting system that displays the PAPP-A results and includes them in the interpretation.
- an eight page handbook (with 11 appendices) that describes the procedures listed above.

The matching algorithm created for this project utilizes the unique preprinted sample receipt number, patient's first and last name, date of birth, ordering physician, and estimated date of delivery. A composite score ranging from 45 (poor match) to 99 (perfect match) is calculated for each data pair (one from the first trimester and one from the second trimester) based on their similarity. At a cut-off score of 57 out of 100, an estimated 99.9 percent of all second trimester samples for which a first trimester sample were received will be matched. All potential matches are individually reviewed by laboratory staff for verification.

Development of Two-part Laboratory Requisition Form: A form for ordering the integrated serum test was created by making minor modifications to an existing second trimester testing requisition slip and adding a removable top page. Both pages have a unique preprinted identification number that aids in matching. The top page also contains basic demographic and pregnancy-related information along with documentation of informed consent. When the integrated serum test is selected, the mark is automatically transferred onto the second page of the form, indicating that the sample is part of the integrated serum test. Once the top page is completed, it is separated and accompanies the serum sample to the laboratory. The remaining portion of the form is retained by the physician and subsequently used to request the second trimester portion of the integrated serum test.

PAPP-A Assay and establishment of reference (median values) data: The PAPP-A kit used in this study is manufactured by Diagnostic Systems Laboratories, Inc (Webster, TX). We previously validated this assay using a case-control study design and showed that Down syndrome screening performance using these reagents was consistent with published studies using other analytic methods. Two quality control samples are run with each assay (a low control with a value corresponding to that found in Down syndrome pregnancies, and a normal control with a value corresponding to that found in unaffected pregnancies). Between-assay CVs for the low and normal controls are 5.7 percent and 4.2 percent, respectively. Initial median values were derived by assaying approximately 200 sera obtained from another laboratory that had established medians using the DSL kit. Our results were compared to their results by regression analysis. The results of this analysis were then used to adjust our colleague's median levels (established using 11,000 samples) to derive initial gestational age specific medians for our study. Approximately two months after enrollment began, a new set of PAPP-A reference ranges were computed (median values) and entered into the clinical patient records system. Because of the wide range of gestational ages (8 to 13 weeks) compared to most published studies (10-13 weeks), a refined regression model using a log-quadratic function was utilized, because it provided a better fit to the data than the previously used log-linear model.

Collection of Outcome Information: After enrollment was completed and the pregnancies had delivered, outcome information was sought on all pregnancies. To identify Down syndrome cases diagnosed as a result of the screening process, outcome information was sought from those few centers performing karyotypes. This included both prenatal samples collected in the second trimester and results from blood samples taken soon after birth. This latter group is likely to have contained those cases missed by the screening process and diagnosed after birth. This information will be also be sought from the Bureau of Vital Record, State of Maine. FBR has a cooperative three-year agreement with the State of Maine to obtain outcome information on Maine pregnancies. In instances were verification is needed, individual health care providers may be contacted. All of the outcome data, along with the clinical information contained in patient records, will be transferred into a research database. The data will then be thoroughly checked, and any inconsistencies or missing data will be obtained and verified. Access to this

database will be limited to study personnel, and no individual patient information will be released.

Patient Satisfaction Survey. Reviewers of this grant expressed concern that women opting for the integrated serum test might be reluctant to wait four to eight weeks after providing a blood sample in the first trimester before they received the final Down syndrome screening results in the second trimester. To address these concerns, a three-part questionnaire consisting of 15 questions was developed in conjunction with one of the study consultants who is experienced in evaluating patient responses to prenatal screening. The questionnaire was divided into three parts: 1) assessment of basic knowledge and where it was obtained, 2) attitude about the integrated serum test, and 3) patient's experience of traditional stand alone second trimester testing as compared to the integrated serum test. The intent of the survey was not to assess whether women experience anxiety when screened. This is well established. Our aim was to determine if any additional anxiety is experienced because of the wait involved with the integrated serum test. Consequently, only women with screen negative test results were evaluated. Initially, the questionnaire was sent to 10 women who meet the following criteria: 1) completed the integrated serum test protocol, 2) delivered a live born infant, and 3) had undergone standard second trimester screening of a past pregnancy within the last 3 years. The intent was to determine how well the questions were answered, and whether any particular question caused confusion. Results were obtained from all 10 women either by return mail or, in a few cases, by a follow-up phone call. There were no missing data. Results of this preliminary study were encouraging. Women were aware that they had been tested using the integrated serum test, but did not report any additional anxiety. Some women appeared to be confused by a true-false question asking if "one of the benefits of the integrated serum test was to reduce the false positive rate"? They answered "Unsure". This question was modified to say that the test was "more accurate.

The modified questionnaire was then sent out to 30 additional women who had delivered their babies. After several weeks, women who had not responded were called to determine the reason. Women contacted by phone were given a choice of filling out the questionnaire and sending it in, or answering the questions directly on the phone. Ninety percent opted to send in the questionnaires, and the remainder answered the questions on line. Approximately one year later, the questionnaire was sent to another group of 30 women, but in this case to those in their third trimester of pregnancy. The goal was to determine if women's responses might be different if they had not yet had a successful delivery of their baby.

D. Statistical Techniques Employed

Primary Data Analyses: The primary aim of the study was to document that the integrated serum test could be successfully implemented in a variety of primary care settings (including rural) through a centralized administrative program. This allows a comparison of the false positive rate with the current standard of care - the triple test. The success of the effort has been analyzed as follows:

- *Integrated serum test Uptake Rate in the First Trimester*: This rate is the number of test requisition test forms with signed informed consent received at the Foundation's Laboratory, divided by the estimated total number of pregnancies seen for care in the first trimester (approximately 11,000). Given that about two-thirds of all pregnancies in Maine are currently screened in the second trimester, we expect that, by the end of the study, up to 50 percent of all pregnancies (85% of those two-thirds) will opt for the integrated serum test.
- *Second Trimester Sample Submission Rate:* Of all women who provide a first trimester sample, most are expected to provide a second trimester sample. The number of matched first and second trimester samples divided by the total number of first trimester samples will be the submission rate. We will document reasons why the second trimester sample did not arrive.
- *Percent of Integrated Serum Tests Reported Out:* This will be the number of integrated serum tests successfully matched, where both samples fall within the acceptable gestation age window. The acceptable range of gestational ages for the first trimester is 8 to 13 weeks, and 15 to 21 weeks for the second trimester sample.
- *The Initial Screen Positive Rate for the Integrated serum test :* Based on mathematical modeling, the initial positive rate is expected to be 2 to 3 percent in the general pregnancy population. Screen positive is the term often used to indicate that the risk is positive, meaning that it is not known whether the pregnancy is affected (true positive) or unaffected (false positive). However, since the prevalence of Down syndrome is low, the screen positive and false positive rates are virtually the same. The initial screen positive rate in the study is defined as the number of women receiving a risk of 1:100 or greater, divided by the total number of women successfully screened, before any further testing (such as ultrasound measurement of gestational age) is performed.
- *The Revised Screen Positive Rate for the Integrated serum test :* The first step in the diagnostic testing protocol for a screen positive pregnancy is to confirm the gestational age. The integrated serum test, like the triple test, is more likely to assign a high risk for Down syndrome, if the pregnancy is incorrectly dated further along than it really is. When a misdated pregnancy is identified, the risk is recalculated and a revised estimate provided. In our study population, we expect 30 to 40 percent of the pregnancies to be dated by last menstrual period. Once ultrasound reclassification due to incorrect dating is accomplished, the revised positive rate will also be computed.
- *Influence of the Maternal Age Distribution on the Expected Positive Rate:* Because the maternal age of the women accepting testing might differ from that in the general pregnancy population (e.g., they may be older), it is possible that the observed screen positive rate may be higher than the 2 to 3 percent expected. As age increases, both the false positive rate and detection rate increase. However, this increase is much less than that seen with the triple test.

Secondary Data Analyses: A secondary aim is to document that the Down syndrome detection rate found in the study is consistent with that predicted by modeling. Based on follow-up information available for the entire study cohort, we can estimate our Down syndrome detection rate using the following formula. This formula takes into account the well described rate of Down syndrome fetal loss

(23%) occurring between the early second trimester and term. Among screened pregnancies, most of the cases of Down syndrome are identified via amniocentesis and karyotype. The remainder will be identified after being born. In some of these cases screen positive women will have refused amniocentesis, others will occur in screen negative mothers.

C_2 = Cases of Down syndrome detected in the second trimester
C_T = Cases of Down syndrome detected at term
Detection Rate (%) = $C_2 / (C_2+(C_T / (1-0.23)))$

The confidence limits on this estimate of the detection rate will be rather broad because only 10 to 15 Down syndrome cases will be expected in the screened group. A more robust estimate of the detection rate can be obtained by mathematical modeling. Modeling has been shown to reliably predict Down syndrome detection rates in many previous studies using two, three, or four serum markers[1,27,28]. The modeled detection rate will be compared with the observed detection rate to determine if the latter rate is consistent with expectation.

Cost Effectiveness of the Integrated Serum Test: The cost effectiveness of the integrated serum test has been calculated as follows:

- The costs of an ultrasound examination (for the purpose of dating the pregnancy of women with positive test results based on last menstrual period dating) and amniocentesis/ karyotype are calculated separately for the triple test and the integrated serum test, using the initial positive rates found in the current study. The overall procedure-related cost is reduced for the integrated serum test, because the screen positive rate is lower than for the triple test.
- The costs of performing the measurements of the markers are separately calculated for the triple test and the integrated serum test. The overall costs for the integrated serum test are higher, because a second serum must be drawn and transported to the laboratory, and the additional markers PAPP-A and dimeric inhibin A measured on all screened women.
- The total costs and the cost per women tested are then directly compared for the two methods of testing. The assumptions for the comparison are given in the table in the Results section.
-

IV. PRESENTATION OF FINDINGS (DETAILED)

A. Computer Matching of First and Second Trimester Samples:

An essential step in the integrated serum test is reliably matching the first and second serum samples (designated a true match). In some instances, no second trimester sample arrives and no match for a first trimester sample is possible (true non-match). There is also the possibility of false positive and false negative matches, and these will be discussed later. Table 3 shows the results of computer matching for the 11,159 women who submitted a first trimester sample (enrolled in the study). All matches were manually reviewed and verified by laboratory personnel. Non-matches were identified by a variety of methods both during and after the recruitment phase of the project. A second trimester sample was matched

for 759,862 of the 11,159 first trimester samples. Of these 9,862, 9,723 (98.6%) were correctly matched by the computer (true match). The remaining 139 women (1 in 70) were falsely matched by the computer, but correctly identified and the mismatch rectified by laboratory personnel (usually 15 to 30 minutes per day was spent manually reviewing matches and contacting physician offices to rectify incorrect information). False matches occurred primarily because the information supplied by the physician office was incorrect (e.g. wrong date of birth, use of a new, rather the second half of the enrollment slip with its unique identifying number, use of hyphenated names, and misspellings). The computer correctly did not find a match for the 1,297 women who submitted only a first trimester sample. Nearly all of these women (1,291) had dropped out of the study. However, the computer algorithm and laboratory validation did not correctly match six women who actually should have received an integrated serum test (false negative matches). These six matches involve complex situations of multiple providers, name changes and, sometimes, multiple first trimester samples. While it is possible that some false negative or false positive matches are still undiscovered, it is our belief that the number is very small. As a whole, the computer matching was highly reliable.

Table 3. Results of Computer Matching of First and Second Trimester Serum Samples From the Same Woman

	True Match		
Computer Match	**Yes**	**No**	**Total**
Yes	9,723	139	9,862
No	6	1,291	1,297
Total	9,729	1,430	11,159

B. Study Subjects Enrolled and Initial Interpretations

The 27-month enrollment phase of the study began in August, 2001 and ended in July, 2003. During this time, 11,159 women provided informed consent and a first trimester blood sample (Table 3). During the 27 months of enrollment, a total of 18,308 women from Maine submitted at least one sample for prenatal screening. Thus, nearly 61 percent of the eligible population enrolled in the study. Figure 1 shows the type of testing completed in these women. A total of 9,723 women (87%) provided the required second trimester blood specimen. After computerized matching of the two specimens, 8,773 women (79% of the 11,159 initial enrollees) were found to have both the first and second trimester samples within the specified gestational age range for generating an integrated serum test report. This number is above the 7,000 completed integrated serum tests projected in our original grant submission. The 8,773 women with an integrated serum test result form the basis of many of the subsequent analyses.

Figure 1. Summary of Initial Interpretations Provided to 11,159 Women Enrolled in the Study

8,773 women (79%)
Received an integrated serum test result (5-marker)

950 Women (9%)
Received a second trimester quadruple test result
First trimester sample too early – 871
First trimester sample too late – 79

1,436 Women (12%)
No second trimester sample received – No test result provided
Spontaneous fetal loss – 575
Declined testing – 459
Elected amniocentesis in the second trimester – 236
Changed provider / residence – 133
Therapeutic termination - 29
Second trimester sample too late - 4

Of the 9,723 women providing both the first and second trimester samples, 950 did not receive an integrated serum test report (Figure 1). The most common explanation was that the first trimester sample was drawn prior to 8 weeks' gestation (871 women). This was considered too early for reliable interpretation. Only 79 of the 950 were outside the acceptable range because the first trimester sample was too late (for reliable interpretation after 14 weeks). These 950 women received an interpretation using the quadruple test, the best second trimester Down syndrome test currently available.

As part of the follow-up phase of this intervention study, we attempted to determine the reasons why a second trimester serum was not received for 1,436 of the initial 11,159 women who enrolled in the first trimester (Figure 1). Overall, this represents 12 percent of all study subjects. These women did not receive any Down syndrome risk estimate. For 40 percent of these women (575) a spontaneous fetal loss was identified after the first trimester sample had been submitted and before the second trimester sample was to have been collected. Another 459 women (32%) declined further testing for various, often unspecified, reasons. Another 236 women (16%) elected to have a diagnostic test (amniocentesis/karotyping) prior to submitting a second trimester sample. Overall, these 236 women represent 2.1 percent of the entire study population. Why they enrolled in a screening study and then chose diagnostic testing is unknown, but 87 percent were age 35 or older. It is likely that this group of older women rethought their decision about screening during the interval between the first and second trimester. A total of 133 women (9%) either moved out of state, or moved within the state and selected a new provider that was not participating or that did not know that the woman was already participating in our

study. Much smaller numbers of women chose selective termination (probably due to social reasons) or had a second trimester sample collected too late.

C. Demographic and Pregnancy-Related Information:

Table 4 contains selected demographic and pregnancy-related information for the 8,773 women who actually received an integrated serum test. Overall, the women were similar to the general pregnancy population in Maine. When birth records are finally released, it will be possible to more formally compare these women with pregnancies statewide. On average, the women waited nearly 7 weeks for the final integrated serum test results to be reported. Additional information that will be available in the near future (due to our agreement with the State of Maine Vital Records) includes birth weight, gestational age at delivery, and neonatal deaths.

Table 4. Selected Demographic and Pregnancy Related Information for the 8,773 Women Receiving an Integrated Serum Test

Characteristic	Result
Average maternal age (sd)	27.8 years (5.5)
Maternal age 35 or older at delivery	11.3 %
Average 1st trimester gestational age	10.0 weeks
Average 2nd trimester gestational age	16.9 weeks
Average time between samples (sd)	6.9 weeks (1.7)
Average maternal weight (sd)	164 (39)
Smokes cigarettes	13%
Maternal race – Caucasian	98%
Vaginal bleeding by the 2nd trimester	12%

D. Initial and Revised Initial Positive Rates:

Table 5 displays the initial and revised positive rates for the integrated serum test at two intervals of gestational age (8 and 9 weeks, and 10 to 13 weeks' gestation). The results are also stratified by initial method of gestational dating. The initial positive rate is the number of women who have a Down syndrome risk at or above 1:100 divided by the total number of women screened (8,773). It is, essentially, the false positive rate because affected pregnancies are relatively rare. When a woman with a pregnancy dated by last menstrual period (LMP) is given a screen positive test result, ultrasound examination is recommended to confirm her gestational age and to identify any obvious explanation for the positive results (twins, fetal demise, fetal anomalies, etc). If no explanation is found, the woman is counseled and offered a diagnostic procedure, typically amniocentesis. The percentage of women who remain screen positive after ultrasound examination is called the revised screen positive rate and is of importance because it corresponds to the women requiring diagnostic studies and more intensive follow-up in the health care system. Typically, the revised positive rate is 20 to 30 percent lower. When the initial dating is based on ultrasound measurements, there is little need for revision.

Our original modeling suggested that the initial positive rate for integrated serum testing at the cut-off level chosen (1:100 in the second trimester) would be approximately 2 to 3 percent. The overall initial positive rate for the entire group is 3.2 percent with a revised rate of 3.0 percent (Table 5, last row). This is just slightly higher than expected. The reason for this is clearly the result of LMP dated pregnancies, especially those enrolled at 8 or 9 weeks' gestation (Table 5, first row). These rates of 5.1 and 4.1 percent are significantly higher than expected, demonstrating the difficulty of interpreting measurements from pregnancies that are not correctly dated. PAPP-A values at 8 to 10 weeks gestation are rapidly changing (75 to 85% per week) at this time in gestation, and small errors in estimation of gestational age have the effect of increasing the screen positive rate. Revised rates are in general lower than initial rates, consistent with expectation. If the analysis were to be restricted to ultrasound-dated pregnancies, the positive rate of 2.7 percent is well within the expected range of 2 to 3 percent. These results indicate that pregnancies earlier than 10 weeks of gestation dated by last menstrual period will yield up to twice the rate predicted.

Table 5. Comparison of Serum Integrated Test Initial and Revised Positive Rates Stratified by Method of Dating and by Gestational Age of the First Trimester Sample

Initial Method Of Dating	Gestational Age (wks)	Number of Women	Positive Rate (%)[1]	
			Initial	Revised
LMP	8 to 9	1,976	5.1	4.1
	10 to 13	1,280	2.5	2.5
	All	3,256	4.1	3.5
US	8 to 9	2,892	2.7	2.7
	10 to 13	2,625	2.7	2.7
	All	**5,517**	**2.7**	**2.7**
Any	8 to 9	4,868	3.7	3.3
	10 to 13	3,905	2.6	2.6
	All	8,773	3.2	3.0

[1] Using a second trimester risk cut-off level of 1:100
LMP = last menstrual period, US = ultrasound

E. Comparing Initial Positive Rates – Integrated Serum Test and the Triple Test:
The primary goal of the current intervention trial is to demonstrate that the integrated serum test could provide important reductions in the false positive rate compared to the current standard of care - the triple test. The initial positive rate is influenced not just by the number and combinations of markers, but by the risk cut-off selected, the method of gestational dating (ultrasound versus last menstrual period), and the maternal age distribution of the screened population. Consequently, the most robust method of comparing the initial positive rates for the integrated serum test and the triple test is to perform a matched analysis using serum measurements from the

8,772 women who completed the integrated serum test. This is accomplished by first calculating a Down syndrome risk using only the measurements from the three markers that constitute the triple test (AFP, uE3, hCG), and then calculating a initial positive rate that would have been obtained at a specified risk cut-off. A risk cut-off of 1:190 has been selected, because the modeled Down syndrome detection rate at this cut-off level is equal to that achieved by the integrated serum test at a cut-off of 1:100 (the cut-off level used in the study and in Table 5). To be valid, comparisons of screen positive rates between different marker combinations must be made at a fixed detection rate.

Our modeling suggested that the false positive rate should be cut in half. As before, the initial positive rate is a reliable surrogate for the false positive rate. Table 6 shows that, overall, the initial positive rate dropped from 4.5 to 3.2 percent. This represents a 28 percent reduction in false positive, less than expected. However, when the comparison is limited to the ultrasound (US) dated pregnancies, the reduction is from 4.5 to 2.7 percent (a 40% drop). This is more consistent with the expected halving of the false positive rate. Again, this analysis shows the problems with interpreting LMP dated pregnancies. The problem is not as apparent in the second trimester triple test (4.6 and 4.5%) because the markers are less associated with gestational age, and methods have been devised to account for the more variable estimates associated with LMP dating. Although these methods can equalize the false positive rates, the detection rate for the triple test has been documented to be substantially higher, when the pregnancy is dated by ultrasound.

Table 6. A Comparison of the Initial Screen Positive Rate Using Serum Integrated Testing and the Triple Test in the Same 8,773 Women

Initial Method Of Dating	Initial Positive Rate (%)	
	Integrated Serum Test[1]	Triple Test[2]
LMP	4.1	4.6
US	**2.7**	**4.5**
All	3.2	4.5

F. Down syndrome Detection Rate and the Integrated Serum Test:

Our original grant proposed an integrated serum test cut-off level (1:100) that was selected to provide a Down syndrome detection rate equivalent to the triple test at a cut-off level of 1:190. Overall, the detection rate was expected to be about 70 percent. This is the lower end of the 70 to 75% detection quoted earlier, because one-third of screened pregnancies were expected to be dated by LMP. Screening is less effective when pregnancies are not dated by ultrasound. Given the relatively small number of low-risk pregnancies undergoing integrated serum testing (8,773) relatively few cases of Down syndrome are expected. Using the maternal age distribution of these women, and the second trimester age-associated risk, we estimate that there should have been 17.8 Down syndrome pregnancies. In order to find all Down syndrome pregnancies in this

group, we solicited information from two centers responsible for nearly all of the karyotypes performed for Maine pregnancies. In addition, we reviewed birth record reports and consulted with Maine genetic counselors. Given that the last screened pregnancy was delivered sometime in early 2004, some follow-up has not yet been completed. So far, we have identified 11 affected pregnancies with Down syndrome. Table 7 shows the observed proportions of Down syndrome pregnancies with positive test results by type of test and method of dating. Both detection rates are somewhat lower than expected, but statistically consistent with a detection rate of 70 percent. Although it is tempting to conclude that integrated serum screening had a higher detection rate than the triple test (65 versus 55%), the difference is due to one additional Down syndrome pregnancy detected. When follow-up is completed two or three additional cases may be identified. These data are consistent with expectations contained in our proposed project and support the finding that the detection rates are similar for these two protocols and are likely to be about 70 percent.

Table 7. Preliminary Estimates of the Down Syndrome (DS) Detection Rate by Testing Protocol, Method of Dating, and Gestational Age

Initial Method Of Dating	Gestational Age (wks)	Number of DS Pregnancies	Detection Rate (%)[1]	
			Triple	IST[2]
LMP	8 to 9	2	0	50
	10 to 13	1	100	100
	All	3	33	67
US	8 to 9	4	75	75
	10 to 13	4	50	50
	All	8	63	63
Any	8 to 9	6	50	67
	10 to 13	5	60	60
	All	**11**	**55**	**64**

G. Patient Satisfaction Survey:

The 15 questions in the two surveys along with the responses are given in Table 8. All 60 women who were asked to complete the survey did so. Answers were similar for almost all questions, and results from the two surveys of 30 women each have been combined. The table also provides the percent of the responders who gave a 'positive' response to the integrated screening test implementation. All 60 women remembered agreeing to have a blood test for risk of Down syndrome in the current pregnancy. Patients also had good recall about the integrated serum test (called the new blood test in the survey), with 59 of 60 women (98%) indicating that they had learned of the test from the physician, nurse, or informational pamphlet. Most (90%) understood that the test was a screening test and that the test would not rule out the possibility of having a baby with Down syndrome (80%). Fewer (63%) understood that the test would be less likely to give positive results. Nearly three-quarters (72%) understood that their final results would be given later. Two-thirds of

women (67%) disagreed with the statement that it was "hard to wait to get their result". However, one in five (20%) neither agreed nor disagreed with that statement. One in eight (12%) indicated that it was hard to wait for the test results. The questions about attitude toward the testing indicated that nearly all (95%) would agree to be tested in a future pregnancy, and that their health care provider left the decision up to them to have the test (98%). The questions about the comparison between the integrated serum test and their previous test indicated that only 5% worried more waiting to get the results of the integrated serum test than their previous test. More than two-thirds of women (69%) chose the integrated serum test because they though it was more accurate. The lowest response was that only about one-third (34%) thought they understood more about the integrated serum test than their previous test. Nearly all (95%) said that they would have the integrated serum test in a future pregnancy. It may be that women are not completely sure of the details of how screening works, but they can comprehend that the integrated serum test is better than a test offered previously. The key findings are:

- all 60 women remembered having the Integrated serum test
- almost all (98%) remembered having a prenatal test in their previous pregnancy
- three quarters of women indicated that they did not experience anxiety because they had to wait for final results until the second trimester, and
- almost all (95%) would consider the integrated serum test in a future pregnancy.

Table 8. Results of the Patient Satisfaction Survey

			Summary of Responses		
Question	**Positive Answer**	**Percent Positive**	**True or Agree**	**False or Disagree**	**Unsure**
I learned about the new blood test test[1] from the Doctor or Nurse, or by reading pamphlet	True	98 %	59	1	0
If results from the new blood test are negative, baby will definitely not have Down syndrome	False	80%	6	48	6
The new blood test will tell me the chance that my baby has Down syndrome	True	90%	54	4	2
The benefit of having the new blood test is that I am less likely to have positive test results	True	63%	38	14	8
I knew I would not get the final results from the new blood test until a second blood sample was tested later in my pregnancy	True	72%	43	6	11

It was hard to wait until after the second blood sample to get result	Disagree	67%	7	13	40
The decision to have the new blood test was a good one for me	Agree	88%	53	7	0
I was satisfied with the amount of information I received about the new blood test	Agree	85%	51	6	3
My health care provider left the final decision to have the new blood test up to me	Agree	98%	59	1	0
I would agree to be tested in a future pregnancy using the new blood test	Agree	95%	57	2	1
I worried more waiting for results of the new blood test than I did when I had my blood tested in a previous pregnancy	Disagree	76%	3	10	45
I chose the new blood test because it is more accurate than the blood test I had in my previous pregnancy.	Agree	69%	40	9	9
I understood more about the new blood test than I understood about the testing from my previous pregnancy.	Agree	34%	20	18	20

1 “new blood test” refers to the integrated serum test.

H. Validation and Reference Ranges for the PAPP-A assay:

The reagents used for the PAPP-A assay were purchased from a commercial source (Diagnostic Systems Laboratory, Webster, Texas). This assay is not currently licensed by the FDA for Down syndrome screening. The Clinical Laboratory Improvement Act of 1988 requires laboratories to validate the clinical usefulness of such assays prior to its use. The Principal Investigator created a case-control set of sera from 52 Down syndrome and 260 matched control pregnancies from a bank of frozen sera collected during an earlier first trimester study [29]. PAPP-A measurements were made on the case/control set using the DSL assay and another validated assay manufactured by Wallac Oy (Turku, Finland) [30]. The Wallac PAPP-A assay is widely used in Europe, and has been approved by the Maternal Fetal Medicine Foundation in London for measuring PAPP-A in first trimester sera. Table 9 shows the results for both the observed and modeled screening performance for the two assays. The two assays give virtually identical clinical screening performance, and either could be used. The DSL assay was chosen for reasons of convenience and availability in the United States.

Table 9. Observed and Modeled Down Syndrome Detection Rates at Two False Positive Rates for Two PAPP-A Assays

False Positive Rate (%)	Down Syndrome Detection Rate (%)	
	DSL Assay	Wallac Assay
Observed		
5	48	44
10	58	52
Modeled		
5	45	46
10	56	57

In addition to clinically validating the research assay, it was essential to carefully monitor the PAPP-A assay to ensure that it gives consistent performance on an ongoing basis. Figure 2 shows an example of the computation of reference data (medians) for the PAPP-A assay over the 8 to 13 week gestational age range where results were interpreted for clinical action. The data were best fit by an log-quadratic model which rises very steeply at 8 and 9 weeks (about 90% per week), and is somewhat less steep at later weeks. PAPP-A is by far the marker with the most gestational age dependence. In comparison, uE3 is the second trimester marker with the steepest slope of about 20 to 25 percent per week. This figure emphasizes the importance of obtaining a correct estimate of gestational age, preferably by ultrasound.

Figure 2. Gestational Age Dependence of PAPP-A Median Levels in the First Trimester. The horizontal axis shows the gestational age in completed weeks. The median PAPP-A result for each completed week (open circle) is plotted on a logarithmic vertical axis.

The appropriateness of medians used for interpretation is constantly monitored by a process called epidemiological monitoring of the screened population using a statistic called the grand MoM. [31] The grand MoM is the median value of all of the patients'

MoM values, and should be approximately 1.0 MoM over time. The goal is to keep this statistic within ten percent of that target (e.g., 0.90 to 1.10 MoM) 95% of the time. Figure 3a shows a temporal graph of the median MoM for PAPP-A for each of the 27 months of the enrollment period of the study. Most grand MoM values fall below 1.0 in the first half of the study more often falling than above the median. This suggests that the median values were slightly higher than what would be appropriate for the screened population. In the last half of the study results tended to fall equally above and below the 1.0 MoM line, consistent with the median values being appropriate for the screened population. The grand MoM falls outside the desired range (0.9 to 1.1 MoM) for 6 of the 27 time periods. This is more often than the one or two times expected by chance and is an indication of the difficulty we had in using this assay. For comparison, the same analysis is provided for unconjugated estriol (uE3), one of the markers that is part of the second trimester component of the integrated serum test. The uE3 assay is considered to be more reproducible and stable over time. This is reflected in Figure 3b. Except for the first and last month (where few data were collected) only one uE3 grand MoM was outside the expected range indicated by the horizontal dashed lines. Note, however, that results are not random, with a string of data above 1.00 MoM followed by a string of data falling below 1.00 MoM beginning at week 13, when an adjustment in median values was made. These results demonstrate that even for a long-established, well-controlled assays, systematic shifts in assays occur which, although within acceptable limits, are not random. The data indicate that the PAPP-A assay and the reference data were reasonable during the study, but improvements need to be made to bring it to the level of performance achievable with currently used screening assays.

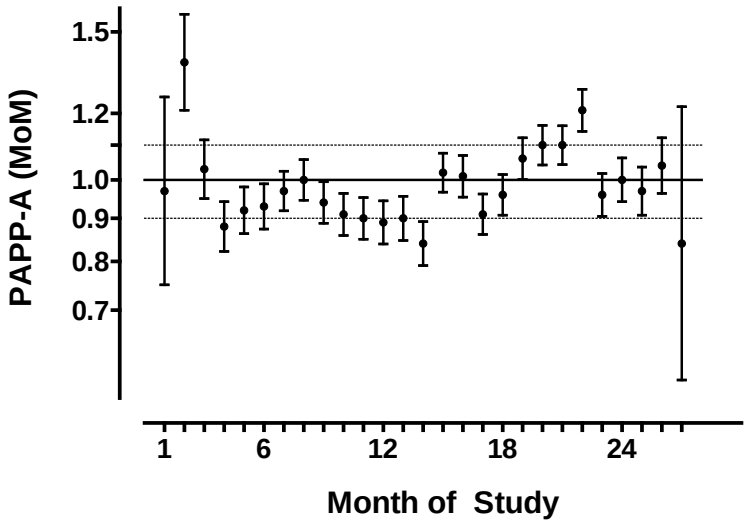

Figure 3a. The Monthly Median PAPP-A MoM Levels During the 27 Month Recruitment Period. The grand MoM (closed circle) and associated 95% confidence interval are shown on a logarithmic vertical axis for each of the 27 months of active enrollment (x-axis). The solid horizontal line at 1.00 MoM indicates the target value with the dashed lines indicating the acceptable range of grand MoM levels.

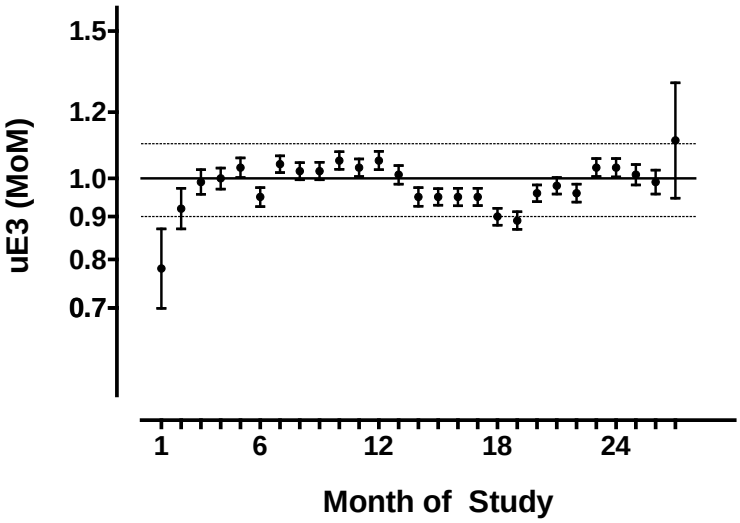

Figure 3b. The Monthly Median uE3 MoM Levels During the 27 Month Recruitment Period. The grand MoM (closed circle) and associated 95% confidence interval are shown on a logarithmic vertical axis for each of the 27 months of active enrollment (y-axis). The solid horizontal line at 1.00 MoM indicates the target value with the dashed lines indicating the acceptable range of grand MoM levels.

I. Cost Effectiveness:

Table 10 displays the additional cost components necessary for expanding the triple test to the quadruple test and to the integrated serum test along with documentation for some of the associated costs. Charges for these tests are likely to be higher. Large volume laboratories can reduce costs through economies of scale.

Table 10. Estimated Laboratory Costs Associated With Expanding the Triple Test to the Quadruple Test and the Integrated Serum Test

Component	Estimated Unit Cost ($)
Triple Test	
Develop Patient/Provider Educational Materials	One-time cost
Transportation/express shipment	$ 3.00
Sample receiving and handling	$ 4.00
Assaying AFP, uE3 and hCG[1]	$ 40.00
Computerized Interpretation	One-time cost
Accession and reporting	$ 2.00
Administrative costs (not counting overhead)	$ 5.00
Total laboratory costs for triple test	**$ 54.00**
Quadruple Test	
Modify Patient/Provider Educational Materials	One-time cost
Assay for DIA[2]	$ 15.00
Modification of Algorithm	One-time cost
Total laboratory costs for quadruple test	**$69.00**
Integrated Serum Test	
Modify Patient/Provider Educational Materials	One-time cost
Modification of Algorithm	One-time cost
Transportation/express shipment	$ 3.00
Sample receiving and handling	$ 4.00
Assay for PAPP-A[3]	$ 20.00
Additional administrative costs (e.g., matching)	$ 5.00
Total laboratory costs for integrated serum test	**$101.00**

[1] Includes costs of AFP ($2.50), uE3 ($2.00), and hCG ($2.50) reagents for singleton assays, depreciation of equipment, disposables, technician salary ($40,000) and benefits (21%), and a 50% overhead charge. It is assumed that 1 FTE can process 15,000 assays per year, including associated paperwork.

[2] Includes costs of DIA reagents ($6.00) and that an FTE could only process 10,000 assays per year.

[3] Includes costs of PAPP-A reagents ($7.00 – assayed in duplicate) and that an FTE could only process 7,000 assays per year. This cost might be reduced depending on assay improvement.

The additional laboratory costs to screen 10,000 women with the integrated serum test rather than the triple test is $470,000 ($1,010,00 - $540,000). To determine the monetary benefits, we assume the initial positive rates for ultrasound dated women of 2.7 percent for the integrated serum test and 4.5 percent for the triple test. The diagnostic costs include amniocentesis/karyotype ($1000), a high-resolution ultrasound ($200) and a half-hour genetic counseling session ($100). Assuming 100% uptake for all diagnostic testing, $585,000 will be spend on diagnostic testing for the 450 women with positive results on the triple test.

Fewer diagnostic tests are required if the integrated serum test is used since only 270 women have positive screening results ($351,000). This results in a net savings in diagnostic costs of $234,000. The screening cut-off level has been chosen (1:100 for the integrated serum test and 1:190 for the triple test) because the detection rate for Down syndrome is similar. This makes it possible to ignore the costs associated with missing (or detecting) a pregnancy with Down syndrome. Overall, implementing integrated serum testing is slightly more expensive ($236,000 per 10,000 women screened or $24 per woman). Given a procedure-related loss of 1:200, one unaffected pregnancy might be lose when integrated serum testing is used, while two procedure-related losses would be expected if triple testing were to be used in the same group of 10,000 women. Although difficult to quantify, 180 (40%) fewer women per 10,000 will experience the anxiety associated with a positive screening test result. Neglected in this analysis is the known positive relationship between the woman's individual Down syndrome risk and uptake of diagnostic testing. It is likely that more Down syndrome cases will be detected among the women screened using the integrated serum test, all of whom receive risks of 1:100 or greater. In contrast, many of the women with positive test results with the triple test will have risks between 1:100 and 1:190, and this group is known to be less likely to choose diagnostic testing.

Some other cost components are not included in this simple analysis of laboratory costs. Generally, they would qualify as health care costs. For example, the cost of blood drawing for the triple or quadruple sample will be incurred again if integrated serum testing is implemented, in order to collect the first trimester sample. There is likely to be additional provider costs associated with offering integrated serum testing as screening will need to be addressed during at least two visits. This preliminary analysis also did not address the issue of an important proportion of the population that will not complete the integrated serum testing process (e.g., a fetal death occurs before 15 weeks). In these instances, health care resources have been spent, with little or no return. Had second trimester testing alone been offered, these women would not have been included in the screening process.

V. DISCUSSION OF FINDINGS

A. Conclusions to be Drawn From Findings

Integrated Serum Screening is Associated with High Uptake in the General Pregnancy Population: The first important finding of this demonstration trial is that women in the general pregnancy population who appear for prenatal care in a routine setting can be offered serum integrated testing from a centralized laboratory. In this setting, a large proportion of women agree to undergo integrated serum screening. During the 27 months the study was operating, 18,301 women agree to some type of prenatal screening and 11,159 submitting a first trimester blood sample, representing 61 percent of all women tested. A total of 8,773 ultimately received a final report including a risk for Down syndrome based on the 5 serum markers included in the integrated serum test. This number substantially exceeds the 7,000 women projected to receive an integrated serum test interpretation in our original proposal.

A Higher Than Anticipated Proportion did Not Submit a Second Sample: Of the 11,159 women agreeing to be screened using the integrated serum test, 1,436 Women (12%) did complete the protocol because a second trimester sample was not received (Figure 1). The three most common reasons for not receiving a second sample were miscarriage, declined further testing, and electing diagnostic testing (amniocentesis). Overall, 575 miscarriages occurred prior to the second trimester. This represents 5 percent of all women enrolling in the study. This result is in line with publications reporting that 3 to 4 percent of pregnancies are miscarried from the 10 to 16 weeks of pregnancy [32]. It is also known that about 25 percent of Down syndrome fetuses are lost between the first and second trimester [33]. The women who declined further testing after submitting a first trimester sample may have changed their mind about having prenatal screening or perhaps did not fully understood what they were agreeing to in the first trimester. However, questionnaires submitted in two patient satisfaction surveys (see below) suggests that the latter explanation is less likely because most women seemed to have understood how the integrated serum test worked. Nearly 9 out of 10 women who elected amniocentesis were age 35 or older. This suggests that most of these high risk women wanted the reassurance that a karytotye gives in ruling out Downs syndrome.

A Higher than Anticipated Proportion of First Trimester Samples were Received with Gestational Ages Outside the Acceptable Range: A total of 950 of the 11,163 women (9%) did not receive an integrated serum test report because the first trimester sample was collected outside the acceptable gestational age window of 8 to 13 weeks of gestation. The vast majority of these were drawing prior to 8 weeks' gestation. Most of these were known to have been drawn too early (the gestational age based on LMP or US was initially reported to be prior to 8 weeks). A smaller proportion were thougt to be at 8 weeks' gestation or later based on LMP dating, but when a routine ultrasound study was performed later in pregnancy, the gestational age was revised and the sample was found to have been drawn too early. In the first trimester, about two-thirds of pregnancies are dated by the last menstrual period (LMP). In contrast, almost two-thirds of women have gestational age estimated by

ultrasound in the second trimester when the integrated serum test is interpreted. Ultrasound-based gestational dating is always used for interpretation if available. It is well established that dating based on LMP tends to overestimate gestational age [34].

Approximately 13 percent of first trimester samples were obtained prior to 8 weeks' gestation during the first few months of the study. As a way to reduce this proportion, we drafted a letter to all participating providers reminding them that samples sent in prior to 8 weeks could not be used. We suggested that drawing samples at 9 weeks or later would ensure that women would fall within the proper screening window. In recognition of the practices' routine, the letter included this statement. "We recognize that in some offices this approach may be incompatible with routine practice or with patient preference, and are only suggesting that this be considered if it does not disrupt office practice". We also stressed that it is important to be aware that if the first serum sample was collected outside the 8 to 13 week window, women are still screened with the second trimester quad marker test, the current best standard of care. Thus, these women's prenatal care was not compromised.

The Initial Positive Rates for Integrated Serum Test can be Lower Than Comparable Protocols: The aim of our original proposal was to offer a serum integrated test that reduced the false positive rate by about half, while maintaining a high Down syndrome detection rate. At that time, all modeling was based on 10 to 13 week pregnancies that had been dated by an early ultrasound. In our study, we allowed screening at 8 and 9 weeks, and in pregnancies dated by LMP. Among the US dated pregnancies receiving an integrated serum test, the initial positive rate was 2.7 percent. This is 40 percent lower than the 4.5 percent initial positive rate had the triple test been used in the same women. For this to be a fair comparison, the risk cut-off level was set so that the detection rates were the same. Among the pregnancies dated by LMP, the reduction in the initial positive rate was smaller. For the integrated serum test, the initial positive rate was 4.1% compared to 4.6% for the triple test (again, computed in the same group of women). The smaller reduction is most likely due to a proportion of these women having incorrect gestational age estimates. This results in incorrectly assigned PAPP-A MoM levels and correspondingly poor Down syndrome risk estimates.

Based on our findings, a reasonable policy would be to require that all pregnancies having an integrated serum test be dated by ultrasound prior to the second trimester interpretation.

The detection rates for the integrated serum and triple test meet expectation: The detection rate for the integrated serum test was expected to be about 70% for the both the integrated and triple test. The detection rate for the integrated serum test was found to be 64%, as contrasted with 55% for the triple test. However, these percentages should be viewed cautiously because the number of Down syndrome cases in the study population is small, and ascertainment may not yet be complete. This limited data is, however, consistent with our projection of a 70% detection rate for either of the two screening protocols.

Women Completing the Integrated Serum Screening Program are Informed and Satisfied: The patient satisfaction survey found that 85 percent of women were satisfied with the amount of information they received about the integrated serum test . The responses to the general knowledge questions were generally good (63% to 98%), particularly since women are being given a great deal of information at this time in their pregnancy. This suggests that the educational component surrounding the introduction of the integrated serum test was adequate and effective. Many women (72%) understood that they would have to wait until the second trimester to get the final results of the integrated serum test, but some (18%) were unsure. Only 7 of 50 women (12%) indicated that they did not understand this characteristic of the integrated serum test. Many women (67%) indicated that they did not find it hard to wait for their results, or, expressed no opinion (22%). Thus, only about 11% of women found it hard to wait for their test results. These two patient surveys indicate that women understand the benefits of the integrated serum test and most, but not all, do suffer unnecessary anxiety about waiting for results until the second trimester. The general level of satisfaction with the test seems high, as indicated by the fact that 95 percent would agree to be tested in a future pregnancy.

The PAPP-A Assay is Acceptable When Carefully Monitored, but Needs Improvement: The PAPP-A assay used for this study was used under the analyte-specific reagent (ASR) rule. No FDA approved kits are available in the United States. For this reason, we had to clinically validate the PAPP-A assay as required by the Clinical Laboratory Improvement Act of 1988. The assay performed satisfactorily, but required very careful monitoring of each new kit lot and frequent adjustment of patient values to keep results within acceptable limits [31]. The Prenatal Screening Laboratory at Foundation for Blood Research is highly specialized and has had 25 years experience in developing and optimizing immunoassays for prenatal screening. We were thus able to bring an intensity to monitoring the PAPP-A assay that will difficult for less experienced laboratories to match. The history of assays used for prenatal screening, beginning with AFP in the 1970s, has been one of continued improvement over time. Prenatal screening laboratories should bring pressure to bear on manufacturers to produce PAPP-A kits that are of the quality available for other screening assays, and to have the manufacturers obtain FDA approval.

A Matching Algorithm Can be Used Successfully with Laboratory Oversight: One of the more challenging aspects of the integrated serum test was developing and routinely using an algorithm for matching the first and second trimester samples from individual women enrolled in the study. We were able to take advantage of a prototype algorithm that had been developed for matching male and female buccal samples submitted for testing for cystic fibrosis mutations in our Molecular Genetics Laboratory. The matching algorithm in combination with laboratory oversight was successful 99.5 percent of the time (11153/11159). Laboratory personnel identified the false positive matches during the mandatory visual inspection step. Most of these occurred because of incorrect information supplied on the requisition slip. Incorrect information can also impact the interpretation of the screening results. For example, the date of birth is used in the matching algorithm, but it is also used to calculate maternal age. Maternal age is used to calculate the *apriori* risk for Down

syndrome. To the extent that these errors are identified by integrated serum testing, the screening process will be improved. The six incorrect (false negative) matches involved very complex situations that even laboratory personnel had difficulty discovering. Laboratories implementing the integrated serum test will need to develop and validate a matching algorithm and associated visual validation. The situation is quite different in some countries (e.g., Canada) where unique identification numbers make matching straightforward

The Additional Laboratory Costs of Integrated Serum Screening are Mostly Offset by the Savings from Reduced Diagnostic Testing: Our analysis shows that the added costs associated with moving from the triple test to the integrated serum test are not completely offset by saving due to reduced diagnostic testing (because of a lowered false positive rate). The additional marginal cost per patient screened by the integrated serum test is about $23. However, there are non-monetary benefits of integrated serum testing because fewer women would be identified with positive test results and half as many procedure-related fetal losses would occur.

Summary: The integrated serum test can be offered statewide as part of routine prenatal care and it is well accepted by pregnant women and health care providers. Women who elect integrated serum testing and whose pregnancy is dated by ultrasound will 40 percent less likely to have a false positive screening test at no loss in the Down syndrome detection. The cost to the health care system will be slightly higher than second trimester triple testing, but the reduction in maternal anxiety, and potential loss of healthy normal fetuses are additional benefits. If a screening program chooses to implement the integrated serum test we recommend that women only be offered the test if they have had an ultrasound confirmation of gestational age before the second trimester interpretation.

B. Explanations of limitations or Possible Distortion of Findings

The Foundation for Blood Research began the statewide prenatal screening program for neural tube defects in the United States in the late 1970's. Since that time, we have pioneered the introduction of Downs syndrome screening using AFP in the mid 1980s, the triple test in the early 90's, and the quad test in 2000. The current intervention study was conducted in the state of Maine where the physicians have historically been receptive to newer methods of prenatal screening. The integrated serum test requires the physician office and the screening program to work closely to bring together the various components of the protocol. Our long history with the physicians in Maine was a major reason why were successful in obtaining high uptake by patients. Other screening programs might be less successful in this effort, particularly the large commercial laboratories that screen a high percentage of the women in the United States. Thus, the integrated serum test may be more appropriate for academic screening programs at medical centers.

We found the process of matching samples to require specialized software and careful examination of proposed matched by laboratory staff. It sometimes required calls to individual health care providers. Given our relatively small number of sample processed (about 400 per month) difficult situations were infrequent. At large reference laboratories, the process of matching samples could be much more

difficult simply because of the large number of samples processed (Over 400 per day in some labs). For example, at our laboratory, it would be uncommon to have more one or two sets of women with the same name awaiting second samples at one time. This would be a much more common event at large laboratories.

Our relatively high proportion of second samples was made possible by have the computer provide reminders to all women who were at 18 weeks' gestation and had not yet provided a second trimester sample. Were other laboratories to implement integrated serum screening without this step, uptake rates may be lower.

C. Comparisons With Findings of Other Studies

The SURUSS study: When the current study was funded in 2001, little information was available on the performance of integrated serum testing in medical practice. This changed in 2003, when the results of the much anticipated Serum, Urine, and Ultrasound Screening Study (SURUSS) was published [24]. The goal of that study was to determine the most effective, safe, and cost effective combination of the currently available first and second trimester serum and ultrasound markers for prenatal screening for Down syndrome. SURUSS included 47,507 women at 25 maternity units (24 in the United Kingdom and 1 in Austria) with 101 cases of Down syndrome. The study was observational in the first trimester with intervention in the second trimester. Serum and ultrasound measurements were obtained in both trimesters allowing a direct comparison of first trimester screening (combined serum and ultrasound markers), second trimester screening (triple, and quad test), and integrated testing (full integrated test and the integrated serum test). The analysis included modeling the false positive rates that would be expected to achieve an 85% detection rate for the various Down syndrome screening methods. This analysis allows a direct companion of the reduction in false positive rate achievable with various test combinations at a fixed detection rate. SURUSS predicted a false positive rate of 2.7 percent for the integrated serum test, identical to the rate we found among the US dated pregnancies in our study. SURUSS predicted the triple test would have a false positive rate of 9.3 percent to achieve the same 85 percent false positive rate. The SURUSS report is not directly comparable to the summary results in the current study because:

- all pregnancies in SURUSS were dated by ultrasound (about one-third of ours were dated by LMP), and
- almost all pregnancies in SURUSS were screened at 10 to 13 weeks' of gestation (about one-half of ours were screened prior to 10 weeks' gestation).

Another study from Italy looked at just the false positive rate in 195 women, and concluded that the best screening combination was an integrated test that included nuchal translucency[35]. However, that study used a different combination of markers than Wald et al [24] and was too small to provide any meaningful comparisons. Another small study concluded that the integrated test as described by Wald was the best test combination for Down syndrome screening [36]. Another addressed the cost effectiveness of different methods of Down syndrome screening, and concluded that integrated test screening was the most cost effective method [37]. The conclusions of this paper were challenged by a number of letters to the Editor, but these were

addressed by the authors in their response [38]. They did not separately evaluate the cost effectiveness of the integrated serum test .

Taken together, these reports, particularly the SURUSS report [24], establish that the integrated test, whether full integrated (serum and ultrasound markers) or integrated serum (first and second trimester serum markers) offers the most discriminatory method for Down syndrome screening.

D. Possible Application of Findings to Actual MCH Health Care Delivery Situations (Including Recommendations When Appropriate) and Policy Implications

Pregnant women and their prenatal care providers are currently confronted with multiple choices for prenatal screening for Down syndrome. A recent editorial in the New England Journal of Medicine [39] entitled "Screening for Down syndrome — Too Many Choices?" addressed this issue. Sorting out the relative merits of age-based screening, first trimester screening, second trimester screening, integrated screening, and variants within each of these categories — coupled with dealing with the tradeoff between the risk of detecting an affected fetus and the risk of losing a normal baby when amniocentesis — can be daunting. The quality of published studies is highly variable because of poor study design or underascertainment of Down syndrome cases. Added to the mix are the often inflated claims of commercial laboratories that exaggerate Down syndrome screening performance to gain a competitive edge.

The Maternal and Child Health Bureau has, as part of its strategic plan, the goal of ensuring quality health care by utilizing evidence based research. Funding a comprehensive intervention trial of a proposed method of Down syndrome screening plays a critical role in providing the screening community with the best possible information for evidence-based policy-making.

E. Suggestions for further research

The current study provides evidence that one form of integrated screening (based solely on maternal serum markers) can be successfully introduced into routine practice in a distributed health care setting. One important finding is that an ultrasound-based estimate of gestational age is required for proper interpretation. Also, nearly one in 10 women who wanted serum integrated screening provided a first trimester sample that was too early to be interpreted. Both of these issues could be addressed by having a first trimester ultrasound examination. In addition, were a first trimester ultrasound to be routinely performed, it would be possible to also include a measurement of nuchal translucency (NT) measurements to improve performance even further. However, introducing NT measurements in to routine practice in the United States is far different from in Europe or in Canada.

We suggest that MCH consider funding a prospective trial of fully integrated testing for Down syndrome in routine practice – something that has not yet been tried in the United States. The aim of the study would be to develop training and ongoing quality assurance measures

for NT measurements that can be combined with integrated serum testing that will result in the most effective Down syndrome screening program that would be cost-effective and that would be available to all pregnant women.

VI. LIST OF PRODUCTS (peer reviewed articles, books, chapters in books, master and doctoral dissertations, conference presentations, etc.)

Knight G.J. Results from the Integrated serum test Study: A U.S. Screening Project. presented at Prenatal Screening for Down Syndrome: Introducing the Integrated Test into Medical Practice. Brown University, Rhode Island. March 28-29. 2003

Knight G.J. Down syndrome Screening: What's New. Maine Medical Center. Obstetrics and Gynecology Grand Rounds. November 14, 2002.

Knight GJ. Integrated serum screening in Maine. Down's Screening News. February 2002. Leeds University, UK. Editor P.bloom@leeds.ac.uk.

Knight GJ, Palomaki GE, Haddow JE. Integrated serum screening for Down syndrome: An intervention trial involving 11,159 women. Manuscript in preparation to be submitted to the New England Journal of Medicine.

VII. LITERATURE CITED

1. Wald NJ, Cuckle HS, Densem JW *et al.* Maternal serum screening for Down's syndrome in early pregnancy, *BMJ* 1988;**297**:883-7.
2. Palomaki GE, Knight GJ, McCarthy JE, et al. Maternal serum screening for Down syndrome in the United states: a 1995 survey. *Am J Obstet Gynecol* 1997;**176**:1045-51.
3. Tabor A, Madsen M, Obel EB, *et al*. Randomized controlled trial of genetic amniocentesis in 4606 low risk women. *Lancet* 1986;**I**:1287-93.
4. Christie WR. Anxiety associated with alpha-fetoprotein prenatal screening. Proceedings of the Second Scarborough Conference (Sept 5-7, 178). Alpha-fetoprotein serum screening in pregnancy. Foundation for Blood Research, Scarborough, ME
5. Marteau TM, Cook R, Kidd J, *et al.* The psychological effects of false-positive results in prenatal screening for fetal abnormality: a prospective study*. Prenat Diagn* 1992;**12**:205-14.
6. Statham H, Green J. Serum screening for Down's syndrome: some women's experiences. *BMJ* 1993; 307:174-6.
7. Goel V, Glazier R, Summers A *et al.* Psychological outcomes following maternal serum screening: a cohort study. *CMAJ* 1998;**159**:651-6.
8. Kornman LH, Wortelboer MJM, Beekhuis JR, Morrssink LP, Matingh A. Woman's opinions and the implications of first- versus second- trimester screening for fetal Down's syndrome. *Prenat Diagn* 1997; 17:1011-18
9. Texixeira JM, Fisk NM, Glover V. Association between maternal anxiety in pregnancy and increased uterine artery resistance index: cohort based study. *BMJ* 1999; 318:153-7.

10. Perkin M. Letter. Cited studies did not show relation between maternal anxiety and birth weight. *BMJ* 1999; 318:1288-89. Haddow JE,
11. Palomaki GE, Knight GJ, Foster DL, Neveux LM. Second trimester screening for Down's syndrome using maternal serum dimeric inhibin A. *J Med Screen* 1998;**5**:115-9.
12. Wald NJ, Watt HC, Hackshaw AK. Integrated screening for Down's syndrome based on tests performed during the first and second trimester. *N Engl J Med* 1999;**341**:461-7.
13. Report of the UK Collaborative Study on alpha-fetoprotein in relation to neural - tube defects. Maternal serum alpha feto-protein measurements in antenatal screening for anencephaly and spina bifida in early pregnancy.; *Lancet* 1977;**i**:323
14. Haddow JE, Kloza EM, Smith DE, Knight GJ. Data from an alpha-fetoprotein pilot screening program in Maine. Obstet Gynecol 1983;**62**:556-60.
15. Haddow JE, Palomaki GE, Knight GJ, *et al.* Prenatal screening for Down's syndrome with use of maternal serum markers. *N Engl J Med* 1992; 327:588-93.
16. Palomaki GE, Neveux LM, Haddow JE. Can reliable Down's syndrome detection rates be determined from prenatal screening intervention trials? *J Med Screen* 1996;**3**:12-7.
17. FBR/CAP Maternal Serum Screening Survey. Set FP-C. 1999. College of American Pathologists, Waukegan, Ill.
18. Cuckle HS, Holding S, Jones R, *et al*. Combining inhibin A with existing second-trimester markers in maternal serum screening for Down's syndrome. *Prenat Diagn* 1996;**16**:1095-100.
19. Wald NJ, Kennard A, Hackshaw A, McGuire A. Antenatal screening for Down's syndrome. *J Med Screen* 1997;**4**:181-246.
20. Wald NJ, Stone R, Cuckle HS, *et al*. First trimester concentrations of pregnancy associated plasma protein A and placental protein 14 in Down's syndrome. *BMJ* 1992;**305**:28.
21. Nicolaides KH, Azar G, Byrne D, Mansur C, Marks K. Fetal nuchal translucency: ultrasound screening for chromosomal defects in first trimester of pregnancy. *BMJ* 1992;*304*:867-9.
22. Snijders RJM, Noble P, Sebire, *et al*. UK Multicentre project on assessment of risk of trisomy 21 by maternal age and fetal nuchal-translucency thickness at 10-14 weeks of gestation. *Lancet* 1998;*351*:343-46.
23. Wald NJ, Hackshaw AK. Combing ultrasound and biochemistry in first-trimester screening for Down's syndrome. *Prenat Diagn* 1997; 17: 821-29.
24. Wald NJ, Rodeck C, Hackshaw AK, *et.al.* First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine, and Ultrasound Screening Study (SURUSS). 2003; *Health Tech Assess* 7(11).
25. Wapner R. Thom E, Simpson JL, *et al.* First-trimester screening for trisomies 21 and 18. *N Engl J Med*. 2003;**349**:1405-13.
26. Wald NJ, Hackshaw A, Stone R et al. Serum alpha-fetoprotein and neural tube defects in the first trimester of pregnancy. *Prenat Diagn* 1993;**13**:1047-50.
27. Haddow JE, Palomaki GE, Knight GJ, et al. Prenatal screening for Down's syndrome with use of maternal serum markers *N Eng J Med* 1992;327:588-93.
28. Palomaki GE, Neveux LM, Haddow JE. Can reliable Down's syndrome detection rates be determined from prenatal screening intervention trials? *J Med Screen* 1996; 3:12-17.

29. Haddow JE, Palomaki GE, Knight GJ *et al*. Screening of maternal serum for fetal Down's syndrome in the first trimester. *N Engl J Med* 1998; 338:955-61
30. Qui-Ping Q, Christiansen M, Oxvig C, Pettersson K, Sottrup-Jensen L, Koch C, Norgaard-Pedersen B. Double-monoclonal immunofluorometric assays for pregnancy-associated plasma protein A/proeosinophil major basic protein (PAPP-A/proMBP) complex in first-trimester maternal serum screening for Down syndrome. *Clin Chem* 1997;**12**:2323-32.
31. Knight, GJ, Palomaki GE. Epidemiologic monitoring of prenatal screening for neural tube defects and Down syndrome. In: Canick JA and Lambert-Messerlain, GM (eds). *Clin Lab Med* 2003;**23**:531-51.
32. Simpson JL, Gray RH, Queenan JT*, et al.* Low fetal loss rates after demonstration of a live fetuses in the first trimester. *JAMA* 1987;**25**:2555-7.
33. Cuckle H. Down syndrome fetal loss rate in early pregnancy. *Prenat Diagn*. 1999;**19**:1177-9.
34. Gardosi J, Mongelli M. Risk assessment adjusted for gestational age in maternal serum screening for Down's syndrome. *BMJ.* 1993;**306**:1509-11.
35. Rode L, Wajdemann R, Shalmi A, *et. al*. Combined first- and second-trimester screening for Down syndrome: an evaluation of proMBP as a marker. *Prenat Diagn* 2003;**23**:593-598.
36. Canini S, Perfumo F, Famularo L, Venturini PL, Palazzese V. Comparison of first trimester, second trimester and integrated screening results in unaffected pregnancies. *Clin Chem Lab Med* 2002; **40**:600-3.
37. Gilbert RE, Augood C,Gupta A, *et. al*. Screening for Down's syndrome: effects, safety, and cost effectiveness of first and second trimester strategies. *BMJ* 2001; **323**:423-5.
38. Screening for Down's syndrome. Letters to and rebuttal by Gilbert et. al. *BMJ* 2002;**324**: 110-112.
39. Mennuti MT, Driscoll DA. Screening for Down's Syndrome — Too Many Choices? *N Engl J Med* 2003;**349**:1471-3.

VIII. Abbreviations and Terminology

(AFP)	alpha feto-protein
(DIA)	dimeric inhibin A
(HCG)	human chorionic gonadotrophin
(IST)	integrated serum test
(LMP)	first day of the last menstrual period
(MoM)	multiple of the median
(NT)	nuchal translucency
(PAPP-A)	pregnancy associated plasma protein A
(NTD)	neural tube defects
(SURUSS)	**S**erum, **U**rine, **U**ltra**S**ound **S**tudy
(UE3)	unconjugated estriol
(US)	ultrasound

- Screening is the systematic application of a test or inquiry, to identify subjects at sufficient risk of a specific disorder to benefit from further investigation or direct preventive action, among persons who have not sought medical attention on account of symptoms of that disorder.
- Detection rate is the proportion of all affected pregnancies that have a positive test result (detection rate is equivalent to test sensitivity).
- False positive rate is the proportion of unaffected pregnancies that have a positive test result (the false positive rate is equivalent to 1-test specificity).
- Initial positive rate or the screen positive rate is the proportion of the screened population that receive a positive test result upon the initial interpretation. It is operationally equivalent to the false positive rate since only a small proportion of initial positives are true, rather than false positive results.
- Revised positive rate is the proportion of the screening population that remains positive even after a dating ultrasound has identified incorrect gestational dating and non-viable fetuses. In programs where most of the dating is by last menstrual period, the revised positive rate may be only one-half the initial positive rate (e.g., initial positive rate of 6% with a revised positive rate of 3%).
- Multiple of the median (MoM) is a method of normalizing assay results in which an individual result is divided by the value expected for the 'average' women of the same gestational age (and other factors as well). For example, if the median AFP value in 100 women at 16 weeks gestation is 20 IU/mL and an individual woman's result was found to be 40 IU/mL, the AFP result in that women is reported as 2.0 MoM (40/20). Median values used to create multiples of the median are assay- and laboratory-dependent.

- Grand MoM is the median value of a set of MoM values obtained from a screened population. By definition, the grand MoM in the general pregnancy population should be 1.00 (within statistical limits) if the median values are appropriate.
- Epidemiologic monitoring is the process of observing key population values that provide insight into how well the screening process is functioning. These values can relate to the assays used (e.g., assay medians and grand MoM), the population being tested (e.g., proportion dated by ultrasound and the maternal age distribution) or to the combination of the two (e.g., initial positive rates and detection rates).
- Positive predictive value is the proportion of women with positive test results that have an affected fetus expressed as a percentage.