Laboratory screening and diagnosis of open neural tube defects, 2019 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG)

Glenn E. Palomaki, PhD[1], Caleb Bupp, MD[2,3], Anthony R. Gregg, MD[4], Mary E. Norton, MD[5], Devin Oglesbee, PhD[6] and Robert. G. Best, PhD[7] on behalf of the ACMG Biochemical Genetics Subcommittee of the Laboratory Quality Assurance Committee

Disclaimer This technical standard is designed primarily as an educational resource for clinical laboratory geneticists to help them provide quality clinical laboratory genetic services. Adherence to this standard is voluntary and does not necessarily assure a successful medical outcome. This standard 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 geneticist should apply his or her own professional judgment to the specific circumstances presented by the patient or specimen.

Clinical laboratory geneticists are encouraged to document in the patient's record the rationale for the use of a particular procedure or test, whether or not it is in conformance with this standard. They also are advised to take notice of the date any particular standard was adopted, and to consider other relevant medical and scientific information that becomes available after that date. It also would be prudent to consider whether intellectual property interests may restrict the performance of certain tests and other procedures.

Open neural tube defects (ONTDs) include open spina bifida (OSB) and anencephaly. These defects are caused by incomplete closure of the neural tube at about 4 weeks of pregnancy. Levels of early second-trimester maternal serum (ms) alpha-fetoprotein (AFP) are sufficiently elevated in affected pregnancies to be used as a population-based screening test. The basic screening methodology was described in the late 1970s and screening programs were active a few years later. By identifying pregnancies with the highest msAFP levels, about 80% of OSB and 95% of anencephaly can be identified as early as 16 weeks gestation. The interpretation of msAFP levels is complicated by the need to consider multiple factors such as gestational age, maternal weight, maternal race, multiple gestations, and more. Testing for AFP and acetylcholinesterase in amniotic fluid and/or identification of the lesion by targeted ultrasound is considered diagnostic of ONTD. When a diagnosis is made, options include termination, surgery after delivery, or in utero surgery, depending on factors such as location and size of the defect, and the presence of any additional anomalies. Screening for ONTD should be performed as part of a comprehensive program linking primary obstetrical care providers, laboratorians, and high-risk clinicians.

OVERVIEW

This Guideline update is intended to augment the current general American College of Medical Genetics and Genomics (ACMG) Standards and Guidelines[1] and to address validation guidelines specific to second-trimester maternal serum alpha-fetoprotein (msAFP) screening for open neural tube defects (ONTD). Individual laboratories are responsible for meeting the Clinical Laboratory Improvement Act/College of American Pathology (CLIA/CAP) quality assurance standards with respect to appropriate sample documentation, assay validation, interpretation of results, general proficiency, and quality control measures.

A number of significant changes have occurred that reframe msAFP screening for ONTD since the original document was published in 2005.[2] This includes a lowered birth prevalence reflecting folic acid fortification of wheat flour in the United States,[3] increased folic acid supplementation,[4] improved public awareness of the roles that dietary folate and folic acid supplementation play in the prevention of neural tube defects,[5] advances in first and second trimester prenatal

ultrasound for the identification and characterization of neural tube defects and other structural anomalies,[6, 7] increased use of electronic patient records, and generally improved access to health-care services (uninsured rate for adult women dropped from 19.3% in 2010 to 14.3% in 2014).[8] In addition, there is an ongoing shift away from biochemical markers toward circulating cell-free DNA (cfDNA) for the detection of fetal aneuploidies[9, 10] that is likely to impact the current process for collecting a maternal serum sample for ONTD and Down syndrome screening, potentially making msAFP a standalone screen for only ONTD. Given the marked differences in access to ultrasound and prenatal care across the United States, and around the world, the routine second-trimester msAFP screening for ONTD is associated with useful information for pregnancy care. For example, such screening can identify incorrectly dated pregnancies, multiple gestations, other birth defects (e.g., open ventral wall defects), and rare maternal conditions (e.g., liver cancer). For women found to be at risk of ONTD due to msAFP results, follow-up diagnostic testing is likely to include targeted ultrasound testing and, in some instances, the testing of amniotic fluid for AFP and acetylcholinesterase (AChE). Guidelines for amniotic fluid testing are found in the online supplement.

METHODS

This document was informed by a review of the literature, including current guidelines and expert opinion. Resources consulted included PubMed searches, the ACMG Standards and Guidelines for Clinical Genetics Laboratories, Clinical and Laboratory Standards Institute (CLSI) guidelines, and CLIA regulations. When the literature provided conflicting evidence about a topic, or when there was insufficient evidence, the authors used expert opinion and current practices to inform these recommendations. Expert opinion included the coauthors of the document, members of the ACMG Laboratory Quality Assurance Committee, as well as experts consulted outside of the Committee, but acknowledged in this document. Any conflicts of interest for workgroup member or consultants are listed. A draft was delivered to the ACMG Board of Directors for review and member comment. The draft document was posted on the ACMG website and an email link was sent inviting ACMG members to provide comment. All comments were assessed by the authors. When appropriate, additional evidence was included to address member comments and the draft was amended. Both member comments and author response were reviewed by a representative of the ACMG Laboratory Quality Assurance Committee and by the ACMG Board of Directors. A final document was approved by the ACMG Board of Directors. This updated technical standard replaces the previous version.[2]

BACKGROUND

Clinical description of neural tube defects

Neural tube defects (NTDs) are among the most common serious birth defects with a prevalence in the United States of 7 per 10,000 live births.[5] Historically, the frequency has been lowest in the western United States, increasing to the south and east, with the highest rates occurring in the southern Appalachian region.[11] NTDs arise when the embryonic neural tube fails to close during the first 28 days postconception. The clinical consequences of NTDs are dependent on the site and severity of the defect. The focus of msAFP screening is primarily OSB. Spina bifida results from failed closure along the posterior neural tube resulting in spinal dysraphism with disruption of the vertebral bones, meninges, and often the spinal cord. Spina bifida occurs at a rate of 4/10,000 live births and in 80% the defect is open (not covered by skin or membrane). These are considered OSB and occur at a rate of 3/10,000 live births. OSB pregnancies can be identified through early second-trimester msAFP screening.[12] Twenty percent of spina bifida cases have closed defects and cannot be detected by msAFP screening. Failed closure of the anterior neural tube results in severe disruption of the development of the brain and cranial vault, resulting in anencephaly. Anencephaly occurs in 2.9/10,000 and is uniformly lethal, usually resulting in miscarriage, stillbirth, or early perinatal death. These pregnancies will also be identified via msAFP screening but can also be easily identified via ultrasound. More rarely (0.1/10,000), disruption occurs along the cranial ridge resulting in encephalocele, most of which are closed defects. Clinical effects of these NTDs vary widely from no impairment of function to lethality, but most often result in some degree of paralysis, hydrocephaly, and incontinence.

Most NTDs follow a complex/multifactorial pattern of inheritance without recognizable Mendelian patterning. The majority occur as isolated birth defects (nonsyndromic). As with most multifactorial birth defects, there are literature reports of apparent Mendelian patterning within the nuclear family. Syndromic NTDs are those associated with other birth defects, and these may occur as part of specific single-gene disorders or chromosome disorders. Environmental factors that have been associated with an increased likelihood that a fetus will be affected with an NTD include inadequate maternal folate intake,[13, 14] exposure to anticonvulsant medications (e.g., valproate)[15–17] (or other folate antimetabolites), early maternal hyperthermia, and increased maternal body mass index (BMI).[18, 19]

Preconception supplementation with folic acid reduces the incidence of ONTD by up to 80%, dependent on adherence, dose, and NTD prevalence.[20–24] Folate supplementation beyond the critical period for neural tube closure (28 days postconception, or 6 weeks from the beginning of the last menstrual period) is ineffective. Fortification of wheat flour in the United States has resulted in a decrease in NTD by approximately 30% nationwide.[4, 25]

What is alpha-fetoprotein?

AFP was first discovered in 1956[26] when protein electrophoresis was used to identify a new band located in the a-1 region in fetal serum and it was recognized that the fetus was producing a protein (69,000 daltons) in high concentration

that was not produced by adults. This band was labeled alpha-1-fetoprotein and it bears structural similarity to albumin. Its circulating half-life is approximately four days. The function of AFP during fetal development has been examined by several investigators. One obvious function is for AFP to maintain oncotic pressure in the intravascular compartment in similar fashion to albumin in the adult. Recently, however, three cases have been documented where the fetus produced minimal or no AFP.[27] All three cases went to term and were associated with healthy newborns of normal weight. As of the present time, a functional role for AFP during fetal development has not been clearly defined. AFP is synthesized in the fetal liver and yolk sac and is the dominant serum protein early in fetal life, reaching a concentration of approximately 2000 µg/mL in fetal serum in the early second trimester. Thereafter, its concentration decreases steadily.

How is measurement of AFP useful in identifying open neural tube defects?

Fetal urine also contains AFP and since amniotic fluid is mainly comprised of fetal urine, AFP would be expected to be measurable. Levels in the amniotic fluid in the early second trimester are about 20 µg/mL, about 100 times lower than in the fetal circulation. In maternal serum, during pregnancy, peak levels are about 100 times lower than in the amniotic fluid (about 0.2 µg/mL). Levels of AFP in the circulation of healthy nonpregnant women is essentially undetectable. Evidence points to the fetus as being the source of AFP during pregnancy, reaching the maternal circulation by diffusion across both the placenta and the amnion.[28] At every junction between the fetus and the external environment, membrane barriers prevent more than a small fraction (~1%) of the circulating AFP from escaping. These large concentration differentials help explain why, when there is an opening in the fetus (such as an ONTD), the amniotic fluid levels will rise, and careful measurement of amniotic fluid AFP is central to the antenatal diagnosis of anencephaly and OSB in early pregnancy.[29] However, since the placenta and amnion are intact, and serve as barriers to passage of AFP into maternal circulation, elevations in maternal serum are less predictable and less direct. Consequently, serum AFP measurement against gestational week-specific norms functions as a screening rather than a diagnostic test.

Screening versus diagnostic testing with AFP

Prenatal testing for ONTDs by measurement of second-trimester AFP is considered a screening test when measured in maternal serum, but part of a diagnostic testing protocol when measured in amniotic fluid. The distinction between a screening and a diagnostic test is important because the goals and expectations for the detection rate (sensitivity) and the false positive rate (1-specificity), costs, and acceptable level of invasiveness differ. msAFP screening results are not diagnostic of any condition. Rather, the screening process identifies pregnancies that are at sufficient risk for OSB or other related birth defects (e.g., anencephaly, open ventral

wall defects, Finnish type congenital nephrosis) to warrant counseling and the offer of additional diagnostic testing such as amniocentesis and further biochemical testing and/or a targeted ultrasound. The detection and false positive rates of msAFP as a screening test will be a function of several factors, including the gestational age when the sample is obtained, the method of gestational age estimation, AFP assay precision, and the msAFP cutoff level used to determine a "screen positive" result. If a woman has a family history of NTD or a screen positive msAFP result, diagnostic testing can be offered. The diagnosis of an ONTD can be made solely by a targeted ultrasound and might include finding of the so-called lemon (frontal scalloping of the calvarium) and banana (shape of cerebellum due to a shallow posterior fossa consistent with the Arnold Chiari malformation) signs[30] as well as visualization of spinal defect. Alternatively, the finding of both elevated AFP and elevated acetylcholinesterase measurements in the amniotic fluid are also considered diagnostic of an open neural tube defect. Usually the defect is confirmed via ultrasound prior to a definitive diagnosis being made.

Prenatal msAFP screening for ONTD is best implemented in the context of a comprehensive program that coordinates preanalytic, analytic, and postanalytic components of the process.[31] Although the prenatal screening laboratory utilizes clinical chemistry methods such as enzyme immunoassays, the role of the laboratory extends beyond the performance of the AFP assay because the results require a unique kind of interpretation. This interpretation puts the results of the test into the appropriate context of a priori risks as determined by race, gestational age, and family history. The laboratory director is often called upon to provide consultation regarding these risks and options for further action. To address these unique requirements, the laboratory director must generally meet the standards set out in section B3 of the ACMG Guidelines.[32] When prenatal screening for ONTDs is performed in a clinical chemistry laboratory in which the director does not meet these standards, the laboratory should have a formal professional relationship with an individual who does meet the standards set out in section B3, and who is available in a timely fashion to aid in interpretation and provide consultation when requested.

Impact of folic acid fortification and supplementation

Conclusive studies have demonstrated primary prevention for a high proportion of recurring[21] or incident[20] neural tube defects via sufficient folate early in pregnancy. Dietary folate alone is generally inadequate, but supplementation with 400 micrograms of folic acid per day is close to optimum. However, the neural tube closes at about four weeks of pregnancy, earlier than pregnancies are usually recognized. Additionally, many pregnancies are unplanned and, consequently, supplementation provides only a partial solution. Fortification of flour has now been implemented in the United States resulting in over 600 prevented spina bifida births each year.[33] Currently, over 80 countries[34] require

fortification of industrially milled wheat flour with folic acid providing women of child-bearing age sufficient folic acid to provide a measurable reduction in the birth rate for ONTD.[35] In the United States, supplementation and fortification together have dramatically reduced the incidence of isolated spina bifida and anencephaly.

Considerations for open spina bifida
When OSB is diagnosed, factors such as other genetic disorders (e.g., trisomy 18), location of the open defect, size of the defect, and other anomalies need to be assessed. Options include termination, surgical repair after delivery, or surgical intervention during pregnancy. The American College of Obstetricians and Gynecologists (ACOG) suggests serial ultrasound examinations and delivery at a tertiary care center equipped to handle any complications.[36]

PREANALYTIC REQUIREMENTS
Sample types
Few published data regarding failure rates for different sample types exist from screening programs, but AFP 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). Since 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 relatively uncommon and can be resolved by repeat testing. In rare instances, a second sample may be requested.

Sample requirements
Blood samples should be collected using standard phlebotomy techniques. Although serum measurements are the norm, each laboratory must specify what samples are acceptable (e.g., whole blood, serum separator tube, spun serum separator tube) based on validation performed in their own laboratories. Specimen containers should be appropriately labeled with at least two patient-specific identifiers, collection date, and follow relevant state and federal guidelines.

Condition of samples: shipping, handling, and storage
Standards for acceptable specimen handling from collection site to the laboratory should be specified, including packaging, mode of transportation, and temperature range. AFP in maternal serum is very stable and can generally be shipped at ambient temperature. However, the same specimen may also be used as a primary screen for common autosomal aneuploidies (e.g., Down syndrome) where some analytes are less stable than msAFP (e.g., human chorionic gonadotropin [hCG] and unconjugated estriol [uE3]). When this is the case, the conditions for sample handling should be restricted to that of the least stable component.[32]

Criteria for sample rejection should be made clear by the laboratory. Variables that can affect the acceptability of a sample for ONTD screening or a specific AFP assay protocol should be established and communicated by the laboratory, including both clinical (e.g., gestational age out of range) and sample-related characteristics (e.g., inappropriate sample type, insufficient quantity, gross hemolysis). Protocols for sample processing should be designed to avoid contamination, tampering, or substitution. Handling samples must be in accordance with Occupational Safety and Health Administration (OSHA) guidelines, with the express understanding that any human fluids may harbor infectious agents. AFP is stable relative to other serum components and can be reliably determined in sera stored at 4–8 °C for several days and at -20 °C for years. Each laboratory should establish its own policies regarding specimen retention.

METHOD VALIDATION
Testing personnel
Laboratory personnel performing msAFP screening for ONTD must receive appropriate training and ongoing competency via an established and documented laboratory protocol. Laboratory personnel must also meet all relevant CLIA requirements for high-complexity testing with a minimum of an associate's degree in laboratory science or medical laboratory technology from an accredited institution. Stricter requirements apply in some states.

Maternal serum AFP: assay methodologies
General guidance on developing assay protocols is available through the American College of Medical Genetics and Genomics Standards and Guidelines for Clinical Genetics Laboratories (see section C).[32]

In the United States, the Food and Drug Administration (FDA) licenses AFP kits as an aid in the diagnosis of ONTDs. As class III devices, these kits are approved to reliably measure AFP in second-trimester maternal serum samples and amniotic fluid. Available kits include immunometric, chemiluminescent, and colorimetric methods, all capable of measuring AFP reliably in the range of values important for ONTD screening (25 to 150 IU/mL).

AFP standards can be calibrated in either mass units (ng/mL) or international units (IU/mL). Each AFP kit manufacturer provides a factor for converting mass units into international units. Conversion factors should be considered manufacturer-specific. Commercially available AFP kits provide calibrators and specific calibration protocols. Laboratories utilizing laboratory developed tests (LDTs) or modifying AFP kit assay protocols are responsible for determining calibration protocols and validating performance.

Internal quality control
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, AFP levels appropriate for ONTD clinical action points, 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.

Repeat assay controls (RACs) are also helpful for 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 serum AFP levels are stable when frozen and thawed, reassaying stored patient samples from the time when the current median values were established can help to identify any long-term assay drift and determine if reference data need to be updated.

Each batch assay should contain at least two quality control (QC) samples that fall at clinical action points (three controls may be required to comply with some licensure requirements). For example, low controls could be targeted at serum AFP values falling at 0.5 multiples of the median (MoM) for 16 weeks of gestation (suitable for Down syndrome screening). Normal or midrange controls could be targeted near the 16-week median (1.0 MoM), and high controls at a value near commonly used ONTD cutoff levels (2.0 to 2.5 MoM).

Following preparation and aliquoting, performance ranges for in-house pooled controls can be set using standard clinical laboratory quality control approaches. Controls received with AFP 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 care should be taken to set appropriate ranges and avoid unnecessary result rejection.

Standard approaches for QC assessments used in the clinical laboratory are appropriate for internal QC of AFP assays, including the type and frequency of assessments. As part of the initial method verification, the laboratory should demonstrate that intra- and interassay variation reported by the manufacturer can be reproduced and specify internal measures of repeatability. Standard approaches to routine equipment calibration and preventive maintenance used in the clinical laboratory are appropriate. In many cases, calibration and maintenance protocols are set by the product/equipment manufacturer.

Because of the impact of as little as 10% systematic change in assay performance on detection and false positive rates, laboratories need to select AFP kits 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). Because AFP MoM values are calculated using reference data collected in the past, it is also important that kits/reagents are stable over a long period of time, and that lot-to-lot variability is minimized.

In-house pooled controls (or commercial products obtained in sufficient quantity to last a year or more) and RACs are valuable for monitoring long-term assay drift and lot-to-lot variability. Median values should be reviewed at regular intervals by the laboratory and recalculated at least annually. Medians should be recalculated if there is a shift in msAFP values greater than 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 AFP values can be monitored by computing the overall median MoM level. Observations from the past should be used to calculate medians only if epidemiological monitoring shows the median MoM has been stable for the time period over which the median values are calculated. Alternative methods of revising medians may be necessary if a significant shift has been observed. Monitoring MoM values should allow for approximately 200 to 500 (or more) samples in each time period. For smaller laboratories, this could result in monthly monitoring. For larger laboratories, this might occur each week.

msAFP: establishing reference ranges

AFP median levels over the appropriate gestational age range should be evaluated with new AFP reagent lots for optimal screening performance. Between 25 and 50 patient samples and current controls can be assayed on the old and new kit/reagent lot and the relationship between the two examined using techniques of regression analysis and method comparison after logarithmic transformations. This relationship can then be applied to the existing medians to derive new medians that can be used until sufficient data are available for optimum analysis.

Normative values of AFP change throughout gestation, increasing approximately 10% to 15% per week between 15 and 22 weeks of gestation.[12] For AFP measurements to be accurately interpreted, each result in mass or IU 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 and race.[37]

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. During startup, it may be acceptable to demonstrate that medians obtained from another source are appropriate for its screened population. Interpretation of AFP values requires the establishment of laboratory-specific AFP medians by gestational age. Package insert (commercial) medians should not be used, even for a short time. Several empirical methods exist that can be utilized to establish reliable medians. The optimal approach to establishing median values is an empirical approach in which each distinct patient population is tested for AFP and median values are established for each week (or, preferably, decimal week) of pregnancy. For example, if numbers are sufficient, separate medians should be computed for Caucasian and African American pregnant women.

Historically, 100 samples for each gestational week from 15 through 20 were used to calculate median values for each independent population identified by the laboratory. Because

AFP is stable, it is possible to use stored frozen specimens collected over several years, although excessive freeze–thaw cycling should be avoided. It is not necessary that all samples used be from unaffected singleton pregnancies because outlying values are uncommon and will have a negligible impact on the median value. Since the vast majority of specimens are drawn in a narrower gestational age range (16–17 weeks), it may be difficult to obtain significant numbers of samples beyond 18 weeks of gestation. Application of regression analysis allows use of fewer samples (e.g., $N = 300$) over the 15- to 20-week period to establish reasonably reliable medians.

Optimal screening performance can be achieved by considering gestational age as weeks and days or decimal weeks (e.g., 15 weeks and 5 days is 15.7 weeks). Less optimal is using gestational age expressed in completed weeks (e.g., 15 weeks and 5 days is 15 completed weeks). Expressing results in rounded weeks (e.g., 15 weeks and 5 days is 16 weeks) is not recommended.

Statistical smoothing of the observed median values by weighted log-linear regression analysis (logarithms of medians regressed versus gestational age in days or completed weeks, weighted by the square root of the number of observations in each category) provides reliable and accurate medians. This method also allows median values for weeks in which few data are available.

Epidemiological monitoring

In recent years, stricter privacy and confidentiality practices have made it much more difficult to collect pregnancy outcome information and information regarding follow-up of medical procedures (such as ultrasound and amniocentesis) performed subsequent to positive screens. This makes computation of detection rate difficult, but ONTD detection rates have been validated in a wide variety of settings over time.

However, to monitor assay and program performance and to identify possible areas of concern, screening programs must perform epidemiological monitoring.[38] Such monitoring, at a minimum, should include (1) periodic computation (monthly or weekly, depending on numbers of samples processed) of the median msAFP MoM, determination of the statistical significance of any deviation from 1.00, and documentation of any necessary corrective action; and (2) periodic computation of the rate of initial screen positive results and comparison of that rate to expected published rates, after taking into account variables such as the screening cutoff level used and the proportion of pregnancies dated by ultrasound.

Analytic validity

Analytic validity defines the test's ability to accurately and reliably measure a specific analyte that is to be used clinically. Each laboratory is responsible for in-house validation of a test methodology but information in the package insert of an FDA-approved kit or from the literature can be used as supplementary supporting evidence.

Analytic sensitivity is commonly defined in the laboratory as an assay's lower limit of detection. However, in the context of maternal serum screening, we are defining analytic sensitivity as the proportion of samples with elevated AFP levels that are correctly classified as being high. Analytic sensitivity can be determined using samples with high consensus AFP levels (e.g., selected proficiency testing samples).

Analytic specificity is commonly defined in the laboratory as the extent to which a method measures an analyte exclusively and does not cross-react with other related compounds. However, in the context of maternal serum screening, we are defining analytic specificity as the proportion of samples with low or normal AFP levels that are correctly classified as being low or normal. 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. Many laboratories also repeat samples with msAFP MoM levels greater than the specified ONTD cutoff level.

Assay robustness measures how resistant testing is 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, failure rates), laboratories should consider the effects of common variables, such as sample type, sample handling (e.g., transit time, conditions), sample quality, reagent lots, or minor changes in assay conditions (e.g., timing or temperature).

Clinical validity

Clinical validity defines the ability of the test to accurately and reliably identify the clinical phenotype of interest. In this instance, it is the ability of msAFP measurements to identify pregnancies in which the fetus is affected with an ONTD.

Clinical sensitivity is the proportion of pregnancies with an ONTD that have a positive test result. Clinical specificity is the proportion of unaffected pregnancies identified by the test as being negative (1 minus false positive rate). Clinical sensitivity (detection rate) and clinical specificity will depend on many factors, including the MoM cutoff level chosen, the method of estimating gestational age, and the gestational age at screening.

Decisions related to cutoffs are generally based on OSB performance, treating anencephaly screening performance as a derivative screen test. The detection rate for OSB is expected to be between 75% and 90% using a 2.0 MoM screening cutoff level, and between 65% and 80% using a 2.5 MoM cutoff level. False positive rates are expected to be between 2% and 5% and between 1% and 3%, respectively. These rates are influenced by many factors (e.g., gestational age at screening, dating method, inclusion/exclusion of multiple gestations) that have been discussed earlier. Using either common ONTD screening cutoff level (2.0 or 2.5 MoM), the detection rate for anencephaly is expected to be 95% or greater. As the

proportion of ultrasound dating increases (especially the use of biparietal diameter), the detection rate moves toward the upper end of the above ranges and the false positive rate moves toward the lower end of these ranges.

The positive predictive value (PPV) and negative predictive value (NPV) of AFP testing in the target population measure the ability of the test to give accurate clinical information with respect to test outcome (i.e., reliability of positive and negative tests). The positive predictive value is the proportion of positive test results that correctly identify a pregnancy with an ONTD [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). The PPV and OAPR are computed using the prevalence of OSB and anencephaly and the respective detection and false positive rates. PPV is known to be heavily influenced by differences in birth prevalence. As an example, in the general population where birth prevalence is 1 in 2000, given a detection and false positive rate for OSB screening of 80% and 3%, respectively, the corresponding PPV would be about 1.3% ([80% / 3%] × 1 in 2000, or 1 in 75). Although this PPV seems low, follow-up testing is often a noninvasive targeted ultrasound with invasive testing being far less common. In contrast, the PPV for a couple whose first child was affected with a neural tube defect (assuming a recurrence risk of 1 in 45) would be about 59% ([80% / 3%] × 1 in 45, or about 1 in 2). These examples show only the population risks and probabilities, but in actual practice patient-specific risks may also be computed with a similar approach based on the patient's own measured AFP level and her estimated risk based on clinical history.

The NPV is the proportion of negative tests that correctly identify an unaffected pregnancy (true negatives / [true negatives + false negatives]). Because the prevalence of the conditions being screened for (OSB and anencephaly) is low, the NPV is not often useful in decision-making, and is generally not computed.

Clinical utility

Clinical utility addresses the risks and benefits associated with testing in 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 answer these questions in the future. 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
- Adopting quality assurance processes that monitor the effectiveness of the laboratory's ongoing testing activities
- Understanding possible adverse health or psychosocial consequences of testing

- Which follow-up testing or interventions in persons with positive and negative test results are reasonable
- Accessibility of testing for the general pregnancy population by payer status
- Alternate testing options, such as ultrasound, and differences in their positive predictive values
- Understanding what is known about the financial costs and economic benefits of testing

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

External proficiency testing

Each laboratory must participate in an external proficiency testing program that evaluates assay performance for msAFP in the second trimester. If this is not possible, the laboratory must utilize other recommended external proficiency testing methods, such as scheduled interlaboratory comparisons or split sample analysis with another laboratory.

SERUM SCREENING TEST INTERPRETATION AND REPORTING

Patient and provider information

Laboratories should either provide educational materials such as brochures or short videos for use by patients in consultation with their providers or, at a minimum, provide information about where such materials can be obtained. The Centers for Disease Control and Prevention, March of Dimes, and other institutions and laboratories[39–42] 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 benefits of testing, and what to expect if the screening test is positive. Laboratories should supply health-care providers they serve with informational materials that include the following:

- Detailed information about the sampling process and how samples should be labeled and transported to the laboratory
- 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 availability of results online, through electronic delivery or through other means of reporting

ACMG TECHNICAL STANDARDS

- Information about expectations for test performance (detection rate, false positive rate, positive predictive value, and failure rate) and reporting formats

Patients should be informed about the benefits (e.g., eventual diagnosis and reproductive choice, reduced anxiety with negative results, identification of unknown twin pregnancies) and limitations (e.g., not 100% sensitive, increased anxiety with a false positive result) of prenatal screening prior to testing. It is the duty of the health-care professional, not the laboratory, to inform and obtain any necessary consent for testing, but the laboratory may be required to document such consent. 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 patient education and informed consent.

Intake information

The collection of complete and accurate intake information is necessary for providing the most reliable interpretation. The laboratory should document their policies for addressing any missing patient information. Laboratories should have a mechanism to collect pretest clinical information through well-designed requisition forms that includes:

- Basic required demographic information
- Relevant medications
- Gestational age (in decimal format)
- Method of dating the pregnancy (preferably by US measurement of biparietal diameter)
- Maternal weight
- Maternal race
- Maternal insulin-dependent diabetes mellitus (IDDM) prior to pregnancy
- Number of fetuses
- Previous screening in the current pregnancy (i.e., initial or repeat serum sample)
- Family history of neural tube defects

Interpretation of msAFP results

The optimal time for ONTD screening by msAFP measurement is 16 to 18 weeks but screening can be performed between 15.0 and 20.9 weeks.[43] Screening performance is significantly decreased in the 14th week of gestation. Under special circumstances, laboratories may accept samples later than 20 weeks of gestation provided sufficient numbers of cases are available to determine medians and maintain quality assurance. The choice of 16 weeks for the optimal screening week would allow time for counseling, follow-up diagnostic testing, and personal decision-making, were the test result to be screen positive.

The means used by the health-care provider to establish the gestational age of the fetus should be considered. A common method for determining gestational age is dating by the first day of the last menstrual period (LMP). Although LMP-based dating is sufficiently accurate for ONTD screening, gestational age estimation based on ultrasound measurements is a more accurate approach.[44] Its use increases detection and reduces false positive rates. Ultrasound measurement of crown–rump length (CRL) in the first trimester provides an accurate estimate of gestational age.[45] It has been established that first-trimester ultrasound dating using a CRL is more accurate than second-trimester dating and guidelines regarding when gestational ages should be modified based on ultrasound measurements have been published.[46] However, the optimal dating method to use for OSB screening is the biparietal diameter (BPD).[47, 48] In affected pregnancies, this dating is about two weeks earlier than the true gestational age resulting in much higher AFP MoM levels in affected pregnancies. Such a measurement would also identify virtually all cases of anencephaly. However, the BPD should not be used to determine whether the pregnancy is at an appropriate gestational age for screening (e.g., 16–18 weeks). Using the BPD measurement in pregnancies with an open defect to date the pregnancy (with the two-week discrepancy) might result in the msAFP test being performed too early in gestation to be reliable (e.g., 14 weeks or earlier).

The method of determining gestational age can be taken into account when providing interpretations in two ways. First, separate medians can be calculated for those pregnancies dated by LMP and those dated by ultrasound measurements. Secondly, separate Gaussian population parameters can be utilized in determining risk. Assigning gestational age based on ultrasound measurements has the effect of "tightening up" the distribution of msAFP measurements in both unaffected and affected pregnancies because the variance in MoM values is lower. For this reason, separate sets of distribution parameters can be used for LMP and ultrasound dated pregnancies. Ultrasound dating based on BPD measurement reduces the screen positive rate and significantly increases the detection rate for OSB.[47, 48] BPD dating also rules out anencephaly because OSB fetuses have, on average, BPD measurements equal to a 2-week younger fetus. If gestational age estimates were based on BPD, there would be significant improvement in OSB detection at any screening cutoff level. Other ultrasound measurements (e.g., crown–rump length or multiple second-trimester measurements) can reliably date the pregnancy but do not have this unique advantage of BPD dating.

Interpretive refinements

There are many interpretive refinements based on patient demographics and other pregnancy-related information that are less critical than gestational age, but which will still improve screening performance by optimizing the interpretation and reducing overall variability. Currently, most laboratories consider maternal weight, maternal race, and maternal insulin-dependent diabetes.

The msAFP levels are, on average, higher in lighter weight women and lower in heavier weight women. This is likely a dilution effect. Adjusting msAFP values for maternal weight

improves ONTD screening performance and should be done.[49–51] Laboratories should only utilize published weight adjustment formulas for a short time until sufficient in-house data are collected and new laboratory-specific formulas derived. Once about 1000 to 2000 samples are available, in-house maternal weight adjustment equations can be derived. Of critical importance in the weight correction equation is an accurate representation of the mean maternal weight for the population being screened. In fact, the maternal weight equation in use can be set to a MoM of 1.00 and solved for the weight. This weight should be within a few pounds of the average weight in the population being tested. Failure to periodically recompute the maternal weight relationship as the average weight in the population changes will diminish the accuracy of the MoM value for each patient.[51] Several publications have documented systematic differences in maternal weight by race (see below), and consequently, separate maternal weight equations by race can be implemented. There is also mounting evidence that maternal weight is positively associated with ONTD (heavier women are more likely to have ONTDs) and laboratories could also account for this association when computing patient-specific risks.[52–54]

Adjustments for maternal race should be incorporated.[55] If sufficient data are available, the preferred adjustment method is to calculate a separate set of medians for each of the groups. If too few observations are available in one or more of the groups, a correction factor may be applied to the MoM when screening those pregnancies. When exploring the relationship between race and msAFP MoM, the analysis should include accounting for the systematic differences in maternal weight expected in various racial groups. Several studies have documented systematic differences in maternal weight by race (in general, Asian women are reported to be the lightest, Black/African American the heaviest, with Hispanic and non-Hispanic Caucasians in between).[56] If a program has a large and relatively diverse population, routine monitoring of the median MoM and proportion with MoMs over 2.0 (or 2.5) for each racial group may prove useful in identifying patient subgroups meriting adjustments or other special considerations.

Pregnant women with insulin-dependent diabetes mellitus (IDDM) prior to pregnancy are at a severalfold higher risk of ONTD.[57, 58] In 1979, msAFP levels were first reported to be 20% to 30% lower in women requiring insulin prior to pregnancy.[59, 60] More recently, conflicting data have been published regarding the association between msAFP and insulin use during pregnancy.[61–63] Many programs apply an adjustment factor for msAFP MoM levels in women with IDDM. However, there is no consensus on whether this correction should be applied to gestational diabetic women or to women who can be controlled by oral agents.[64]

Although supporting data are sparse, most screening programs assume that the effects of weight, race, and IDDM are independent, and thus all may be applied to the same patient. If a program is sufficiently large to support separate analyses, observed effects could be used in place of correction factors.

The laboratory may choose to contact health-care providers if critical patient information does not accompany the specimen. If the laboratory does not receive critical patient information, the written report should indicate that the information is missing and what assumptions, if any, were used in the interpretation. In some cases, including a statement on the report about the potential impact of the missing information may be warranted (e.g., maternal weight, race). In other cases, full interpretation may not be possible (e.g., no gestational age).

Laboratories should be able to compute patient-specific risks for anencephaly and spina bifida even though they may not be routinely reported. Patient-specific risks that take into account the patient's own measured AFP level and a priori risk for ONTD are often dramatically different than the computed risks for screen positive and screen negative tests for the screening program as a whole. Some licensing agencies may require the ability to report patient-specific ONTD risks and it is important to be able to provide this information to health-care professionals for clinical counseling and pregnancy management. Patient-specific risks are generated by complex mathematical algorithms that are integral to prenatal screening. The use of specialized software applications is generally considered a necessity for ONTD screening, due to the complex nature of calculating and interpreting the results, the need for patient-specific interpretive reports, and because of the large number of samples processed. Software to perform these calculations can be obtained commercially or developed in-house. Software must be verified prior to routine clinical use. Usually, separate estimates of risk for OSB are added to the risks for anencephaly to create a patient-specific ONTD risk.

Care should be taken to verify that the most accurate a priori risk value is used in each computation as this heavily influences the calculated risk. The patient's family history of NTD, medical history, and race/ethnicity can be used to provide a more accurate a priori risk. Note that most estimates of population risk were established prior to the introduction of folate fortification, and are therefore likely overestimated.[4] A priori risks do not generally consider individual patient folate intake, which would significantly alter the computation of a patient-specific risk. Laboratories should be aware of the assumptions applied to their calculations of risk and the limitations imposed by any assumptions.

Computing patient-specific risks

The commonly used algorithm to estimate the patient-specific risk utilizes a Bayesian approach to modify the prior risk for each condition using a likelihood ratio calculated from the woman's specific MoM value, after being adjusted for variables such as weight, race, and IDDM status as discussed above. The likelihood ratio is derived from the overlapping

Gaussian distributions described by the affected and unaffected distribution parameters.[43, 65]

Risk algorithms utilize published or in-house population parameters for msAFP, expressed as log means (or medians) and log standard deviations of the msAFP distributions in unaffected pregnancies and in pregnancies affected with OSB or anencephaly. Population parameters for each of these disorders can vary based on factors such as gestational age at the time of testing and gestational dating method.[43] There is no formal consensus on which adjustments to the result or prior risk to include, specifically how to include them, or how inclusion influences screening performance. These decisions are left to the laboratory director's discretion. Incorporating data regarding folate fortification or supplementation into patient-specific risk calculations is also left to the discretion of the laboratory director.

Determination of "screen positive" results most commonly relies on a preset msAFP MoM cutoff level. Few laboratories choose the ONTD risk estimate as the screening variable because of the inherent difficulty in computing reliable risks. Typically, msAFP cutoff levels for ONTD screening range between 2.0 and 2.5 MoM. Screen positive results are defined as those with an msAFP MoM greater than or equal to the cutoff level.

The background risk (or birth prevalence in the absence of prenatal diagnosis and selective termination) may be higher or lower in certain populations. The published literature indicates that the birth prevalence of ONTDs is increased severalfold in women with IDDM. Race may also influence birth prevalence estimates. For example, using the 1995–2011 birth prevalence data in Caucasians as a reference point, the combined rate of OSB and anencephaly is approximately 25% lower in Black/African American pregnancies, and 40% higher in Hispanics.[5] Family history of ONTD may also increase the prior risk, depending on the number of affected relatives and the degree of relatedness. Family history can be incorporated into the ONTD risk estimate using available algorithms.[66] The laboratory may routinely include a recommendation for genetic counseling, if a positive family history is identified.

To address differences in prior risk, the screening cutoff level may be modified to keep the risk of an ONTD at the MoM cutoff roughly equal (iso-risk screening). Use of a higher screening cutoff would be a means of addressing the lower positive predictive value that occurs in a population with a lower birth prevalence. Alternatively, the AFP MoM cutoff could be kept constant and the overall detection rate would remain constant (isodetection). This would ensure that the same proportion of affected fetuses would be identified in both populations and would be insensitive to the differences in positive predictive value. Either approach is acceptable, but the laboratory should understand the tradeoffs associated with the different approaches. Screening programs that use a 2.0-MoM cutoff level in Caucasian pregnancies might use 2.5 MoM for Black/African American pregnancies since the risk is similar for the two groups at these specified levels.

Alternatively, the same cutoff level could be used for the two groups, resulting in a lower frequency of true positives but identical detection rates. It is not possible to have equal positive predictive values and equal detection rates for populations with different birth prevalence.

Screening twin pregnancies

Twin pregnancies are known to have msAFP levels approximately two times the levels in singleton pregnancies. Distribution parameters for msAFP measurements have been defined for unaffected twin pregnancies and for twin pregnancies in which one or both of the fetuses are affected with OSB or anencephaly.[67] The birth prevalence of ONTD is also higher in twin pregnancies, with one report estimating that an ONTD is 2.28 times more likely (per fetus) than in a singleton pregnancy.[67] The msAFP cutoff levels for consideration of amniocentesis should be determined separately for twin pregnancies. Typically, cutoff levels fall between 4.0 and 5.0 MoM. Other factors, such as the acceptability within the medical community of performing amniocentesis on twin gestations and the difficult options should one affected fetus be identified, should also be considered in setting the cutoff level.

Repeat testing

Obtaining a second specimen for repeat testing may be beneficial when the initial specimen has a slightly elevated msAFP (relative to the screening cutoff level) and the gestational age is early enough to allow time for appropriate follow-up. Most laboratories do not combine results for the two tests but rather employ a simple set of rules for interpreting results of repeat testing. Methods for combining the results of the two tests have been published[68] and the laboratory should develop a policy that indicates the method to be used. If the pregnancy was originally misdated and the revised gestational age is too early for interpretation (e.g., 14 weeks or earlier), the subsequent sample can be considered to be the first usable sample. Repeat testing may be of particular use when access to ultrasound is limited. Where ultrasound follow-up is readily available, repeat testing is generally not useful.

Results reporting

Reports should contain appropriate patient and specimen information as described in the American College of Medical Genetics and Genomics Standards and Guidelines for Clinical Genetics Laboratories, section C[32] and as specified by CLIA. Final reports of test results must be clear to a nongeneticist health-care 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 was ordered
- Type of specimen
- Date when sample was obtained

- Laboratory accession number(s)
- Demographic and pregnancy-related information used in the interpretation (e.g., gestational age, method of dating, maternal race, maternal weight)
- Analytic results in both mass units (e.g., ng/mL) and interpretive units (i.e., MoM) upon which all adjustments/corrections have been performed
- Clinical interpretation, including whether the result is screen positive or screen negative, the msAFP MoM level, the MoM cutoff level, and the patient-specific risk
- Potential follow-up steps could include a detailed ultrasound capable of detecting ONTD and/or diagnostic testing of amniotic fluid for amniotic fluid alphafetoprotein (AFAFP) and amniotic fluid AChE

Screen negative patient reports can be transmitted to the referring physician by electronic transmission, US mail, courier, or overnight carrier.

Screen positive results should be promptly transmitted to the referring health-care provider by some method that ensures prompt receipt by the referring provider, 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:

- A dating ultrasound to confirm gestational age and fetal viability and to rule out twins, anencephaly, and other fetal defects
- Referral for genetic counseling
- Referral for targeted ultrasound examination
- Amniocentesis with AFAFP and AChE testing
- Referral to maternal fetal medicine specialist for consideration of counseling, targeted ultrasound, or additional testing
- Repeat sampling

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 as negative. Reclassification usually occurs when an LMP dated pregnancy is subsequently dated by ultrasound, and the difference between the LMP and ultrasound dating exceeds a set standard. As guidance to laboratories, reclassification should not be considered unless the revised estimate of gestational age is different by at least a week. Many laboratories use 10 days (e.g., 1.5 weeks) as the standard. One way to help avoid reclassification and improve overall screening performance is to encourage physicians to base their initial gestational age estimates on ultrasound measurements, preferably BPD measurements.

Conditions other than an ONTD that are associated with elevated msAFP MoM levels include:

- Underestimated gestational dating
- Multiple gestation
- Recent fetal demise

- Ventral wall defects (e.g., omphalocele, gastroschisis)
- Finnish type congenital nephrosis

In addition, women with unexplained elevated msAFP level have an increased incidence of poor pregnancy outcome and other complications (e.g., poor fetal growth, stillbirth, hypertension associated conditions, maternal liver disease, and placental abruption).

Markedly low levels of msAFP do not generally merit clinical workups apart from the inclusion of msAFP as a marker in aneuploidy screening. Rarely, pregnancies may produce no msAFP, and this is not known to be associated with any risk of adverse outcome.[69]

REFERENCES
1. American College of Medical Genetics and Genomics. Technical standards and guidelines. 2008 edition with upgrades through March 2007. https://www.acmg.net/ACMG/Medical-Genetics-Practice-Resources/Technical_Standards_and_Guidelines.aspx?WebsiteKey=6e814a8c-3077-4552-ba39-f7fcacff42d6. Accessed 1 October 2019.
2. Bradley LA, Palomaki GE, McDowell GA, Group OW, ACMG Laboratory Quality Assurance Committee. Technical standards and guidelines: prenatal screening for open neural tube defects. Genet Med. 2005;7:355–369.
3. United States Food and Drug Administration. Code of Federal Regulations Title 21. 2015. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=101.79. Accessed 1 October 2019.
4. Crider KS, Bailey LB, Berry RJ. Folic acid food fortification—its history, effect, concerns, and future directions. Nutrients. 2011;3:370–384.
5. Williams J, Mai CT, Mulinare J, et al. Updated estimates of neural tube defects prevented by mandatory folic acid fortification—United States, 1995-2011. MMWR Morb Mortal Wkly Rep. 2015;64:1–5.
6. Fleurke-Rozema JH, Vogel TA, Voskamp BJ, et al. Impact of introduction of mid-trimester scan on pregnancy outcome of open spina bifida in The Netherlands. Ultrasound Obstet Gynecol. 2014;43:553–556.
7. Wald NJ, Cuckle HS, Haddow JE, Doherty RA, Knight GJ, Palomaki GE. Sensitivity of ultrasound in detecting spina bifida. N Engl J Med. 1991;324:769–772.
8. Agency for Healthcare Research and Quality. Chartbook on women's health care: access to health care. 2015. http://www.ahrq.gov/research/findings/nhqrdr/2014chartbooks/womenhealth/wm-access.html. Accessed 1 October 2019.

9. Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. Lancet. 1997;350:485–487.

10. Palomaki GE, Kloza EM, Lambert-Messerlian GM, et al. DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. Genet Med. 2011;13:913–920.

11. Greenberg F, James LM, Oakley GP Jr. Estimates of birth prevalence rates of spina bifida in the United States from computer-generated maps. Am J Obstet Gynecol. 1983;145:570–573.

12. Wald NJ, Cuckle H, Brock JH, Peto R, Polani PE, Woodford FP. Maternal serum-alpha-fetoprotein measurement in antenatal screening for anencephaly and spina bifida in early pregnancy. Report of U.K. collaborative study on alpha-fetoprotein in relation to neural-tube defects. Lancet. 1977;1:1323–1332.

13. Zhang T, Xin R, Gu X, et al. Maternal serum vitamin B12, folate and homocysteine and the risk of neural tube defects in the offspring in a high-risk area of China. Public Health Nutr. 2009;12:680–686.

14. McMahon DM, Liu J, Zhang H, Torres ME, Best RG. Maternal obesity, folate intake, and neural tube defects in offspring. Birth Defects Res A Clin Mol Teratol. 2013;97:115–122.

15. Gilboa SM, Broussard CS, Devine OJ, et al. Influencing clinical practice regarding the use of antiepileptic medications during pregnancy: modeling the potential impact on the prevalences of spina bifida and cleft palate in the United States. Am J Med Genet C Semin Med Genet. 2011;157C:234–246.

16. Jentink J, Loane MA, Dolk H, et al. Valproic acid monotherapy in pregnancy and major congenital malformations. N Engl J Med. 2010;362:2185–2193.

17. Macfarlane A, Greenhalgh T. Sodium valproate in pregnancy: what are the risks and should we use a shared decision-making approach? BMC Pregnancy Childbirth. 2018;18:200.

18. Shaw GM, Velie EM, Schaffer D. Risk of neural tube defect-affected pregnancies among obese women. JAMA. 1996;275:1093–1096.

19. Stothard KJ, Tennant PW, Bell R, Rankin J. Maternal overweight and obesity and the risk of congenital anomalies: a systematic review and meta-analysis. JAMA. 2009;301:636–650.

20. Czeizel AE, Dudas I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. N Engl J Med. 1992;327:1832–1835.

21. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. MRC Vitamin Study Research Group. Lancet. 1991;338:131–137.

22. Bestwick JP, Huttly WJ, Morris JK, Wald NJ. Prevention of neural tube defects: a cross-sectional study of the uptake of folic acid supplementation in nearly half a million women. PLoS ONE. 2014;9: e89354.

23. Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Folic acid supplementation for the prevention of neural tube defects: US Preventive Services Task Force Recommendation Statement. JAMA. 2017;317:183–189.

24. Wald NJ, Law MR, Morris JK, Wald DS. Quantifying the effect of folic acid. Lancet. 2001;358:2069–2073.

25. Williams LJ, Mai CT, Edmonds LD, et al. Prevalence of spina bifida and anencephaly during the transition to mandatory folic acid fortification in the United States. Teratology. 2002;66:33–39.

26. Bergstrand CG, Czar B. Demonstration of a new protein fraction in serum from the human fetus. Scand J Clin Lab Invest. 1956;8:174.

27. Greenberg F, Faucett WA, Rose E. Congenital deficiency of alpha-fetoprotein. Pediatr Res. 1989;25:52A.

28. Haddow JE, Macri JN, Munson M. The amnion regulates movement of fetally derived alpha-fetoprotein into maternal blood. J Lab Clin Med. 1979;94:344–347.

29. Amniotic-fluid alpha-fetoprotein measurement in antenatal diagnosis of anencephaly and open spina bifida in early pregnancy. Second report of the U.K. Collaborative Study on Alpha-fetoprotein in Relation to Neural-tube Defects. Lancet. 1979;2:651–662.

30. Van den Hof MC, Nicolaides KH, Campbell J, Campbell S. Evaluation of the lemon and banana signs in one hundred thirty fetuses with open spina bifida. Am J Obstet Gynecol. 1990;162:322–327.

31. American College of Obstetricians and Gynecologists. Prenatal dection of neural tube defects. ACOG Technical Bulletin No. 67. October 1982. Washington DC.

32. American College of Medical Genetics and Genomics. Standards and guidelines for clinical genetics laboratories. 2008. https://www.acmg.net/PDFLibrary/General%20Policies.pdf. Accessed 1 October 2019.

33. Centers for Disease Control and Prevention. Impact of folic acid fortification in the U.S. 2018. https://www.cdc.gov/globalhealth/infographics/birth-defects/birth-defects-infographic.htm. Accessed 1 October 2019.

34. Food Fortification Initiative. Global progress of industrially milled cereal grain fortification. 2017. http://www.ffinetwork.org/global_progress/. Accessed 1 October 2019.

35. Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LY. Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. JAMA. 2001;285:2981–2986.

36. Committee on Practice Bulletins–Obstetrics. Practice bulletin no. 187: neural tube defects. Obstet Gynecol. 2017;130:e279–e290.

37. Crandall BF, Lebherz TB, Schroth PC, Matsumoto M. Alpha-fetoprotein concentrations in maternal serum: relation to race and body weight. Clin Chem. 1983;29:531–533.

38. Knight GJ, Palomaki GE. Epidemiologic monitoring of prenatal screening for neural tube defects and Down syndrome. Clin Lab Med. 2003;23:531–551.

39. Centers for Disease Control and Prevention. What is spina bifida? 2017. https://www.cdc.gov/ncbddd/spinabifida/facts.html. Accessed 1 October 2019.

40. March of Dimes. Neural tube defects. 2016. https://www.marchofdimes.org/complications/neural-tube-defects.aspx. Accessed 1 October 2018.

41. National Birth Defects Prevention Network. Neural Tube Defect Ascertainment Project. 2013. https://www.nbdpn.org/docs/NTD_Fact_Sheet_11-13_for_website.pdf. Accessed 1 October 2019.

42. Eunice Kennedy Shriver National Institute of Child Health and Human Development. How do health care providers diagnose neural tube defects (NTDs)? 2017. https://www.nichd.nih.gov/health/topics/ntds/conditioninfo/diagnosed. Accessed 1 October 2019.

43. Wald NJ, Cuckle HS, Densem JW, Kennard A, Smith D. Maternal serum screening for Down's syndrome: the effect of routine ultrasound scan determination of gestational age and adjustment for maternal weight. Br J Obstet Gynaecol. 1992;99:144–149.

44. Wald N, Cuckle H, Boreham J, Turnbull AC. Effect of estimating gestational age by ultrasound cephalometry on the specificity of alpha-fetoprotein screening for open neural-tube defects. Br J Obstet Gynaecol. 1982;89:1050–1053.

45. Chalouhi GE, Bernard JP, Benoist G, Nasr B, Ville Y, Salomon LJ. A comparison of first trimester measurements for prediction of delivery date. J Matern Fetal Neonatal Med. 2011;24:51–57.

46. Committee on Obstetric Practice, the American Institute of Ultrasound in Medicine, and the Society for Maternal-Fetal Medicine. Committee opinion no. 700: methods for estimating the due date. Obstet Gynecol. 2017;129:e150–154.

47. Roberts AB, Campbell S. Small biparietal diameter of fetuses with spina bifida: implications for antenatal screening. Br J Obstet Gynaecol. 1980;87:927–928.

48. Wald N, Cuckle H, Boreham J, Stirrat G. Small biparietal diameter of fetuses with spina bifida: implications for antenatal screening. Br J Obstet Gynaecol. 1980;87:219–221.

49. Haddow JE, Kloza EM, Knight GJ, Smith DE. Relation between maternal weight and serum alpha-fetoprotein concentration during the second trimester. Clin Chem. 1981;27:133–134.

50. Johnson AM, Palomaki GE, Haddow JE. The effect of adjusting maternal serum alpha-fetoprotein levels for maternal weight in pregnancies with fetal open spina bifida. A United States collaborative study. Am J Obstet Gynecol. 1990;163:9–11.

51. 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.

52. Waller DK, Mills JL, Simpson JL, et al. Are obese women at higher risk for producing malformed offspring? Am J Obstet Gynecol. 1994;170:541–548.

53. Watkins ML, Scanlon KS, Mulinare J, Khoury MJ. Is maternal obesity a risk factor for anencephaly and spina bifida? Epidemiology. 1996;7:507–512.

54. Werler MM, Louik C, Shapiro S, Mitchell AA. Prepregnant weight in relation to risk of neural tube defects. JAMA. 1996;275:1089–1092.

55. Feuchtbaum LB, Currier RJ, Riggle S, Roberson M, Lorey FW, Cunningham GC. Neural tube defect prevalence in California

(1990-1994): eliciting patterns by type of defect and maternal race/ethnicity. Genet Test. 1999;3:265–272.

56. Byrne JL, Waller DK, Rose E, Elias S. Effect of Hispanic ethnicity on interpretation of maternal serum screening. Fetal Diagn Ther. 1997;12:102–106.

57. Kucera J. Rate and type of congenital anomalies among offspring of diabetic women. J Reprod Med. 1971;7:73–82.

58. Mills JL, Baker L, Goldman AS. Malformations in infants of diabetic mothers occur before the seventh gestational week. Implications for treatment. Diabetes. 1979;28:292–293.

59. Wald NJ, Cuckle H, Boreham J, Stirrat GM, Turnbull AC. Maternal serum alpha-fetoprotein and diabetes mellitus. Br J Obstet Gynaecol. 1979;86:101–105.

60. Greene MF, Haddow JE, Palomaki GE, Knight GJ. Maternal serum alpha-fetoprotein levels in diabetic pregnancies. Lancet. 1988;2:345–346.

61. Evans MI, Harrison HH, O'Brien JE, et al. Correction for insulin-dependent diabetes in maternal serum alpha-fetoprotein testing has outlived its usefulness. Am J Obstet Gynecol. 2002;187:1084–1086.

62. Sancken U, Bartels I. Biochemical screening for chromosomal disorders and neural tube defects (NTD): is adjustment of maternal alpha-fetoprotein (AFP) still appropriate in insulin-dependent diabetes mellitus (IDDM)? Prenat Diagn. 2001;21:383–386.

63. Huttly W, Rudnicka A, Wald NJ. Second-trimester prenatal screening markers for Down syndrome in women with insulin-dependent diabetes mellitus. Prenat Diagn. 2004;24:804–807.

64. Sprawka N, Lambert-Messerlian G, Palomaki GE, Eklund EE, Canick JA. Adjustment of maternal serum alpha-fetoprotein levels in women with pregestational diabetes. Prenat Diagn. 2011;31:282–285.

65. Wald NJ, Cuckle HS, Boreham J. Alpha-fetoprotein screening for open spina bifida: effect of routine biparietal diameter measurement to estimate gestational age. Rev Epidemiol Sante Publique. 1984;32:62–69.

66. Harper PS. Prenatal diagnosis and reproductive aspects of medical genetics. In: Hodder A(ed.) Practical genetic counselling. Boca Raton, FL: CRC Press; 2010. p. 115–135.

67. Cuckle H, Wald N, Stevenson JD, et al. Maternal serum alpha-fetoprotein screening for open neural tube defects in twin pregnancies. Prenat Diagn. 1990;10:71–77.

68. Estimating an individual's risk of having a fetus with open spina bifida and the value of repeat alpha-fetoprotein testing. Fourth report of the UK collaborative study on alpha-fetoprotein in relation to neural tube defects. J Epidemiol Community Health1982;36:87–95.

69. Greenberg F, Faucett A, Rose E, et al. Congenital deficiency of alpha-fetoprotein. Am J Obstet Gynecol. 1992;167:509–511.