

Technical standards and guidelines: Prenatal screening for Down syndrome that includes first-trimester biochemistry and/or ultrasound measurements

Glenn E. Palomaki, BA, BS¹, Jo Ellen S. Lee, PhD², Jacob A. Canick, PhD¹,
Geraldine A. McDowell, PhD³, and Alan E. Donnemfeld, MD⁴,
for the ACMG Laboratory Quality Assurance Committee

Disclaimer: These Standards and Guidelines are designed primarily as an educational resource for clinical laboratory directors to help them provide quality clinical laboratory genetic services. Adherence to these standards and guidelines does not necessarily ensure a successful medical outcome. These standards and guidelines should not be considered inclusive of all proper procedures and tests or exclusive of other procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific procedure or test, the clinical laboratory director should apply his or her own professional judgment to the specific clinical circumstances presented by the individual patient or specimen. It may be prudent, however, to document in the laboratory record the rationale for any significant deviation from these standards and guidelines.

Abstract: This statement is intended to augment the current general ACMG Standards and Guidelines for Clinical Genetics Laboratories and to address guidelines specific to first-trimester screening for Down syndrome. The aim is to provide the laboratory the necessary information to ensure accurate and reliable Down syndrome screening results given a screening protocol (e.g., combined first trimester and integrated testing). Information about various test combinations and their expected performance are provided, but other issues such as availability of reagents, patient interest in early test results, access to open neural tube defect screening, and availability of chorionic villus sampling are all contextual factors in deciding which screening protocol(s) will be selected by individual health care providers. Individual laboratories are responsible for meeting the quality assurance standards described by the Clinical Laboratory Improvement Act, the College of American Pathologists, and other regulatory agencies, with respect to appropriate sample documentation, assay validation, general proficiency, and quality control measures. These guidelines address first-trimester screening that includes ultrasound measurement and interpretation of nuchal translucency thickness and protocols that combine markers from both the first and second trimesters. Laboratories can use their professional judgment to make modification or additions. *Genet Med* 2009;11(9):669–681.

Key Words: Down syndrome, prenatal screening, nuchal translucency, maternal serum, first trimester

From the ¹Women and Infants Hospital and Albert Medical School of Brown University, Providence, Rhode Island; ²Genzyme Genetics, Santa Fe, New Mexico; ³Laboratory Corporation of America, Research Triangle Park, North Carolina; and ⁴Genzyme Genetics, Philadelphia, Pennsylvania.

Glenn E. Palomaki, BA, BS, Woman & Infants Hospital, 101 Dudley Street, Providence, RI 02905. E-mail: gpalomaki@ipmms.org.

Disclosure: The authors declare no conflict of interest.

Submitted for publication March 6, 2009.

Accepted for publication March 6, 2009.

Published online ahead of print August 17, 2009.

DOI: 10.1097/GIM.0b013e3181ad5246

This statement is intended to augment the current general ACMG Standards and Guidelines for Clinical Genetics Laboratories and to address guidelines specific to first-trimester screening for Down syndrome. The aim is to provide the laboratory the necessary information to ensure accurate and reliable Down syndrome screening results given a screening protocol (e.g., combined first trimester and integrated testing). In contrast, the ACMG Practice Guidelines are aimed at helping health care professional and laboratories to determine which screening protocols to offer. This document attempts to provide information about all reasonable test combinations and their expected performance. However, other issues such as availability of reagents, patient interest in early test results, access to open neural tube defect screening, and availability of chorionic villus sampling (CVS) are all contextual factors in deciding which screening protocol(s) will be selected by individual health care providers. Individual laboratories are responsible for meeting the quality assurance standards described by the Clinical Laboratory Improvement Act, the College of American Pathologists, and other regulatory agencies, with respect to appropriate sample documentation, assay validation, general proficiency, and quality control measures. Other ACMG Standards and Guidelines¹ address second-trimester screening for Down syndrome. This document addresses first-trimester screening and protocols that combine markers from both the first and second trimesters. The First Trimester Maternal Serum Screening for Down Syndrome section addresses first-trimester biochemical screening for Down syndrome. The Additional Down Syndrome Screening Protocols Using First-trimester Down Syndrome Markers section addresses ultrasound measurements and interpretation of nuchal translucency (NT) thickness, and Down syndrome screening protocols that combine information from the two trimesters (namely, integrated screening,² serum integrated screening,³ sequential screening, and contingent screening).^{4–6} As stated in the disclaimer, this document provides a set of standards and guidelines, but laboratories can use their professional judgment to make modification or additions. Through-

out the document, the term “serum” (e.g., collect a serum sample) is often used, and unless specifically noted, the term “dried blood spot” is interchangeable.

BACKGROUND ON DOWN SYNDROME

Brief clinical description

Down syndrome (Trisomy 21) (OMIM number: 190685) is one of the most common genetic causes of moderate to severe mental retardation. Virtually, all individuals with Down syndrome are hypotonic and have minor dysmorphic features that include upslanting palpebral fissures, epicanthal folds, flat nasal bridge, Brushfield spots of the iris, shortened, incurving fifth finger, and transverse palmar crease. Congenital heart disease is present in 40% of cases, and 5% have gastrointestinal anomalies such as duodenal atresia or Hirschsprung disease. The incidence of childhood leukemia is increased up to 20 times over that of the general population. Adults with Down syndrome experience neuronal degeneration identical to that present in Alzheimer disease. Individuals with Down syndrome without congenital heart disease can live beyond 60 years of age.

Etiology

Down syndrome is caused by the presence of an extra copy of Chromosome 21, either as a free chromosome, a Robertsonian translocation, or as a reciprocal translocation involving Chromosome 21. Approximately 95% of cases result from sporadic nondisjunction during parental meiosis. The nondisjunction is maternal in 95% of cases, and 77% of the maternal nondisjunction occurs during Meiosis I. The risk of having a child with Down syndrome increases with advancing maternal age. Down syndrome can be inherited when one parent carries a Robertsonian translocation, the risk to the offspring is dependent on the sex of the carrier parent (unless it is a 21:21 translocation).

Laboratory director

Although the prenatal screening laboratory uses clinical chemistry methods such as enzyme immunoassays, its function differs because the results require a unique kind of interpretation. These interpretations put the results of the test into the appropriate context of a priori risks as determined by maternal age, gestational age, and family history. The laboratory director is often called on to provide consultation regarding these risks and options for further action. To address these unique requirements, the laboratory director should meet the standards set out in Section B3 of the ACMG Guidelines. When prenatal screening for Down syndrome is carried out in a clinical chemistry laboratory in which the director does not meet these standards, the laboratory should have a demonstrated relationship with an individual who does meet the standards set out in Section B3. That person should be available to aid in interpretation and provide consultation when requested.

Screening versus diagnostic testing

First-trimester prenatal testing for Down syndrome using maternal serum markers in combination with sonographic fetal NT measurements is considered a screening test because results are not definitive. Asking a woman her age and then offering amniocentesis if she reports being age 35 years or older at delivery is also a screening test. The diagnostic test for Down syndrome is the karyotyping of fetal cells, usually obtained by CVS or amniocentesis. Newer diagnostic technologies using

only fluorescence in situ hybridization or quantitative fluorescent polymerase chain reaction may also have a place. The distinction between a screening and a diagnostic test is important, because the goals and expectations differ for clinical sensitivity and specificity, costs, and acceptable level of invasiveness. The interpretation of maternal serum markers and NT measurement is not diagnostic of any specific disorder. Rather, the screening process identifies pregnancies that are at sufficient risk for Down syndrome to warrant genetic counseling and the offer of additional diagnostic testing, such as CVS or amniocentesis and karyotyping. The clinical sensitivity and specificity of this screening test will be a function of several factors, including maternal age, number of analytes measured, assay precision, method of estimating gestational age, technique of obtaining an NT measurement, appropriateness of reference data, and risk cutoff level used to determine “screen-positive” results.

FIRST-TRIMESTER MATERNAL SERUM SCREENING FOR DOWN SYNDROME

Prenatal screening for Down syndrome is best implemented in the context of a comprehensive program that coordinates preanalytic, analytic, and postanalytic components of the process. This section will address first-trimester stand-alone Down syndrome risk assessments that combine maternal age-associated risk with interpretations of a fetal NT ultrasound measurement and the levels of first-trimester maternal serum screening markers. These serum markers include pregnancy-associated plasma protein A (PAPP-A) and one of various forms of human chorionic gonadotropin, including the free beta-subunit of human chorionic gonadotropin (free beta-hCG) and total or intact hCG (hCG). Although not common, some laboratories have chosen to add or substitute other markers (e.g., dimeric inhibin-A [DIA] or hyperglycosylated hCG [h-hCG]) in their first-trimester screening panel.

The choice of the forms of hCG (free beta-hCG, intact hCG, or total hCG) measured in the first trimester is dependent on the markers performance at each week of gestation (Screening Performance section, Table 1). Before 11 weeks, free beta-hCG is discriminatory but hCG is not. Between 11 and 13 gestational weeks, free beta-hCG is univariately a more discriminatory Down syndrome screening marker than hCG. When combined with maternal age, NT and PAPP-A, free beta-hCG performs better than hCG (2–3% higher detection) at 11 weeks. At 13 weeks, hCG may perform slightly better than free beta-hCG (1–2% higher detection). Only a limited number of laboratories provide testing using free beta-hCG, but these laboratories provide testing for a relatively large number of patients.

Patient and provider information

Patient information

Laboratories should either provide educational materials (e.g., brochures and videotape) for patient use in consultation with their providers or, at a minimum, provide information about where such materials can be obtained. Many laboratories and professional organizations (e.g., American College of Obstetricians and Gynecologists [ACOG], National Society of Genetic Counselors, regional genetics groups) have produced, and in some cases formally evaluated, materials that are in effective formats, at appropriate reading levels, and available in multiple languages. These materials provide general information about the disorder, test performance, patient rights, eligibility, test interpretation, treatment options, costs, risks and

Table 1 Expected Down syndrome detection rate for first-trimester combined screening at two different gestational ages, two different false-positive rates, and for two marker combinations

	DR (%) at 11 wks		DR (%) at 12 wks		DR (%) at 13 wks	
	1% FPR	5% FPR	1% FPR	5% FPR	1% FPR	5% FPR
Maternal age, NT, PAPP-A and						
Free beta-hCG ^a	NR	86	NR	84	NR	82
hCG ^a	NR	84	NR	83	NR	84
Free beta-hCG ^b	68	85	68	83	66	81
hCG ^b	62	82	64	81	64	81
Free beta-hCG ^c	NR	90	NR	88	NR	NR
hCG ^c	NR	87	NR	85	NR	NR
Free beta-hCG ^d	68	83	69	83	69	84
hCG ^d	64	81	67	83	71	86
Free beta-hCG ^e	74	87	72	84	66	80
hCG ^e	NR	NR	70	83	68	82

^aBased on the parameters from the FASTER study by Canick et al.¹⁷

^bFrom a meta-analysis published by Palomaki et al.²⁰

^cFrom a meta-analysis published by Evans et al.⁵²

^dBased on the parameters from the SURUSS study by Wald et al.^{40,51}

^eFrom a meta-analysis published by Cuckle et al.⁵³

NR, not reported; DR, detection rate.

benefits of testing, and what to expect if the screening test is positive.

Informational materials for health care providers

Laboratories should supply providers with informational materials that include the following:

- The sampling process and how samples should be labeled and transported.
- Samples of test requisitions that must accompany samples to provide information needed for identification and accurate test interpretation.
- General information on testing, such as laboratory turnaround time and whether results will be telephoned/faxed or mailed.
- Information about expectations for clinical test performance (sensitivity, specificity, and failure rate) and reporting formats.

Informed consent

Patients should be informed about the benefits and limitations of prenatal screening before testing. It is the duty of the health care professional, not the laboratory, to inform and obtain consent for testing, but the laboratory sometimes is required to document such consent (e.g., New York State). It is the laboratory's responsibility to provide sufficient information about prenatal screening to the health care provider to ensure that an appropriate specimen is obtained and to facilitate educating the patient and obtaining consent.

Requisition forms and intake information (also see Recommended Report Formats section)

For the most reliable interpretation, laboratories should have a mechanism to collect pretest clinical information that includes:

- Basic required demographic information (see section C2.4 and C3 [The 2008 edition of the ACMG Standards and Guidelines for Clinical Laboratories is available at the ACMG website: www.acmg.net]).
- Gestational age and method on which it is based (see Time of Testing to Incorporating Dating Method section).
- Maternal weight.
- Number of fetuses.
- Family history of Down syndrome (previous affected pregnancy).
- NT measurement (in mm).
- Crown-rump length (CRL) measurement (in mm).
- Identification of sonographer and sonographer's center.
- The laboratory may also choose to collect information regarding maternal race, presence of insulin-dependent diabetes before pregnancy, assisted reproductive techniques (e.g., age of egg donor, in vitro fertilization), and smoking status. At this time, however, there is no clear consensus as to whether, or how, these factors should be incorporated into the risk calculation.

The laboratory may choose to contact the provider if critical patient information does not accompany the specimen. If the laboratory does not obtain this information, the written report should indicate that the information is missing and what information, if any, was used in the interpretation. In some cases, including information on the report about the potential impact of the missing information may be warranted (e.g., maternal weight); in other cases, full interpretation may not be possible (e.g., no maternal age or no gestational age).

Specimen collection and transportation

Specimen collection

Blood samples should be collected using standard phlebotomy techniques. The laboratory should specify to clients what samples are acceptable (e.g., whole blood, serum separator tube, centrifuged serum separator tube, and dried whole blood spots). For most assays, serum is the most common sample type. When compared with serum, dried blood spots may yield lower screening performance for PAPP-A and hCG⁷ but may improve performance for the free beta-subunit of hCG, because of stability issues.⁸ Dried blood spots also have the advantages of less invasive collection methods and greater ease of transportation. Specimen containers should be labeled with the patient's name and draw date.

Specimen transportation

Acceptable specimen handling from collection site to the laboratory should be specified, including packaging, mode of transportation (e.g., courier, standard mail, overnight transport), and temperature range (the serum and dried blood spot markers are relatively stable⁹⁻¹¹ and either sample type can be shipped at ambient temperature). See also Sample Stability section.

Specimen processing and storage

Criteria for sample rejection

Variables that can affect the acceptability of a sample for Down syndrome screening should be established by the laboratory and may include both clinical (e.g., gestational age out of range) and sample-related characteristics (e.g., inappropriate

sample type, insufficient quantity, and gross hemolysis). See also Sections C2.4 to 2.6.

Specimen processing

Protocols should be designed to avoid contamination, tampering, or substitution. Handling samples must be in accordance with Occupational Safety and Health Administration guidelines, with the express understanding that any human fluids may harbor infectious agents.

Sample stability

PAPP-A and the intact (or total) forms of hCG can be reliably determined in sera stored at 4 to 8°C for at least 6 days⁹ and at -20°C for several months. If frozen samples are to be used to derive medians, possible freeze/thaw effects should be examined. For optimal performance, shipping time should be minimized (e.g., express mail and courier service) and samples should not be exposed to high temperatures. Free beta hCG subunit is spontaneously generated at high temperatures (e.g., daytime summer temperatures in the southern United States), due to dissociation of intact hCG.¹² If free beta is to be measured in serum, samples must be protected from high temperatures (e.g., cool packs with overnight shipment in the summertime). Shipping samples in the form of blood spots can also result in improved stability.¹³ See also Sections C2.4 to 2.6.

Establishment of laboratory policies regarding specimen retention

See Sections C2.7 to 2.8 for more information.

Assay methodologies

Detailed analytic procedures

Guidance on developing assay protocols is available. See Sections C5, C6, and Validation C8.3.

Methodology and reagents

In the United States, neither PAPP-A, free beta-hCG nor intact hCG have been licensed by the Food and Drug Administration (FDA) for specific use in Down syndrome screening. The FDA has, in fact, not licensed PAPP-A or free beta-hCG reagents for any use; they are available as research use only kits, as laboratory developed tests, or as analyte-specific reagents. Only the hCG (intact and total) and DIA kits have Class II approvals (510K), albeit for applications other than Down syndrome screening. Manufacturers are thus restricted from making clinical claims about Down syndrome screening.

Laboratory-developed tests are currently not subject to these FDA restrictions. Some manufacturers provide reagents that can be used for screening as analyte-specific reagents. These require full in-house validation. Published guidelines from the Clinical and Laboratory Standards Institute provide procedures for evaluating manufactured reagents including precision, linearity, stated performance characteristics, and guidelines on clinical sensitivity and specificity.¹⁴ Kits from several manufacturers are available for serum PAPP-A and intact or total hCG assays. Because of patent/licensing agreements, fewer options are available to laboratories for purchasing reagents for measuring serum DIA, h-hCG, and free beta-hCG. This also holds for quantifying PAPP-A, free beta-hCG, and intact hCG in dried blood spots.

Standards and calibration procedures

Most serum hCG kits are calibrated in mIU/mL, using the First International Reference Preparation (equivalent to the

Third International Standard). Results can be reported as IU/mL (e.g., 3.2) or mIU/mL (e.g., 3200). Some free beta hCG kits are also calibrated using the First International Reference Preparation (equivalent to the Third International Standard) and results are reported in mIU/mL or ng/mL. Serum PAPP-A is reported in mIU/mL. The DIA assay is standardized against the First International Standard for Inhibin (Code 91/624) and results can be reported in pg/mL or IU/mL. Some kits may report results in uIU/mL for PAPP-A and mIU/mL for free beta/intact-hCG, but values can be converted to those used with serum samples by correcting for dilution. Commercially available kits provide calibrators and specific calibration protocols. Laboratories using "home brew" or "laboratory-developed tests" assays or modifying kit assay protocols are responsible for determining calibration protocols and validating performance.

Preparation, characterization, and use of controls

Assay controls. Serum: In-house pooled controls, commercially available controls, or controls received in kits serve as checks on reagents and technical performance. Advantages of in-house pooled controls include a sample matrix that more closely resembles patient samples, levels specifically targeted for Down syndrome clinical action points (e.g., lower PAPP-A and higher hCG or free beta-hCG levels), and control lots prepared with long expiration dating to aid in assessment of kit master reagent lot changes and long-term assay drift. An alternative for long-term monitoring is commercial controls bought in sufficient quantity to last a year or more.

Dried blood spots: No currently available commercial controls specifically include inhibin, PAPP-A, or free beta-hCG. No commercial controls are available for the dried blood spot assays. Controls for these assays include those received in kits, which are liquid and only check for reagent and technical performance. In-house dried blood spot controls should be made to check for extraction efficiency.

Repeat assay controls. Repeat assay controls can also be helpful in monitoring performance variability. To assess short-term performance, unfrozen patient samples are chosen at random from recent assays and reassayed to monitor intra- and interassay precision. Because the serum analytes currently in wide use are essentially stable when frozen and thawed (see Section "Sample stability"), reassaying stored patient samples from the time period when the current median values were established can also help to identify long-term drift and determine whether reference data need to be updated.

Concentration of controls. Each assay should contain at least two quality control samples that fall at clinical action points (three controls may be needed to comply with some licensure requirements). For example, an hCG or free beta-hCG high control could be targeted at a value equivalent to 2.0 or 2.5 multiples of the median (MoM) value at 12 weeks, along with a second mid control near the median (1.0 MoM).

Characterization of control materials. After preparation and aliquotting, performance ranges for in-house pooled controls can be set using standard clinical laboratory quality control approaches. Controls received with licensed kits have an acceptable target range specified by the manufacturers, but laboratories may wish to establish an in-house range. This information is used to accept or reject individual control results or a whole assay, so that care should be taken to set appropriate ranges and avoid unnecessary result rejection.

Quality control

Type and frequency of QC assessments. Standard approaches used in the clinical laboratory are appropriate for internal QC of these assays.

Measures of repeatability both within and between runs. As part of the initial method validation, the laboratory should demonstrate that intra- and interassay variation reported by the manufacturer can be reproduced.

Routine equipment calibration and preventive maintenance. Standard approaches used in the clinical laboratory are appropriate. In many cases, calibration and maintenance protocols are set by the product/equipment manufacturer (see Characterization of Control Materials section).

Assay results

Converting assay results to multiples of the median

In order for each of the analytical measurements to be interpreted, each result in mass or international units must first be converted to a MoM for a given gestational age. The resulting MoM levels can then be adjusted for other factors such as maternal weight.

Normative data

It has been established that values obtained from different lots from the same manufacturer or from different manufacturers may demonstrate systematic bias. Therefore, it is essential that each laboratory establish its own normative data or, at a minimum, demonstrate that data obtained from another source are appropriate for its screened population.

Source of medians. Package insert (commercial) medians should not be used, even for a short time. Several methods exist that can be used to establish reliable medians.

Sample size. Ideally, 100 samples for each completed gestational week from 11 through 13 would be used to calculate median values. Because these analytes are stable, it may be possible to use stored frozen specimens. A consecutive series of samples may also be appropriate; it is not necessary that all samples be documented as being drawn from unaffected singleton pregnancies. Using regression analysis allows the use of fewer samples (e.g., 300 spread over the 11–13 completed week period) to establish reasonable medians.¹⁵

Computing medians. “Smoothing” the observed median values by weighted regression analysis provides reliable and accurate medians. Appropriate models for each of the analytes can be found in the literature (all analytes fit a log-linear model between 11 and 13 weeks’ gestation).¹⁶ These methods also allow median values to be extrapolated for weeks in which little data are available. Using median values that are specific to each day of gestation will further improve screening performance and their use is encouraged. For some laboratories, it may be necessary to establish median values earlier than 11 weeks’ gestation (see Gestational Age section). If so, actual observed data are required, as extrapolation of the above models may not provide accurate results.

Expected change in medians by gestation. Maternal serum PAPP-A levels increase by a constant percentage per week between 11 and 13 completed weeks’ gestation (e.g., 40–50%).¹⁶ Before 11 completed weeks’ gestation, PAPP-A levels increase even more each week. Levels of free beta-hCG, hCG, h-hCG, and DIA decrease by about 20% to 40% from 11 to 13 completed weeks.^{16–19}

Variables that have significant impact on calculation of the MoM level

Time of testing. Clinical sensitivity and specificity for Down syndrome screening are best at 11 completed weeks’ and reduced by 13 completed weeks’ gestation. Measurements of both NT and PAPP-A perform less well, whereas free beta-hCG and hCG perform better, as gestational age increases from 11 through 13 completed weeks’ gestation.²⁰

Gestational age. The best screening performance is achieved by expressing gestational age as weeks and days or decimal weeks (12 weeks, 5 days is 12.7 weeks). If a laboratory chooses to report gestational age as whole weeks, it should be expressed in completed weeks (12 weeks, 5 days is 12 completed weeks) rather than in rounded weeks (12 weeks, 5 days is 13 completed weeks).

Dating method to use. In the first trimester, the most common method for determining gestational age is dating by ultrasound measurement of CRL. CRL measurements at 11 to 13 completed weeks’ gestation (or earlier) provide an accurate estimate of gestational age to within 7 days.²¹

Incorporating dating method. In general, all first-trimester screening (and most integrated screening) will be based on gestational age as estimated from the CRL measurement. Measuring the CRL early in pregnancy provides a highly reliable estimate of gestational age. In addition, this measurement has been shown to be an unbiased measurement in both Down syndrome and unaffected pregnancies.²² Programs should make efforts to avoid using gestational age estimates based on ultrasound measurements known to be systematically different in Down syndrome pregnancies (e.g., humeral and femur length). Given that nearly all gestational age estimates will be based on ultrasound measurements, programs will likely have only one set of medians for each marker and use only one set of parameters.

Factors that may be used to adjust the MoM levels

The following are interpretive refinements based on patient demographics and other pregnancy-related information that are less critical than taking gestational age into account, but will improve screening performance by optimizing the interpretation. Currently, many laboratories take the following factors into account.

Maternal weight. For each of the analytes, levels are on average higher in lighter weight women and lower in heavier weight women. Adjusting for maternal weight provides only a slight improvement in first-trimester Down syndrome screening. However, weight adjustments should be performed for other reasons (e.g., Trisomy 18). The relationship between PAPP-A measurements and maternal weight is stronger than for any other marker.³ All first-trimester markers should be adjusted for weight when screening for Down syndrome and for Trisomy 18. Laboratories should use published weight adjustment formulas only until in-house data are collected and new laboratory-specific formulas are derived.²³

Maternal race and maternal insulin-dependent diabetes mellitus. Insulin-dependent diabetes mellitus effects for the first-trimester biochemical markers are not yet well described. Racial/ethnic effects are better defined, with clear increases in PAPP-A levels in African Americans of 25% or more^{24–26} and lesser changes for Asian Americans and Hispanic Americans for PAPP-A, hCG, and free beta-hCG. Such changes should be accounted for in the Down syndrome interpretations. Labora-

tory directors should periodically review the literature on these topics.

Maternal cigarette smoking. Among smokers, levels of both free beta hCG and PAPP-A are significantly reduced by about 10% and 15% to 20%, respectively.²⁷ The effect is not dependent on the numbers of cigarettes smoked. The birth prevalence of Down syndrome does not seem to be influenced by smoking status.²⁸ At this time, laboratories should consider adding a simple yes/no question concerning cigarette smoking status to the requisition slip. Methods to incorporate smoking status responses in a Down syndrome screening program, and the expected results have been published²⁷ and can be used as a guide. Accounting for smoking status when interpreting Down syndrome risks is up to the discretion of the laboratory director.

Assisted reproductive techniques. The levels of the first-trimester markers have been studied in various types of assisted reproductive technologies ART pregnancies, primarily in vitro fertilization, but a consensus has not yet emerged. Laboratory directors have discretion to implement ART adjustment, but should periodically review the literature on this topic.^{29–38}

Use of multiple correction factors to calculate the MoM. In some situations, multiple adjustments to the analytic result (expressed in MoM) might be warranted. Although data are sparse, programs can make the assumption that the effects are independent. For example, although most data for cigarette smoking are derived mainly from the studies of white women, the assumption is usually made that a similar effect will be seen in Black/African American women.

Prenatal screening software for computing and reporting patient-specific risk for Down syndrome

Laboratories must be able to compute risks for Down syndrome. The use of specialized software applications is generally considered a necessity for Down syndrome screening due to the complex nature of calculating and interpreting the results, the need for patient-specific interpretive reports, and the large number of samples processed.

Down syndrome risks. Patient-specific Down syndrome risks are generated by complex computer algorithms that are integral to prenatal screening. Such software applications can be obtained commercially or developed in-house and must be verified before routine clinical use.

Risk algorithm. The commonly used algorithm to assign a patient-specific risk uses the MoM results (adjusted for variables such as weight and race, as discussed earlier) to calculate a likelihood ratio based on the overlapping multivariate Gaussian distributions defined by the affected and unaffected population parameters. The a priori risk for Down syndrome, based on maternal age, is then multiplied by the corresponding likelihood ratio to generate the patient-specific risks. Equations to compute the a priori risk for a given maternal age have been published.³⁹

Population parameters. Risk algorithms use published^{40–43} or in-house population parameters for each of the analytes, expressed as log means and log standard deviations for unaffected pregnancies and pregnancies affected with Down syndrome. In addition, pairwise correlation coefficients in both affected and unaffected pregnancies and truncation limits are needed to generate reliable risks. These population parameters will vary, based on the gestational age in the first trimester and use of week-specific parameters are recommended (see Method of Assigning Gestational Age section).

Combinations of factors. There is no formal consensus on which adjustments to the result or a priori risk to include in an interpretive software program. Specifically how to include them, or whether to include them, is a decision left to the laboratory director.

Selection of screening cut-off levels

Screening for Down syndrome relies on the patient-specific risk as the screening variable. This risk is computed by multiplying the age-associated a priori risk with the likelihood ratio derived from the multiple analyte measurements.

Down syndrome risks. Down syndrome risks can be expressed as risks in the first trimester, second trimester, or at term.⁴⁴ If the term risk for an average 35-year-old woman is reported to be 1:380, the corresponding (and equivalent) first-trimester risk is $1:380 \times 0.57$ or 1:220. The factor of 0.57 (or 57%) represents a reliable estimate of the proportion of Down syndrome fetuses that survive from the late first trimester to term (e.g., 43% of Down syndrome pregnancies are lost between the late first trimester and term).⁴⁴ Lower estimates for this proportion have also been published.⁴⁵ Term Down syndrome risks should be computed for maternal age in decimal years rather than completed years. Published equations can be used to compute these risks.^{39,46,47}

Risk cut-off level. Historically, women have been offered diagnostic testing if they were 35 years of age or older (term risk of about 1:380; first-trimester risk of about 1:220). Screening programs that combine NT measurements with PAPP-A and free beta-hCG or hCG measurements in combination with maternal age (at 11 through 13 completed weeks, combined) may continue to use this cutoff level with the expectation of about a 85% Down syndrome detection rate (DR) and a 5% false-positive rate. It might also be acceptable to choose a higher risk cutoff level (e.g., first-trimester risk of 1:150) that is associated with a lower DR of about 81% and a lower false-positive rate of about 4%.⁴⁰ Detection and associated false-positive rates (FPR) vary not only by screening cutoff level chosen but also by the age distribution of the women tested, the gestational age at sampling, and the combination of markers chosen. Reports should contain a numerical risk rather than just a screen-positive or screen-negative interpretation.

Variables that impact either the prior risk or the screening parameters

Time of testing. The optimal time in the first trimester for Down syndrome screening based on biochemistry alone is 11 completed weeks' gestation or even earlier.⁴⁰ However, professional organizations representing sonographers suggest that NT measurements be done between 11 and 13 completed weeks (see Introduction section). The gestational age window for NT measurements will likely also define the window for biochemical testing, although some suggest collecting a blood sample several weeks before the NT measurement being performed. If testing at 10 weeks' gestation or earlier is to be offered, the number of samples used to compute reliable medians should be increased and observed data collected within each completed week in which an interpretation is to be provided. Before 11 completed weeks' gestation, free beta-hCG is useful for combined screening, but hCG is not.

Multiple gestation. Overall, Down syndrome screening in twin pregnancies is less effective than for singleton pregnancies. The levels of serum markers in twin pregnancies are pregnancy specific. In contrast, NT measurements are fetal specific. Algo-

rithms for assigning a pregnancy-specific risk when combining NT and biochemistry have been published.⁴⁸ Such algorithms can also take into account the zygosity of twins by observing chorionicity. Whether or not, and how, a program will interpret samples from known twin pregnancies is left to the laboratory director.

Repeat testing. Obtaining a second specimen for repeat testing is discouraged as part of Down syndrome screening (one exception is when sample mix-up is suspected).¹⁶ If an initial sample is later found to have been collected too early, the subsequent sample is considered the first interpretable sample. Should a known repeat sample be submitted to the laboratory, it should not be interpreted as though it were an initial specimen. Published methods exist to combine the information from the two samples to provide a reliable interpretation. Other methods of interpretation may be acceptable.

Family history. Family history of Down syndrome increases the a priori risk, depending on the degree of relatedness and the mode of inheritance. Laboratories may choose to include a recommendation for genetic counseling in the report if a family history suggests an a priori risk higher than the patient's age-related risk. Alternatively, laboratories may choose to incorporate a previous affected pregnancy into the Down syndrome risk estimate using published algorithms.⁴⁹

Method of assigning gestational age. First-trimester screening using NT and biochemical measurements are nearly always interpreted based on a CRL estimate of gestational age. The sonographer may report both a CRL (in mm) and an associated gestational age (in weeks + days, or decimal weeks). There is a possibility of confusion arising, however, because not all sonographers use the same reference data to convert CRL to gestational age (discussed in Conversion of CRL to Gestational Age section).

Technical limitations of the methodology for the intended use

Laboratories need to select reagents for maternal serum screening to meet performance requirements that are more stringent than for other intended uses. Kits need to be both precise and relatively accurate (different kits need not give identical values on the same sample provided in-house reference data are established using the same kit). Coefficients of variation of 10% can adversely impact both detection and screen-positive rates. It is also important that kits/reagents are stable over a long period of time, and that lot-to-lot variability is minimized.

Long-term assessment of variability and performance

Assay controls. In-house pooled controls (or commercial products obtained in sufficient quantity to last a year or more) and repeat assay controls are valuable for monitoring long-term assay drift and lot-to-lot variability (see Assay Controls to Characterization of Control Materials section).

Normative data review. Median values should be reviewed at regular intervals by the laboratory and recalculated when necessary, at least annually. Medians should be recalculated if there is a shift in analyte values >10% or a shift between 5% and 10% that is consistent over time, whether due to observed assay drift or reagent lot change. Shifts in analyte values can be monitored by computing the overall median MoM level (see Epidemiological Monitoring section). Observations from samples tested in the previous months or years should be used only if epidemiological monitoring (defined as the process of gath-

ering data from the screened population to allow for calculation of parameters such as the false-positive rate that serve as extended quality control measures) shows that the median MoM has been stable. Alternative methods of revising medians may be necessary if a significant shift has been observed (see Normative Data section). Shifts in medians for NT measurements are discussed in a later section (Establishing Sonographer-Specific, Center-Specific or Single Set of Medians for NT Measurements).

Evaluating medians with new reagent lots. Between 25 and 50 patient samples and current controls can be run on the old and new kit/reagent lot and the relationship between the two examined using techniques of regression analysis and method comparison. That relationship can then be applied to the existing medians to derive temporary new medians that can be used until sufficient data are available from the new lot for the optimum analysis (see Sample Size section).

Epidemiological monitoring. To monitor assay and program performance and to identify possible areas of concern, screening programs must perform epidemiological monitoring. Such monitoring, at a minimum, should include the periodic computation (monthly or weekly depending on numbers of samples processed) of the median MoM for each of the analytes and the determination of the statistical significance of any deviation from 1.00. Any corrective action needs to be documented. Laboratories should periodically compute their initial positive rate and compare it to expected published rates, after taking into account variables such as the number of analytes and screening cutoff level used.

Long-term monitoring

In recent years, stricter privacy and confidentiality policies and in some cases laws have made it much more difficult to collect pregnancy outcome information and even information regarding follow-up of medical procedures (such as ultrasound and amniocentesis) performed subsequent to positive screens. If possible, laboratories should collect pregnancy outcome information on the women with initial screen-positive results. This information might include the proportion of pregnancies reclassified as screen negative, the diagnostic testing uptake rate, and the number of affected pregnancies identified either in the second trimester or at term. For those laboratories that have sufficient resources, complete pregnancy follow-up is recommended and will allow the determination of the Down syndrome DRs. An alternative approach acknowledged by some regulatory agencies is to use epidemiological monitoring data as performance measures.⁵⁰ This can be accomplished by comparing published rates with in-house statistics for such measurements as the median MoM for each analyte, population parameters (log means and log standard deviations), and initial and revised positive rates (see Epidemiological Monitoring section).

Failure rates for different sample types

Few published data exist from screening programs, but kit manufacturers do provide information about acceptable sample types (e.g., serum versus plasma), minimum sample volumes required, and conditions that can affect assay performance (e.g., hemolysis). Because laboratories should have specific sample processing protocols, many identifiable problem samples will be rejected before testing. Other testing "failures," such as results falling below the lower limit of sensitivity of the assay due to a sampling error, are likely to be uncommon and resolvable by repeat testing. In rare cases, a second sample may be requested.

External proficiency testing

Each laboratory must participate in a proficiency testing or interlaboratory comparison program that evaluates assay performance for the serum analytes in the first trimester. The proficiency testing should also include Down syndrome risk estimates (see Section C4).

Analytic validity

The analytic validity of a test defines its ability to accurately and reliably measure a specific analyte that is to be used clinically. Each laboratory is responsible for documenting in-house validation of a test methodology using standard clinical chemistry criteria, which should include determination of inter- and intra-assay precision, establishment of linear range and the lower limit of detection, analytic specificity, and accuracy (e.g., recovery or method comparison). Information in the package insert of an FDA-approved kit or from the literature can be used as supporting evidence.

Confirmatory testing

Samples with results less than the lower limit of sensitivity of the assay must be repeated to rule out a technical error (e.g., sampling probe error) and to confirm the value. Results above the highest standard on the calibration curve must be repeated at dilution. Samples with a high coefficient of variation between replicate values (generally >10%) are routinely retested by most laboratories to confirm the value. Confirmatory testing is particularly important for the majority of laboratories using methodologies that test in singlicate to minimize analytic errors.

Assay robustness measures how testing is resistant to small changes in preanalytic and analytic variables. In an attempt to define performance requirements and minimize possible impact on assay performance (e.g., analytic validity, reproducibility and failure rates), laboratories should consider the effects of common variables such as sample type, sample handling (e.g., transit time or conditions), sample quality, reagent lots, or minor changes in assay conditions (e.g., timing or temperature).

Clinical validity

The clinical validity of a genetic test defines its ability to accurately and reliably identify the clinical phenotype of interest. In this instance, it is the ability of the assigned Down syndrome risk (based on maternal age, ultrasound marker(s), and multiple maternal serum analyte measurements) to identify pregnancies in which the fetus is affected with Down syndrome.

Clinical sensitivity (or the detection rate [DR]) is the proportion of pregnancies with Down syndrome that has a positive test result (risk at or above the specified risk cutoff level).

Clinical specificity (or 1—false-positive rate) is the proportion of unaffected pregnancies that has a negative test result (risk below the specified risk cutoff level).

Screening performance

Clinical sensitivity (DR) and the corresponding clinical specificity (1—false-positive rate) will vary, depending on the combination of serum analytes used, the Down syndrome risk cutoff level chosen, and the distribution of maternal ages in the population being tested. Readers are directed toward published studies that have compared various marker combinations and risk cutoffs. Care should be used in interpreting published results because of differences in the gestational weeks studied and in the parameters used. Table 1 contains performance estimates for selected first-trimester Down syndrome screening protocols, based on published parameters,^{17,20,40,41,51–53} and should

be used only as a guide for initial comparisons. DRs for screens done at 11 and 13 weeks and at a fixed 1% and 5% false-positive rate are provided.

Positive predictive value (PPV) and negative predictive value (NPVs) of testing in the target population measure the ability of the test to give accurate clinical information.

Positive predictive value. The PPV is the proportion of positive test results that correctly identifies pregnancies with Down syndrome [number of true positives/(true positives + false positives)]. The PPV can also be expressed as an odds ratio and is referred to as the odds of being affected given a positive result (OAPR).

Negative predictive value. The NPV is the proportion of negative tests that correctly identifies unaffected pregnancies [number of true negatives/(true negatives + false negatives)]. Because the prevalence of Down syndrome is low, the NPV is generally not computed.

Modifying factors are important to understand

These may be genetic, environmental, or other factors. Several of these have been discussed earlier and include maternal race (see Maternal Race and Maternal Insulin-dependent Diabetes Mellitus section), family history (see Family History section), and twin pregnancies (see Multiple Gestation section).

Result reporting

Recommended report formats

Final reports of test results (see Validation C8.5.7) must be clear to a non-geneticist professional and must include:

- Patient's name, date of birth, and other unique identifiers.
- Name of referring physician/health center to receive the report.
- The test that is ordered.
- Type of specimen.
- Date when sample was obtained.
- Laboratory accession number(s) that uniquely identifies the sample.
- Demographic and pregnancy-related information used in the interpretation (e.g., CRL, maternal age, and maternal weight).
- NT measurement and interpretive units (e.g., NT in mm, NT in MoM).
- Sonographer's name or ID (e.g., credential number).
- Analytic results in both mass units (e.g., ng/mL) and interpretive units (e.g., MoM) on which all adjustments/corrections have been performed.
- Clinical interpretation, including whether the result is screen positive or screen negative, the risk cutoff level, and the patient-specific risk for Down syndrome.

Reporting screen-negative results

Written reports of screen-negative results can be transmitted to the referring physician by mail, courier, electronic transmission, or overnight carrier.

Reporting screen-positive results

Screen-positive results should be promptly transmitted to the referring physician, usually by phone and/or fax, within one working day after completion of the test. Appropriate recommendations for follow-up of screen-positive results may include the following.

Genetic counseling. In a small proportion of the screened population, it may be reasonable to include a recommendation on the laboratory report to refer the woman for genetic counseling. For example, a referral might be reasonable when there is a family history of numerous individuals affected with Down syndrome (which may indicate a familial translocation).

Chorionic villus sampling or amniocentesis and karyotype. Prenatal chromosome analysis is generally available through two invasive testing options: CVS and amniocentesis. CVS is usually offered between 10 weeks 0 days and 12 weeks 6 days of gestation although many physicians perform CVS up to 13 weeks 6 days.⁵⁴ Beginning at 15 weeks, amniocentesis is generally offered as a diagnostic test to determine the chromosome status of the fetus.

Reclassification of positive results

Laboratories should be aware of the potential problems associated with reclassifying screen-positive women as screen negative. There is a chance of reclassifying a true positive (the fetus has Down syndrome) as a false negative. Reclassification is primarily a concern for second-trimester screening and usually occurs when a last menstrual period (LMP)-dated pregnancy is subsequently dated by ultrasound and the difference between the LMP and the ultrasound dating exceeds a set standard. The need to reclassify first-trimester screening results is expected to be infrequent when NT and CRL measurements are included.

Other conditions associated with a high Down syndrome risk estimate

- Delivery of an unaffected infant is the most common outcome associated with high Down syndrome risk.
- Other chromosome abnormalities (e.g., Turner syndrome, and triploidy).
- Other fetal/placental/maternal abnormalities or conditions (these may occasionally be identified, but the strengths of these associations are low).

Other conditions that can be identified in conjunction with a Down syndrome screening program

Trisomy 18 (Edwards syndrome). Trisomy 18 is a serious chromosome abnormality that is nearly always fatal in the first few days or months after birth. In the absence of prenatal diagnosis and termination, the age-specific birth prevalence of Trisomy 18 is about 10 times lower than for Down syndrome. An estimated 70% of Trisomy 18 fetuses alive in the second trimester will be spontaneously lost by term. The high rate of cesarean section in women with an undiagnosed Trisomy 18 pregnancy may result in unnecessary maternal morbidity. Although it would be difficult to justify a screening program aimed solely at identifying Trisomy 18, such opportunistic testing is a common addition to second-trimester serum maternal serum screening programs for Down syndrome. Because the pattern of markers is different in Trisomy 18 compared with Down syndrome, a separate algorithm is required. Algorithms are available that estimate an individual pregnancy's risk of Trisomy 18.^{55,56} Data regarding detection and FPR associated with Trisomy 18 screening in the first trimester are limited. Published data suggest that laboratories may be able to detect at least 75% to 80% of Trisomy 18 fetuses using maternal age in combination with NT, PAPP-A, and hCG measurements at a false-positive rate of <1%.⁵⁶ Because of the high fetal loss rate

for Trisomy 18 pregnancies from the late first trimester to term, the true DR is difficult to reliably quantify.

Trisomy 13 (Patau syndrome). Trisomy 13 is a serious chromosome abnormality with a lower birth prevalence than Trisomy 18 (about 1 per 10,000 or lower). The existing data (limited to intervention trials subject to ascertainment bias) indicate that the biochemical and ultrasound measurements in Trisomy 13 pregnancies can be useful in identifying a high-risk group that could be offered diagnostic testing.⁵⁷ Programs can consider reporting risks for Trisomy 13 but should be mindful that these risks may not be as reliable as those reported for Down syndrome and the associated clinical utility is less certain. The risk cutoff chosen should be associated with a low positive rate and a high PPV.

Clinical utility

Clinical utility addresses the risks and benefits associated with testing in the routine clinical practice. This information may be requested by those ordering or paying for testing, and the laboratory should be able to provide a reasonably accurate summary of the published literature. When clear gaps in knowledge exist, the laboratory may want to collect data in such a way as to address these questions. The following is a list of selected clinical utility topics that often are applicable:

- Knowing whether pilot trials have been undertaken and, if so, what the results were.
- Establishing or adopting quality assurance processes that monitor the effectiveness of the laboratory's ongoing testing activities.
- Understanding possible adverse health or psychosocial consequences of testing.
- Describing what follow-up testing or interventions might be reasonable in persons with positive test results.
- Understanding what is known about the financial costs and economic benefits of testing.

Ethical, legal, and social implications

The laboratory should be familiar with the ethical, legal, and social issues regarding genetic testing in general, and those specifically applicable to maternal serum screening for Down syndrome. These may include informed consent, insurability, discrimination, labeling, confidentiality, and obligations to disclose. Legal issues such as patents, licensing, sample ownership and storage, proprietary testing, and reporting requirements should be carefully examined.

ADDITIONAL DOWN SYNDROME SCREENING PROTOCOLS USING FIRST-TRIMESTER DOWN SYNDROME MARKERS

Introduction

NT is defined as "the subcutaneous collection of fluid in the fetal nuchal region," and its sonographic measurement in the late first trimester is a marker of risk of fetal Down syndrome. Bronshtein et al.⁵⁸ and Szabo and Gellen⁵⁹ first reported an association between increased NT in the first trimester and abnormal karyotype. By 1992, studies had demonstrated that NT measurement could serve as a useful screening marker of fetal Down syndrome.⁶⁰

The quality of NT as a screening marker depends on sonographer training, ongoing quality review of the sonographic image, and analysis and quality assurance of the quantitative measurements. It is beyond the scope of these guidelines to

describe the method of NT image capture and the review of image quality, but the issues involved in analysis and quality assurance of the quantitative aspects of NT as a marker will be discussed.

Coordinating information with, and collecting data from, sonographers

Proof of training, credentialing, or certification

Laboratories should make an effort to obtain, and have the right to require, documentation of specialized training and successful submission of NT images (and associated data) that qualifies participating sonographers and their supervising physicians specifically for this purpose. Such documentation could consist of the sonographer/physician providing such information and/or by searching databases listing qualified sonographers (e.g., NT Quality Review [NTQR] Program, Fetal Medicine Foundation [FMF]). If a sonographer cannot document his/her qualifications, the laboratory can choose not to accept the prenatal sample for testing and interpretation.

Providing NT and crown-rump length data to the laboratory

At a minimum, the sonographer must report the NT and CRL measurement in mm, and the number of fetuses. In addition, the laboratory should have methods in place to identify each sonographer providing NT measurements. This may be a laboratory-specific code or the code assigned by an NT training program (e.g., NTQR). Sonographer initials should not be used, because they are likely not unique. Additional data that may be useful include a center code, a supervisor code, and the estimated gestational age.

Conversion of CRL to gestational age

There is no universally accepted equation to convert CRL (in mm) to gestational age (in weeks + days, or decimal weeks). This can cause problems when determining whether an NT measurement was taken within the acceptable gestational age window. Most groups suggest obtaining an NT measurement between 11 + 0 and 13 + 6 (weeks + days). This range, when converted to CRL, is dependent on the equations used. Given that the CRL, and not the gestational age, is being measured, it is appropriate to use the CRL to define the screening window. Equations commonly used by sonographers include one published in 1992 by Hadlock et al.²¹ This yields CRL limits of 39 to 84 mm. Other conversions are also available.^{61,62}

The CRL or NT measurement is missing

If the CRL is missing, gestational age cannot be reliably estimated. As a result, it is not possible to confidently interpret either the maternal serum or NT measurements. Thus, the laboratory may consider the sample to be inadequate. The NT measurement might be initially missing, but provided later (with an accompanying CRL). Thus, laboratories should have the capability of interpreting an NT measurement that arrives at a different time in gestation from the biochemistry sample, so long as that original sample and the NT measurement were collected within the specified gestational age window.

Establishing sonographer-specific, center-specific, or single set of medians for NT measurements

Most proponents of NT measurements (NTQR; FMF) argue that only a single "master" set of medians is necessary, if sonographers are properly trained.⁶³ That is, all sonographers can be trained to measure NT in exactly the same way. This is

the rationale for programs of the NTQR and the FMF. Although standardization of technique among sonographers is important, it may not be necessary (or possible) to achieve the uniform performance level required to allow use of only one universal set of medians. Indeed, at least two research trials have analyzed the NT data by both center-specific and sonographer-specific medians. In both, the use of a single master set of medians was associated with lower screening performance than with the use of sonographer-specific medians.^{40,41} If laboratories find that important differences in NT measurements exist between sonographers, it is reasonable for them to use center-specific or sonographer-specific medians (see Section Repeat assay controls). Laboratory software should at least be able to use center-specific medians. The use of a delta NT (difference between expected and observed NT measurements) has been suggested as an alternative to NT MoM (the ratio of expected to observed NT measurements).⁶⁴

Decision to incorporate clinical NT results

Optimally, sonographers should provide paired NT/CRL measurements from 30 (or more) pregnancies to the laboratory to validate appropriate reference data before clinical interpretations. This number was chosen to ensure that the 95% CI of the median NT MoM is between 0.9 and 1.1 MoM. If this is not possible, the laboratory can use an existing set of medians that is likely to be representative of that sonographer (e.g., those recommended by the training program that credentialed the sonographer, those in use at the center where the sonographer practices). As a last resort, a published "generic" set of medians can be used. For all sonographers new to the laboratory, epidemiological monitoring should be performed as soon as adequate data are available.

Quantitative monitoring of NT performance

Performance guidelines for QA/epidemiological monitoring methods

Laboratories can monitor individual sonographer performance with three epidemiological parameters: (1) percent increase of NT measurements per week, (2) median NT MoM assigned, and (3) logarithmic standard deviation of NT MoM levels. This methodology has been used successfully in the First and Second Trimester of Risk (FASTER) trial.⁴¹ Target ranges for the three parameters are (1) 15% to 35%, (2) 0.90 to 1.10 MoM, and (3) standard deviations of 0.08 to 0.13.⁶⁵ Neither the percent per week increase nor logarithmic standard deviation are under the direct control of the screening laboratory. However, the laboratory can adjust the median values when the median MoM for a given sonographer is outside of the expected range. Laboratory directors should monitor the literature to update acceptable ranges for these epidemiological parameters.

Frequency of QA/epidemiological monitoring

The frequency of monitoring depends on numbers of NT measurements reported. If monitoring is performed with too few samples, large variations from the expected might be due to random chance. Alternatively, if monitoring is performed too infrequently, interpretations may not be of reasonable quality. Optimally, at least 50 NT observations should be evaluated quarterly (200 per year). If too few samples are available for quarterly monitoring, then biannually is an acceptable alternative (100 per year). If fewer than 50 samples are available over a 1-year time period, it may not be possible to assess the quality of NT measurements from that sonographer. Written laboratory guidelines need to be in place for dealing with sonographers

providing small numbers of observations. It may be possible to obtain data from sonographers who have sent NT measurements to other laboratories by the laboratory's participation in the NTQR program (www.ntqr.org). In all instances, sonographers should receive routine feedback from the laboratory regarding number of women tested and the three epidemiological parameters.

Addressing significant departures from expected NT performance

If the epidemiological monitoring of a sonographer's NT measurements identifies a statistically significant departure from the expected values (for the increase per week and the logarithmic standard deviation), the laboratory should carefully review the data collection and analyses. If these departures persist over time, the sonographer should be contacted by the laboratory director and the data reviewed for any obvious explanation (e.g., a referral sonographer who sees a higher proportion of elevated NT measurements, inferior ultrasound equipment, multiple sonographers reporting under a single code). The laboratory can then also consider contacting those organizations overseeing the training of sonographers in the United States (e.g., www.ntqr.org, fetalmedicine.com/usa). If no explanation is found and the departure from the expected cannot be rectified, the laboratory could choose not to accept the prenatal samples for testing and interpretation from that sonographer. Laboratories should have a written plan to deal with sonographers whose epidemiological monitoring results fall consistently outside the specified target range.

Decision rules for changing NT medians (center- or operator-specific)

If center- or sonographer-specific medians are used, they should be updated each year, or when monitoring indicates that the median MoM for that sonographer is consistently outside the target range (0.90–1.10 MoM).

Who should handle NT QA training and monitoring?

NT training

Screening laboratories will not usually have the resources, credibility, or expertise to be involved in educating, training, and credentialing sonographers (see Proof of Training, Credentialing, or Certification section). These activities should be handled by groups with special expertise in prenatal ultrasound (e.g., www.ntqr.org, www.fetalmedicine.com/usa).

NT monitoring

Although laboratories should monitor the NT data they receive from each sonographer (and recalculate medians when appropriate), identifying the causes for the deviations from the expected ranges for epidemiological monitoring and any subsequent remedial actions should be handled by professional/credentialing/training organizations.

Nasal bone

The identification of an absent nasal bone between 11 and 13 completed weeks' gestation is reported to be a useful Down syndrome marker.⁶⁶ However, reliable nasal bone evaluation and measurement require additional training and demonstration of proficiency. Inclusion of nasal bone measurement as a standard part of a routine Down syndrome screening in the United States will require further validation studies.^{67–69}

Combining first-trimester and second-trimester markers

Integrated screening

Integrated screening refers to a two-stage process² that “integrates” the most informative Down syndrome markers (from both the first and second trimesters) together into a single risk assessment. Integrated screening can be done using first- and second-trimester serum markers only (the serum integrated test) or serum markers in combination with first-trimester NT measurement (the full integrated test). Integrated testing does not include free beta-hCG or hCG measurement in the first trimester, because either free beta-hCG or hCG are a better marker in the second than the first trimester. Because screening results are reported in the early second trimester, the follow-up diagnostic procedure for screen-positive women is usually amniocentesis. Laboratories offering integrated screening should also be prepared to have the first-trimester sample and NT measurement collected earlier in gestation and establish protocols for dealing with the absence of a second-trimester sample, or its late arrival. Offering integrated screening may require the laboratory to obtain a license, as it is covered by international patent. Sequential or contingent protocols may, or may not, be covered.

Serum integrated screening (NT measurements not available)

For the serum integrated test, PAPP-A is measured in a blood sample obtained in the first trimester, but the result is not reported. A second-trimester blood sample is then obtained from the same woman, and additional markers (e.g., AFP, uE3, free beta-hCG or hCG, and DIA) are measured. First-trimester PAPP-A measurements plus the four second-trimester analytes are then combined with maternal age to produce a single patient-specific risk in the second trimester. At a fixed 5% FPR, the DR is about 89% (if the first-trimester sample is obtained at 10 weeks' gestation).⁴⁰ This is comparable with the screening performance of the first-trimester “Combined Test” (Table 1), which requires NT measurements. The serum integrated test can be offered to women who present for prenatal care in the first trimester, but who, for whatever reason, cannot obtain reliable NT measurements.

The laboratory should decide the acceptable gestational age range for which it will accept samples for serum integrated screening. Because NT measurements are not involved, the lower gestational age limit for the first-trimester sample collection can be earlier than 11 weeks' gestation, a time when PAPP-A measurements are more discriminatory.⁷⁰ However, laboratories would need to obtain appropriate Down syndrome risk equations and develop reliable median levels before offering the test clinically. Another issue is whether to accept samples that are dated by LMP. The gestational dating method of choice for the serum integrated test is an ultrasound measurement of either first-trimester CRL or, less desirably, a second-trimester BPD. Because PAPP-A levels increase rapidly during the late first trimester, the use of last menstrual period dating may result in an inappropriately high screen-positive rate and is not recommended.

When a serum integrated test has been ordered but the second-trimester serum sample is not received promptly, the laboratory should contact the referring physician for further information to resolve the issue. If the expected sample is still not forthcoming, the laboratory can choose either to issue a report stating that the test result cannot be calculated or to retrieve the first-trimester serum sample from storage and perform further testing. A first-trimester test based only on mater-

nal age, PAPP-A, and hCG (or free beta-hCG) measurements has an estimated DR of about 60% at a 5% false-positive rate.⁴⁰

Full integrated test (NT measurements available). For the full integrated test, an NT measurement collected within the appropriate gestational age range is included, along with the five serum markers. Together with maternal age, these are used to calculate a single patient-specific risk. At a fixed 5% FPR, the estimated DR is about 92%, with the first-trimester data collected at 12 week's gestation. At a fixed 1% FPR, the corresponding DR is 83%.⁴⁰ Depending on the maternal age distribution of the population being screened, a 1 in 100 second-trimester risk cutoff will provide a DR of 85% to 90%, with a 1% to 2% false-positive rate.⁴⁰

Sequential screening (NT measurements available)

Sequential (or step-wise) screening incorporates aspects of first-trimester and second-trimester screening in a two-step strategy, in an effort to preserve the benefits of each type of screening (early diagnosis of Down syndrome pregnancies and highest screening performance, respectively). Measurement of first-trimester NT is required for sequential screening. In addition, CVS should be available if women are offered sequential testing. To maintain an acceptable balance between detection and false positives, only women with very high Down syndrome risks in the first trimester (e.g., >1:25 or 1:50) are offered diagnostic testing in the first trimester. More than half of the Down syndrome pregnancies will be found in this group, at a corresponding false-positive rate of 0.5%, or lower.⁶ A second-trimester blood sample is then obtained from the remaining women (99.5%) for measurement of the four additional markers. The risk is calculated using information from both trimesters. Variations of the sequential test exist. Free beta-hCG or hCG is often measured in the first-trimester sample as a way to improve the first-trimester performance. The interim first-trimester risk may, or may not, be reported to all women. The laboratory should address the risks and benefits of these and other important decisions before routine implementation.

Contingent screening (NT measurements available)

Contingent screening is similar to sequential screening. However, in contingent screening, the first-trimester results are divided into three outcomes: screen-positive, screen-negative, and intermediate/pending risk. Those patients with intermediate risks will then provide a second-trimester sample for testing to compute an integrated risk. This strategy allows for early diagnosis of Down syndrome among the small high-risk group (screen positives) while concurrently offering early reassurance to the large low-risk group (screen negatives). Contingent screening attempts to maintain high performance by having those with intermediate/pending first-trimester risks benefit from the integrated test. Contingent screening is more complex for the laboratory to implement as it provides multiple risk estimates to many women, and leaves open the question of second-trimester screening for open neural tube defects. Because of these issues, formal pilot trials might be warranted before routine implementation.

EXISTING GUIDELINES

Two recent clinical guidelines have been published. These differ in focus from these laboratory technical standards and guidelines, but their review by laboratory directors may prove beneficial.

American College of Obstetricians and Gynecologists

ACOG has prepared Clinical Management Guidelines for obstetrician-gynecologists for screening for fetal chromosomal abnormalities.⁶⁷ These policy guidelines include recommendations for screening in the first and/or second trimesters, ultrasound markers and dating, as well as analyte combinations.

Canadian Society of Obstetricians and Gynaecologists of Canada Clinical Practice Guideline for Prenatal Screening for Fetal Aneuploidy.⁷¹ These guidelines include recommendations for screening in the first and/or second trimesters, ultrasound markers and dating, as well as analyte combinations.

REFERENCES

1. Palomaki GE, Bradley LA, McDowell GA; Down Syndrome Working Group; ACMG Laboratory Quality Assurance Committee. Technical standards and guidelines: prenatal screening for Down syndrome. *Genet Med* 2005;7:344–354.
2. Wald NJ, Watt HC, Hackshaw AK. Integrated screening for Down's syndrome on the basis of tests performed during the first and second trimesters. *N Engl J Med* 1999;341:461–467.
3. Knight GJ, Palomaki GE, Neveux LM, et al. Integrated serum screening for Down syndrome in primary obstetric practice. *Prenat Diagn* 2005;25:1162–1167.
4. Wright D, Bradbury I, Benn P, Cuckle H, Ritchie K. Contingent screening for Down syndrome is an efficient alternative to non-disclosure sequential screening. *Prenat Diagn* 2004;24:762–766.
5. Wald NJ, Rudnicka AR, Bestwick JP. Sequential and contingent prenatal screening for Down syndrome. *Prenat Diagn* 2006;26:769–777.
6. Palomaki GE, Steinort K, Knight GJ, Haddow JE. Comparing three screening strategies for combining first- and second-trimester Down syndrome markers. *Obstet Gynecol* 2006;107:367–375.
7. Palomaki GE, Neveux LM, Knight GJ, Haddow JE, Lee J. Estimating first trimester combined screening performance for Down syndrome in dried blood spots versus fresh sera. *Genet Med* 2007;9:458–463.
8. Krantz DA, Hallahan TW, Orlandi F, et al. First-trimester Down syndrome screening using dried blood biochemistry and nuchal translucency. *Obstet Gynecol* 2000;96:207–213.
9. Lambert-Messerlian GM, Eklund EE, Malone FD, Palomaki GE, Canick JA, D'Alton ME. Stability of first- and second-trimester serum markers after storage and shipment. *Prenat Diagn* 2006;26:17–21.
10. Macri JN, Anderson RW, Krantz DA, Larsen JW, Buchanan PD. Prenatal maternal dried blood screening with alpha-fetoprotein and free beta-human chorionic gonadotropin for open neural tube defect and Down syndrome. *Am J Obstet Gynecol* 1996;174:566–572.
11. Verloes A, Schoos R, Herens C, Vintens A, Koulischer L. A prenatal trisomy 21 screening program using alpha-fetoprotein, human chorionic gonadotropin, and free estriol assays on maternal dried blood. *Am J Obstet Gynecol* 1995;172:167–174.
12. Beaman JM, Akhtar N, Goldie DJ. Down's syndrome screening using free beta hCG: instability can significantly increase the Down's risk estimate. *Ann Clin Biochem* 1996;33:525–529.
13. Spencer K, Macri JN, Carpenter P, Anderson R, Krantz DA. Stability of intact chorionic gonadotropin (hCG) in serum, liquid whole blood, and dried whole-blood filter-paper spots: impact on screening for Down syndrome by measurement of free beta-hCG subunit. *Clin Chem* 1993;39:1064–1068.
14. NCCLS SC1-L. Evaluation protocols. Wayne, Pennsylvania: NCCLS Specialty Collection, 2003.
15. Haddow JE, Palomaki GE, Knight GJ, Canick JA. Measuring maternal serum markers for Down syndrome in the second trimester. In: Prenatal screening for major fetal disorders Vol. II: Down syndrome. Scarborough: Foundation for Blood Research, 1998:19–24 through 19–26.
16. Haddow JE, Palomaki GE, Knight GJ, Williams J, Miller WA, Johnson A. Screening of maternal serum for fetal Down's syndrome in the first trimester. *N Engl J Med* 1998;338:955–961.
17. Canick JA, Lambert-Messerlian GM, Palomaki GE, et al. Comparison of serum markers in first-trimester down syndrome screening. *Obstet Gynecol* 2006;108:1192–1199.
18. Palomaki GE, Knight GJ, Neveux LM, Pandian R, Haddow JE. Maternal serum invasive trophoblast antigen and first-trimester Down syndrome screening. *Clin Chem* 2005;51:1499–1504.
19. Weinans MJ, Butler SA, Mantingh A, Cole LA. Urinary hyperglycosylated hCG in first trimester screening for chromosomal abnormalities. *Prenat Diagn* 2000;20:976–978.
20. Palomaki GE, Lambert-Messerlian GM, Canick JA. A summary analysis of Down syndrome markers in the late first trimester. *Adv Clin Chem* 2007;43:177–210.

21. Hadlock FP, Shah YP, Kanon DJ, Lindsey JV. Fetal crown-rump length: reevaluation of relation to menstrual age (5–18 weeks) with high-resolution real-time US. *Radiology* 1992;182:501–505.
22. Wald NJ, Smith D, Kennard A, et al. Biparietal diameter and crown-rump length in fetuses with Down's syndrome: implications for antenatal serum screening for Down's syndrome. *Br J Obstet Gynaecol* 1993;100:430–435.
23. Neveux LM, Palomaki GE, Larrivee DA, Knight GJ, Haddow JE. Refinements in managing maternal weight adjustment for interpreting prenatal screening results. *Prenat Diagn* 1996;16:1115–1119.
24. Wald NJ, Kennard A, Hackshaw A, McGuire A. Antenatal screening for Down's syndrome. *J Med Screen* 1997;4:181–246.
25. Krantz DA, Hallahan TW, Macri VJ, Macri JN. Maternal weight and ethnic adjustment within a first-trimester Down syndrome and trisomy 18 screening program. *Prenat Diagn* 2005;25:635–640.
26. Spencer K, Heath V, El-Sheikhah A, Ong CY, Nicolaides KH. Ethnicity and the need for correction of biochemical and ultrasound markers of chromosomal anomalies in the first trimester: a study of Oriental, Asian and Afro-Caribbean populations. *Prenat Diagn* 2005;25:365–369.
27. Spencer K, Bindra R, Cacho AM, Nicolaides KH. The impact of correcting for smoking status when screening for chromosomal anomalies using maternal serum biochemistry and fetal nuchal translucency thickness in the first trimester of pregnancy. *Prenat Diagn* 2004;24:169–173.
28. Rudnicka AR, Wald NJ, Huttly W, Hackshaw AK. Influence of maternal smoking on the birth prevalence of Down syndrome and on second trimester screening performance. *Prenat Diagn* 2002;22:893–897.
29. Bellver J, Lara C, Soares SR, et al. First trimester biochemical screening for Down's syndrome in singleton pregnancies conceived by assisted reproduction. *Hum Reprod* 2005;20:2623–2627.
30. Bersinger NA, Wunder D, Vanderlick F, et al. Maternal serum levels of placental proteins after in vitro fertilisation and their implications for prenatal screening. *Prenat Diagn* 2004;24:471–477.
31. Frishman G, Blazar A. The history of assisted reproductive technology and its future. *Med Health R I* 1997;80:393–396.
32. Ghisoni L, Ferrazzi E, Castagna C, Levi Setti PE, Masini AC, Pigni A. Prenatal diagnosis after ART success: the role of early combined screening tests in counselling pregnant patients. *Placenta* 2003;24(suppl B):S99–S103.
33. Hui PW, Lam YH, Tang MH, NG EH, Yeung WS, Ho PC. Maternal serum pregnancy-associated plasma protein-A and free beta-human chorionic gonadotropin in pregnancies conceived with fresh and frozen-thawed embryos from in vitro fertilization and intracytoplasmic sperm injection. *Prenat Diagn* 2005;25:390–393.
34. Hui PW, Tang MH, Lam YH, Yeung WS, Ng EH, Ho PC. Nuchal translucency in pregnancies conceived after assisted reproduction technology. *Ultrasound Obstet Gynecol* 2005;25:234–238.
35. Maymon R, Shulman A. Integrated first- and second-trimester Down syndrome screening test among unaffected IVF pregnancies. *Prenat Diagn* 2004;24:125–129.
36. Orlandi F, Rossi C, Allegra A, et al. First trimester screening with free beta-hCG, PAPP-A and nuchal translucency in pregnancies conceived with assisted reproduction. *Prenat Diagn* 2002;22:718–721.
37. Wald NJ, White N, Morris JK, Huttly WJ, Canick JA. Serum markers for Down's syndrome in women who have had in vitro fertilisation: implications for antenatal screening. *Br J Obstet Gynaecol* 1999;106:1304–1306.
38. Wojdemann KR, Larsen SO, Shalmi A, Sundberg K, Christiansen M, Tabor A. First trimester screening for Down syndrome and assisted reproduction: no basis for concern. *Prenat Diagn* 2001;21:563–565.
39. Hecht CA, Hook EB. Rates of Down syndrome at livebirth by one-year maternal age intervals in studies with apparent close to complete ascertainment in populations of European origin: a proposed revised rate schedule for use in genetic and prenatal screening. *Am J Med Genet* 1996;62:376–385.
40. Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM. First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). *J Med Screen* 2003;10:56–104.
41. Malone FD, Canick JA, Ball RH, et al. First-trimester or second-trimester screening, or both, for Down's syndrome. *N Engl J Med* 2005;353:2001–2011.
42. Knight GJ, Palomaki GE, Neveux LM, Fodor KK, Haddow JE. hCG and the free beta-subunit as screening tests for Down syndrome. *Prenat Diagn* 1998;18:235–245.
43. Haddow JE, 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;5:115–119.
44. Morris JK, Wald NJ, Watt HC. Fetal loss in Down syndrome pregnancies. *Prenat Diagn* 1999;19:142–145.
45. Snijders R. Fetal loss in Down syndrome pregnancies. *Prenat Diagn* 1999;19:1180.
46. Morris JK, Mutton DE, Alberman E. Revised estimates of the maternal age specific live birth prevalence of Down's syndrome. *J Med Screen* 2002;9:2–6.
47. Cuckle HS, Wald NJ, Thompson SG. Estimating a woman's risk of having a pregnancy associated with Down's syndrome using her age and serum alpha-fetoprotein level. *Br J Obstet Gynaecol* 1987;94:387–402.
48. Wald NJ, Rish S, Hackshaw AK. Combining nuchal translucency and serum markers in prenatal screening for Down syndrome in twin pregnancies. *Prenat Diagn* 2003;23:588–592.
49. Arbuzova S, Cuckle H, Mueller R, Sehmi I. Familial Down syndrome: evidence supporting cytoplasmic inheritance. *Clin Genet* 2001;60:456–462.
50. Knight GJ, Palomaki GE. Epidemiologic monitoring of prenatal screening for neural tube defects and Down syndrome. *Clin Lab Med* 2003;23:531–551, xi.
51. Wald NJ, Rodeck C, Hackshaw AK, Rudnicka A. SURUSS in perspective. *BJOG* 2004;111:521–531.
52. Evans MI, Krantz DA, Hallahan TW, Galen RS. Meta-analysis of first trimester Down syndrome screening studies: free beta-human chorionic gonadotropin significantly outperforms intact human chorionic gonadotropin in a multimarker protocol. *Am J Obstet Gynecol* 2007;196:198–205.
53. Cuckle H, Benn P, Wright D. Down syndrome screening in the first and/or second trimester: model predicted performance using meta-analysis parameters. *Semin Perinatol* 2005;29:252–257.
54. ACOG Committee on Practice Bulletin. ACOG practice bulletin number 160: screening for fetal chromosomal abnormalities. *Obstet Gynecol* 1995;109:217–227.
55. Palomaki GE, Knight GJ, Neveux L, Pandian R, Haddow J. Maternal serum invasive trophoblast antigen (ITA) in first trimester trisomy 18 pregnancies (poster). Paper presented at the annual meeting of The American Society of Human Genetics, Toronto, Canada, October 29, 2004.
56. Tul N, Spencer K, Noble P, Chan C, Nicolaides K. Screening for trisomy 18 by fetal nuchal translucency and maternal serum free beta-hCG and PAPP-A at 10–14 weeks of gestation. *Prenat Diagn* 1999;19:1035–1042.
57. Spencer K, Nicolaides KH. A first trimester trisomy 13/trisomy 18 risk algorithm combining fetal nuchal translucency thickness, maternal serum free beta-hCG and PAPP-A. *Prenat Diagn* 2002;22:877–879.
58. Bronshtein M, Rottem S, Yoffe N, Blumenfeld Z. First-trimester and early second-trimester diagnosis of nuchal cystic hygroma by transvaginal sonography: diverse prognosis of the septated from the nonseptated lesion. *Am J Obstet Gynecol* 1989;161:78–82.
59. Szabo J, Gellen J. Nuchal fluid accumulation in trisomy-21 detected by vaginosonography in first trimester. *Lancet* 1990;336:1133.
60. 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;304:867–869.
61. Daya S. Accuracy of gestational age estimation by means of fetal crown-rump length measurement. *Am J Obstet Gynecol* 1993;168:903–908.
62. Robinson HP, Fleming JE. A critical evaluation of sonar "crown-rump length" measurements. *Br J Obstet Gynaecol* 1975;82:702–710.
63. Nicolaides K. The 11–13+6 weeks scan. London: Fetal Medicine Foundation, 2004.
64. Spencer K, Bindra R, Nix AB, Heath V, Nicolaides KH. Delta-NT or NT MoM: which is the most appropriate method for calculating accurate patient-specific risks for trisomy 21 in the first trimester? *Ultrasound Obstet Gynecol* 2003;22:142–148.
65. Palomaki GE, Neveux LM, Donnemfeld A, et al. Quality assessment of routine nuchal translucency measurements: a North American laboratory perspective. *Genet Med* 2008;10:131–138.
66. Cicero S, Curcio P, Papageorghiou A, Sonek J, Nicolaides K. Absence of nasal bone in fetuses with trisomy 21 at 11–14 weeks of gestation: an observational study. *Lancet* 2001;358:1665–1667.
67. ACOG Committee on Practice Bulletin. ACOG practice bulletin number 77: screening for fetal chromosomal abnormalities. *Obstet Gynecol* 2007;109:217–227.
68. Malone FD, Ball RH, Nyberg DA, et al. First-trimester nasal bone evaluation for aneuploidy in the general population. *Obstet Gynecol* 2004;104:1222–1228.
69. Prefumo F, Sairam S, Bhide A, Thilaganathan B. First-trimester nuchal translucency, nasal bones, and trisomy 21 in selected and unselected populations. *Am J Obstet Gynecol* 2006;194:828–833.
70. Spencer K, Crossley JA, Aitken DA, Nix AB, Dunstan FD, Williams K. The effect of temporal variation in biochemical markers of trisomy 21 across the first and second trimesters of pregnancy on the estimation of individual patient-specific risks and detection rates for Down's syndrome. *Ann Clin Biochem* 2003;40:219–231.
71. Summers AM, Langlois S, Wyatt P, Wilson RD; Society of Obstetricians and Gynaecologists of Canada. Prenatal screening for fetal aneuploidy. *J Obstet Gynaecol Can* 2007;29:146–179.