This guidance was written prior to the February 27, 1997 implementation of FDA's Good Guidance Practices, GGP's. It does not create or confer rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both. This guidance will be updated in the next revision to include the standard elements of GGP's.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Food and Drug Administration

Memorandum

Date · July 15, 1994

From Chief, Immunology Branch, Division of Clinical Laboratory Devices, Office of Device Evaluation, Center for Devices and Radiological Health.

Subject Review Criteria for Assessment of Alpha-Fetoprotein (AFP) in vitro Diagnostic Devices for Fetal Open Neural Tube Defects Using Immunological Test Methodologies.

To Interested Manufacturers

We have developed a draft document entitled, "Review Criteria for Assessment of Alpha-Fetoprotein (AFP) in vitro Diagnostic Device for Fetal Open Neural Tube Defects Using Immunological Test Methodologies". Since the document lists items we will be reviewing, it is intended to assist manufacturers in the preparation of marketing submissions for these types of devices. This document is also available from the Division of Small Manufacturers Assistance (DSMA), telephone 800-638-2041.

We are soliciting your ideas, recommendations, and comments regarding the attached review criteria. We will appreciate receiving your comments so that we can incorporate as many improvements as possible in a revision.

Please address comments to:

Peter E. Maxim, Ph.D.
Chief, Immunology Branch
Division of Clinical Laboratory Devices (HFZ 440)
Office of Device Evaluation
Center for Devices and Radiological Health
2098 Gaither Road
Rockville, MD 20850

Peter E. Maxim, Ph.D.

Attachment

Original version, July 1994

DRAFT

REVIEW CRITERIA FOR ASSESSMENT OF ALPHA-FETOPROTEIN (AFP) IN VITRO DIAGNOSTIC DEVICES FOR FETAL OPEN NEURAL TUBE DEFECTS USING IMMUNOLOGICAL TEST METHODOLOGIES

This is a flexible document representing the current major concerns and suggestions regarding premarket approval applications (PMA) for alpha-fetoprotein (AFP) in vitro diagnostic devices as an aid in detecting fetal open neural tube defects (ONTD). The devices addressed by this guidance document are those employing immunological test methodologies. It is based on 1) current basic science, 2) clinical experience, and 3) the Safe Medical Devices Act of 1990 and FDA regulations in the Code of Federal Regulations (CFR). As advances are made in science and medicine and changes in implementation of Congressional legislation, these review criteria will be re-evaluated and revised as necessary.

PURPOSE

This document provides guidance and clarification on information necessary before the Food and Drug Administration (FDA) may file and process a premarket approval application (PMA). While it is not to supersede the Code of Federal Regulations (CFR), this document provides additional guidance and clarification on what information is necessary before the FDA can issue an approval order for marketing such a device.

A PMA submission must show the device has clinical utility and there is reasonable assurance of its safety and effectiveness. A PMA must be submitted by all sponsors of AFP devices indicated for fetal open neural tube defects (ONTD). A PMA must provide evidence that the device is accurate, safe, and effective for its stated intended use. Safety and effectiveness information refers to information in the PMA submission, including adverse information that is relevant to an assessment of the performance characteristics of the device in the target population of maternal patients.

Following FDA approval, a Summary of Safety and Effectiveness document (SS&E) is released to the public. This is a mandatory requirement and should be a separate portion of the submission. The SS&E should not contain information the sponsor considers to be confidential or proprietary. A Federal Register announcement will announce the availability of the SS&E information to requesting persons to support FDA's decision that the device is safe and effective for its stated intended use.

The general content requirements for PMAs are set out at sections 515(c) (1) (A)-(G), 513(a) (3), 21 CFR 814.20 and 21 CFR 860.7. To help address the basic content requirements under 21 CFR 814.20, of the act, copies of the PMA Manual and PMA

Manufacturer's Supplement are available from the Office of Small Manufacturer's Assistance (800-638-2041). We believe these manuals, in conjunction with this Guidance Document, will enable the sponsor to submit a complete and well-organized application.

DEFINITION:

This generic type of device is intended for use in clinical laboratories, in response to a physician's prescription, as an aid in the detection of fetal open-neural tube defects (ONTD), to determine the level of AFP in maternal serum and amniotic fluid samples by an in vitro diagnostic test for quantitative measurement of AFP by various immunological test methodologies such as radioimmunoassay, enzyme immunoassay, and fluorescent immunoassay.[36]

PRODUCT CODE: 82-LOK

PANEL: Immunology Devices Panel

REVIEW REQUIRED: **PMA**

I. BACKGROUND:

There is no regulation for AFP devices intended for the detection of ONTD; nevertheless, FDA has the authority to regulate these devices under 21 CFR section 515(c) (1) (A)-(G), section 513(a) (3) of the act, also under 21 CFR 814.20 and 21 CFR 860.7.

An AFP test kit is a device that consists of the reagents used to quantitatively measure the concentration of AFP in maternal serum, plasma and in amniotic fluid by immunochemical techniques. Measurement of AFP in maternal specimens is an aid in the detection of fetal ONTD.

Significance of an increased or decreased level and medical and societal issues:

A National Conference on maternal serum alpha-fetoprotein (MSAFP)[13 14] was held in July 1980 in Washington, D.C. to discuss scientific, medical, ethical, legal, and economic issues associated with MSAFP testing for prenatal detection of ONTD. Potentially, all pregnant women might be offered this blood test as a way of detecting fetal ONTD, with diagnostic ultrasound and amniotic fluid AFP testing offered to women undergoing critical informed consent and genetic counseling sessions with their physicians. Although some individuals afflicted with open spina bifida lead productive and satisfying lives, this disorder is one of the most common serious congenital malformations: more than 2,000 pregnancies are affected with this condition each year in the United States (another 2,000 are affected with anencephaly). More than 95 percent of all ONTD occur among pregnant women whose

preceding offspring have not been affected. All of the infants born with anencephaly will die at or shortly after birth. The life expectancy for about 1,400 of the infants born with open spina bifida is at least 5 years. The majority will experience significant physical and mental handicaps. Identifying open spina bifida prenatally allows the family to choose between terminating or continuing the pregnancy. When the latter choice is made, the family and physician can prepare for the birth of an affected child. This advance notice permits these women to have their babies in centers that can offer surgical, medical, and other care needed to minimize the infants' disability. When anencephaly is identified, nearly all families opt to terminate the pregnancy. In addition, elevated MSAFP levels may help to identify pregnancies at higher risk for perinatal complications[15-17] and also about 50 percent of twin or higher multiple-birth pregnancies.[18 19 72]

A. Introduction:

1. Clinical indication and clinical significance of the intended use:

The two major types of fetal ONTD are anencephaly and open spina bifida. The value of AFP in prenatal testing for ONTD was first suggested in 1972 with a report by Brock and Sutcliffe[1], who documented that amniotic fluid AFP levels were increased in the presence of ONTD. Following confirmation of this discovery, measurement of amniotic fluid AFP levels rapidly became part of second trimester prenatal maternal care. At that time, such testing was reserved for women who had already borne a child with a neural tube defect, since these women are at increased risk for having another similarly affected pregnancy. AFP analyses, nevertheless, soon began to be performed on amniotic fluid samples being processed primarily for other diagnostic purposes (e.g., cell culture for chromosomal analysis).[6 7] Although amniotic fluid AFP measurements proved to be clinically useful as a prelude to genetic counseling, occasional false-positive results led to the need for continuing re-appraisal, including repeat amniocentesis in some cases. Fetal blood contamination of the amniotic fluid sample was found to be the most frequent explanation of such false positives. Other fetal lesions (e.g., open ventral wall defects and the Finnish type congenital nephrosis) were also found to be associated with elevated amniotic fluid AFP concentrations, diminishing the diagnostic specificity of the test but increasing the range of identifiable fetal problems.[6] Requirements for timely sample collection (gestational weeks of pregnancy) and handling, as well as methodology were such that existing prenatal testing facilities in the United States were able to develop ultrasonography dating of pregnancy and subsequent amniotic fluid AFP testing capabilities with relative ease.[37]

2. Role in normal pregnancies:

The biological basis for differing prevalence of ONTD for different populations and geographic regions is largely unknown at this time; nevertheless, each sponsor of a PMA for ONTD must explain to each of the clinical investigators that these factors are an important consideration in the accurate interpretation of MSAFP results. When physicians and clinical laboratorians are educated to accurately interpret MSAFP (within the context of recommended testing protocols[36]), AFP testing is not different from any other clinical laboratory procedure whose goal is to improve the quality of maternal health care.

3. Background description:

Provide a background summary of the disease including basic science such as biochemistry, physiology, vitamin deficiency, pathophysiology, and pharmacology.

4. Etiology of Disease:

Provide a description of the etiology for fetal ONTD, for example:

There is considerable evidence that the periconceptual administration of multivitamin supplements[72 73] or a folic acid supplement alone[74] prevents recurrent neural-tube defects. However, about 95 percent of the women who deliver infants with neural-tube defects have not previously delivered infants with these defects. A recent report by Czeizel and Dudas[75] demonstrated that periconceptual vitamin use decreases the incidence of a first occurrence of neural-tube defects. The manufacturer or sponsor should provide an updated literature review of the most recent studies available.

Historically, Leek, Ruoss, Kitau, Chard[2] and Brock, Bolton, and Monaghan[3] independently demonstrated increased maternal serum AFP (MSAFP) levels in the presence of fetal anencephaly. In 1974, Wald, Brock, and Bonnar[4] and Brock and Scrimgeour[5] presented data showing that MSAFP levels were also increased when the fetus was affected by open spina bifida. Strengths and limitations of the testing protocol were investigated and addressed by numerous centers during the ensuing years, and it became clear that MSAFP measurement could be applied usefully as a routine technique for improved maternal health care.[12 69 70] Two major United Kingdom collaborative studies, one completed in 1977 addressing MSAFP testing[4], and the other in 1979 involving measurement of amniotic fluid levels of AFP, established the overall reliability of both analytic processes and formed a general basis for a testing protocol that allowed estimating detection rates and false positive rates when applied to the maternal population within the United States.[7 66 67 68] The application of AFP measurements to

the detection of ONTD was comprehensively evaluated by several manufacturers of devices that quantitatively measure AFP levels in maternal specimens. Measurement of AFP in maternal serum and amniotic fluid samples, following FDA[42] prescribed testing protocol, aids in the detection of fetal ONTD.[65 66 36]

Table 1 summarizes the FDA approved devices for fetal ONTD. Three international conferences (1977, 1978, and 1980) in Scarborough, Maine, as a prelude to the introduction of MSAFP testing in this country, coincided with FDA's evaluation of ***in vitro*** devices for MSAFP testing. Haddow and Macri[11] and Wald and Cuckle[12] also reviewed the current state of knowledge concerning practical applications of MSAFP testing.

Table 1. ALPHA-FETOPROTEIN DEVICES FOR WHICH THERE IS A FDA-APPROVED PMA FOR USE AS AN AID IN THE DETECTION OF FETAL OPEN NEURAL-TUBE DEFECT

SERUM MARKER	MANUFACTURER/DEVICE	TYPE	FDA APPROVAL	PMA NUMBER
AFP	ABBOTT/DAIN-ABOTT	RIA	6/25/84	P820060
AFP	WAMPOLE/AFP-TEST™	FARR TECHNIQUE	1/13/84	P760001
AFP	KALLESTAD/QUANTITOPE™ [^{125} I-AFP] (Sanofi Diagnostics Pasteur)	RIA	1/13/84	P800025
AFP	TRAVENOL-GENENTECH/ GAMMADAB™ [^{125}I-AFP]	RIA	4/30/86	P790032
AFP	HYBRITECH/TANDEM-E AFP	IMMUNO-ENZY-METRIC	4/29/88	P840035
AFP	AMERSHAM/AMERLEX (EASTMAN-KODAK, USA)	RIA	4/07/84	P780006

5. Significance of false positive and false negative results:

When a manufacturer initiates clinical studies in support of a PMA for a device intended as an aid in the detection of fetal ONTD, the MSAFP testing program should be implemented according to guidance offered by FDA. It is FDA's view that it is necessary at the outset to assist in the education of physicians, genetic counselors, nurses, laboratory staff, and the pregnancy age population with brochures and device labeling for interpreting laboratory results. FDA believes that the testing of pregnant women for fetal ONTD should not be referred to as

"screening" for fetal ONTD. Screening implies that all pregnant women should undergo AFP testing, and this is not the position of FDA.

II. DEVICE DESCRIPTION:

Key issues in the review of a new AFP device are the specific intended use, the specimen source tested, and the technology utilized. The following descriptive information must be included in a submission to adequately characterize the new in vitro device for detection of AFP in clinical specimens. Appropriate peer-reviewed literature references that relate to the technology of the device must be submitted, in addition to the descriptive information requested below, to adequately describe the new in vitro device.

A. Intended Use:

An AFP test system is a device that consists of the reagents used to measure by immunochemical techniques the AFP in maternal serum and amniotic fluid samples. Measurement of AFP in maternal serum and amniotic fluid samples, following a FDA prescribed testing protocol, aids in the detection of fetal ONTD.[36]

B. Principle of Operation:

An AFP Immunological Test System may use the following general principles, which vary according to methodology employed. The principle of the test methodology requires adequate discussion. This document addresses some of the difficulties surrounding test system assessment, such as assuring that reagents are immunochemically specific to AFP. Due to the stability of the molecule, AFP has always been measured by some form of immunoassay. Early studies quantitated AFP in amniotic fluid using electroimmunodiffusion (rocket electrophoresis), a type of precipitation immunoassay employing AFP-specific antibodies immobilized in a gel. Rocket electrophoresis is still occasionally used to measure AFP in amniotic fluid. This technique is not sufficiently sensitive to measure the lower levels of AFP found in maternal serum. Most laboratories now use radioimmunoassay (RIA) or enzyme immunoassay (EIA). Several radioimmunoassay (RIA) and enzyme-linked immunosorbent assay (ELISA) devices and reagents are approved and their manufacturing monitored by the Food and Drug Administration (FDA). These methods are capable of quantitating AFP in both maternal serum and amniotic fluid. In practice, a test kit must be capable of measuring AFP values reliably between 10 and 150 IU/mL.

Antigen-antibody binding systems can be divided into two types: competitive and noncompetitive assays. In the competitive assay a stoichiometric relationship is established between the

monospecific anti-AFP reagent and a specific amount of labelled (e.g., radioiodinated AFP standard). The extent to which a sample containing unknown levels of AFP competes with labeled AFP is compared, and the concentration of AFP is read from a standard curve. The success of this type of competitive binding procedure depends on the separation of bound and free labelled AFP. This is usually accomplished by the addition of a second monospecific anti-globulin reagent which precipitates the immunoglobulin: monospecific anti-immunoglobulin reagent complex (double antibody radioimmunoassay). Additional aspects of this technique are addressed later in this guidance document: for example, purity, specificity and affinity binding of antibody, antigen purity, pH of reaction mixture, protein concentrations.

The competitive assay (as well as chemiluminescence), employing monoclonal antibodies, has also been adapted for use with a solid phase system, immunoradiometric assay (IRMA). The AFP in the test system competes with a fixed amount of radiolabelled (or light-emitting reagent) AFP for a limited number of binding sites on the immobilized solid support containing the anti-AFP reagent.

When the competitive binding assays utilize radioisotopes, they are classified as radioimmunoassay; however, it is also possible to substitute fluorochrome labelled anti-immunoglobulin reagents (fluoroimmunoassay) or hybridoma derived enzyme conjugated anti-immunoglobulin reagents. These reagents allow the estimation of the analyte concentration by monitoring colorimetric reactions upon addition of the appropriate substrate with an appropriate spectrophotometer (enzyme immunoassay). Under those assay conditions, binding systems amplified by radiolabelled (radioimmunoassay) or enzyme-linked reagents (enzymeimmunoassay) offer the required sensitivity. Substrate/Indicator reagents used in AFP devices must be described, including pH values (e.g., fluorogenic substrates and others).

An overview of the methods for the quantitative measurement of human AFP in serum and amniotic fluid indicates techniques are becoming increasingly sensitive and accurate due to improved instrumentation, standards, monoclonal antibody reagents and knowledge of interfering substances. Immunochemical, bioluminescence and chemiluminescence methods for the quantitation and characterization of human AFP are rapidly becoming the methods of choice in automated medical devices. These methodological innovations result in improved precision and overall reliability in estimating AFP MOM values. The limitations connected with the newer and innovative methodologies should be narratively explored and compared to the older technology.

III. NONCLINICAL LABORATORY STUDIES: SPECIFIC PERFORMANCE CHARACTERISTICS:

FDA requests different types and amounts of data and results of statistical analyses in PMA applications seeking approval for a Class III (premarket approval application). The amount and type of data requested depends on: 1) the test analyte, 2) the intended use (which determines whether the application is a PMA, and 3) the technological characteristics of the device.

A. Sampling and Data Collection Design

Factors That Influence AFP Measurements During Pregnancy:

The manufacturer or sponsor of AFP devices should provide patient information, laboratory data, and statistical evaluation of AFP results due to the changes of the average AFP concentration in maternal serum as well as amniotic fluid throughout gestation.

During each week of gestation there is overlap in the distribution of AFP concentrations between unaffected pregnancies and pregnancies affected by ONTD. This is particularly true for maternal sera and, to a much lesser extent for amniotic fluid. Consequently, whatever AFP cut-off is chosen, that cut-off carries a definable false-positive rate, as well as, detection rate.

1. Maternal weight influences the MSAFP (but not amniotic fluid AFP) concentration. This is due to a dilutional effect as the blood volume increases. Black women have MSAFP measurements that are about 10 to 15% higher than other racial and ethnic groups in the pregnancy population. This difference is independent of weight or other known variables. Insulin-dependent diabetes in a pregnant woman is associated with lower MSAFP levels (by about 20%). FDA recommends that all of these factors be taken into account when generating tables of data to be used for interpreting MSAFP test results by adjusting the AFP values. Median values of MSAFP test results, when tabulated by gestational week, may be interpreted by comparing the patient's AFP value to the multiple of the unaffected population median values at a given gestational age. FDA expects sponsors of AFP kits to narratively instruct physicians and genetic counselors in understanding and interpreting MSAFP values.

2. Maternal age, parity, geographic factors and factors associated with seasonal variation are variables that may have an influence on AFP level distribution in serum and amniotic fluid samples from affected pregnancies; however, these variables are known not to influence distributions of serum or amniotic fluid AFP measurements in unaffected pregnancies.

3. Methods of Dating Gestational Age: Sponsors of AFP devices should be aware there are two conventions presently in use for defining which gestational week to assign a given serum sample

sent for MSAFP measurement. In the first, completed weeks determine assignment of gestational age (e.g., 17 weeks and 4 days = 17 weeks). In the second, gestational age is rounded to the nearest week (e.g., 17 weeks and 4 days = 18 weeks). Results from FDA mandated Post-Approval Studies and from proficiency testing programs in the United States indicate that completed weeks of gestation are used by almost all laboratories. FDA makes no formal recommendation which method to adopt at a particular testing laboratory other than acknowledging that completed weeks of gestation may be used to achieve consistency in interpreting MSAFP AFP levels.

The gestational age is traditionally estimated by counting from the first day of the last menstrual period (LMP). About 40% of the specimens with an initially elevated MSAFP level are subsequently discovered after ultrasound dating to be due to underestimated gestational age. Furthermore, it is known that when the LMP method is used a significant proportion of pregnancy dates are overestimated. This can lead to false-negative MSAFP testing results. Routine ultrasonography is recommended by FDA as the next step in the testing protocol to (a) corroborate gestational dates, (b) illuminate twins, (c) evaluate the viability of the fetus, and (d) possibly detect fetal anencephaly. If the elevated MSAFP value remains unexplained, level 2 or gray-scale ultrasonography is performed to search for other fetal malformations, most prominently open spina bifida and open ventral wall defects. Following genetic counseling with the maternal patient's physician, amniocentesis for the purpose of measuring amniotic fluid AFP is also offered.

4. Timeliness of Reporting: For a manufacturer's clinical trials coordinator, testing site investigators should be educated to the necessity for timely reporting of test results. All AFP assays need to be performed and reported promptly, because elevated test results prompt physicians to adhere to a testing protocol of diagnostic procedures that must be completed within a short time span. If the gestational dating is correct, a specific risk of neural tube defect or ventral wall defect can be calculated from the MOM, rather than relying on a specific cut-off for the MOM, such as 2.5 or 2.0 multiples of the normative MOM. This method or paradigm uses a known prevalence of ONTD within the population, the distributions of concentrations of MSAFP from both the affected and unaffected pregnancies, and Bayes' theorem.[61] Whether a MOM cut-off or a specific risk calculation is used, the MOM must be adjusted for maternal weight (volume of distribution); race (multiply MOM by calculated factor for each racial group for which normative data is available);[62] [63] and diabetes (divide MOM by factor derived from population specific AFP values).[64] Using the specific risk-based method proposed by Adams'[62] group, risks of undiagnosed twins and ventral wall defects (VWD) may be calculated from the MOM. Risks may then be combined, and a total risk for combination of either

VWD or ONTD may be quoted. The reliability of this paradigm has been questioned recently by Bishop [76] and therefore must be verified by prospectively developed data (100 maternal patients for each gestational week for which the device is recommended) at each of three geographically diverse testing centers. Computer-generated risk factors must be verified in like manner. For an excellent general discussion of the predictive value and clinical efficiency of medical diagnosis from the laboratory, see also Galen and Gambino[25] .

5. Normative Data: Each manufacturer or sponsor of an AFP device should instruct its clinical studies coordinator that the median MSAFP, calculated from healthy women who have uncomplicated pregnancies with singleton outcomes, increases approximately 15% per gestational week during the second trimester.[51 52 64] From clinical information and AFP determinations, this should be validated during data analyses following data collection and verification.

About 90 percent of sera from unaffected singleton pregnancies at a given gestational week have AFP values varying from half to double the median value (about 35 IU/mL at completed weeks). The distribution of MSAFP values during each gestational week results in a skewing of the data when plotting AFP values arithmetically, with a significant number of "outlier" values beyond the upper end of a bell-shaped curve. The observed distribution should be found to be approximately log-normal.

6. Quality Control: Reagents (calibrators, controls) must be identified and instructions for use should be included in a separate section of the labeling. Standardization of AFP measuring systems is influenced by three major factors: 1) the AFP preparation used for calibration, 2) the anti-AFP reagent, and 3) the limitations of the given assay procedure. The advent of hybridoma-derived and bacteria-gene splicing derived monoclonal antibodies has overcome some of the difficulties in the standardization of reference preparations. Nevertheless, monoclonal antibodies do not precipitate in agar, making the demonstration of specificity more difficult, and many recognize epitopes of specific AFPs that are not necessarily shared by all AFP proteins in a particular reference preparation.[76] When AFPs are measured by radial immunodiffusion in agar utilizing precipitating antisera, a particular group of epitopes are detected. Thus, different AFP assays measure different epitopes of the AFP molecule, resulting in different MoM values of AFP in maternal samples. Using radioimmunoassays, as well as solid-phase radiolabeled and enzyme-labeled immunoassays (ELISA), AFPs may be calibrated against several standards. All calibrators and control sera should be run in duplicate in order to minimize random errors.

BRITISH STANDARD

The First British Standard for Human Cord Serum (72/227) was produced from the same batch of cord serum as the WHO material but was lyophilized by a different technique. The British reference material is calibrated against WHO (72/225).

WHO STANDARD

The World Health Organization (WHO) Standard for Human AFP (72/225) is a lyophilized cord serum preparation that is available on a limited basis. In this preparation, International Units (IU) have been collaboratively assigned.[26-29]

U.S. REFERENCE MATERIAL

The U.S. National Reference Material for AFP in Mid-Pregnancy Maternal Serum is a pool of hepatitis-free human normal adult serum, spiked with cord sera to contain 392.8 IU of AFP/mL. This sterile-filtered, lyophilized pool was packaged in 10,000 x 0.5 mL vaccine vials. Six of these vials may be obtained every six months at no cost by testing programs with a workload of at least 50 patient specimens per week.

Composition of Proposed U.S. National Reference Materials

For manufacturers or sponsors of AFP devices, two body fluids, maternal serum and amniotic fluid should be analyzed for AFP content during pregnancy. Stable calibrators traceable to the WHO (72/225) standard that stimulate the analytical conditions defined by the composition of these fluids are desirable.

7. Diluent: AFP content in amniotic fluid samples is generally measured using high sensitivity assays designed to measure the lower AFP levels found in maternal serum. This requires an initial dilution step (1/100 to 1/200) for amniotic fluid samples. Most manufacturers either supply diluent with their devices or sell diluent separately.

If a diluent other than that recommended or supplied by the manufacturer is used, the laboratory must show by direct comparison that the alternative diluent is suitable.

Caution: Manufacturers or sponsors of AFP devices should guard against introducing inadvertent claims (for additional intended uses and other associated test procedures). This may be avoided by thoroughly editing the narrative text of the device's description, evaluating labeling, and product brochures. FDA may require the manufacturer to modify or entirely delete certain additional intended use claims that are clinically inappropriate or constitute an implied claim (see 21 CFR 801.4).

B. Analytical/Laboratory In Vitro Studies

1. Clinical Performance of Manufacturer's Test Kits Submitted For Approval To FDA:

a. Purity and homogeneity of AFP antigen

The preparation of immunogens represents the first step in antiserum preparation. Only human AFP is to be used as immunogen. AFP isolated from amniotic fluid AFP is the preferred source (rather than AFP from the sera of patients with AFP producing hepatomas). The manufacturer should demonstrate antigen purity or characterize any impurities present and justify why they will not interfere with the assay.

b. Purity and specificity of AFP antiserum

Assessment of the specificity of polyclonal antisera should be performed by sensitive methods, such as crossed two-dimensional immunoelectrophoresis[39] or immunoblotting. The use of less sensitive and less precise methods, such as immunoelectrophoresis and immunodiffusion, is not acceptable.

Reagents that comprise these devices are increasingly derived from hybridomas. If anti-AFP hybridoma antibodies are used in the device, the manufacturer must submit a summary of characterization data including the following information:

i. identification of source of parent myeloma cell.

ii. source of antibody (mouse, etc.)

iii. antibody characterization

iv. description of cloning and criteria used for selection.

v. stability data (real-time studies)

vi. precision, accuracy, reproducibility, and linearity data.

vii. summary of data demonstrating lot-to-lot consistency (3 separate lots)

viii. sensitivity, specificity, cross-reactivity and interference.

ix. comparisons [Demonstrate the antiserum specificity using assay of comparable sensitivity to the submitted device.]

x. cross-reactivity [Test anti-serum for the possible cross-reactivity to other normally occurring pregnancy-associated proteins.]

xi. strength (affinity binding constant) [Provide the titer or affinity binding constant for the anti-serum as appropriate.]

TEST REAGENTS AND TEST DEVICE SET-UP

The sponsor must provide the source of the antisera, either polyclonal or hybridoma derived, and sufficient information about the dilution of the antisera, filtration used and the standard curves generated from the antisera for each lot of AFP assayed.

The frequency of preparing standard curves for each instrument described should be stated (e.g., standards run every 40 samples). Description of a pooled sera Quality Control sample should be provided by the sponsor and the frequency of use provided (e.g., run every 10 samples). Blanks or samples without serum should be run on the entire test series. Parameters that should be discussed include the entire reagent set-up, washing procedures, diluting solution used, and the final dilution of serum samples examined for immunoglobulin concentration. Also look for a description of active and non-reactive ingredients (carriers). The reference preparation for each lot of AFP should be identified (e.g., US National Reference Preparation for AFP from Centers for Disease Control), the construction of standard curves should be described, and the procedures for fitting the assay points to a graph explained (e.g., calculation of results using a third order polynomial curve fitting procedure were performed).

2. Performance Characteristics of the Device

a. Analytical Sensitivity (Detection Limit):

For Immunoassays (radioimmunoassays, enzyme immunoassays, etc.):

The analytical sensitivity is defined as the lowest quantity differentiated from Zero (95% confidence intervals or 2 standard deviations are commonly used)[53][54]. Run the Zero standard (Zero diluent) at least 25 times in the same run and calculate the mean of the Zero Standard and two standard deviations (SD) of the mean (counts, optical densities, etc.).

b. Linear Range:

Validate the linear range of the assay with normal and abnormal specimens covering the whole range of the standard curve.[55]

c. Accuracy/Recovery Studies:

Determine the spiking recovery[53 54 56 57] of known amounts of AFP. Analytical recovery studies involve analyses after known amounts of AFP are added to the maternal fluid on which the determination will be performed. Include samples from patients with grossly abnormal and pathological conditions, as well as normal samples from healthy individuals. Calculate the percent recovery by the formula: amount added-amount present/amount added x 100.[58]

d. Dilution/Parallelism Studies:

In certain in vitro assays, different antigen preparations or specimen types may yield different dose-response curves. Statistically compare the slope of the calibration curve to the slopes of dilutions of AFP containing patient samples with elevated analyte levels for all claimed sample types. These maternal patients should represent a variety of ethnic origins. Alternatively, add a statement to the Limitations Section of the package insert that the device has not been tested for cross-reactivity or interference of one or more substances.

Hybridoma-derived monoclonal antibodies should be tested for specificity by competitive inhibition techniques, using radioimmunoassay and ELISA or by immunoblotting. The binding properties of the selected antibodies should be well characterized to be sure that the epitope(s) reacting in a particular assay are expressed on all particles of a given type in all serum and amniotic samples.

3. Specificity/Cross-Reactivity/Interference Studies:

The testing protocol should describe the methods used to determine the amount of assay interference contributed by other serum proteins, anticoagulants, other drugs or chemicals or differences in assay temperature, time for assay incubation, washing of reactants, elapsed time of measurement of analyte and time of adding solutions that halt the assay procedure.

a. Evaluate drugs or their metabolites that might be commonly used by patients tested for interference with test system that quantitates AFP. Alternatively, add a statement to the Limitations Section of the package insert that the device has not been tested for cross-reactivity or interference.

b. Other endogenous substances:

Evaluate potentially cross-reacting endogenous substances at high concentrations including common serum components, such as lipids, hemoglobin, bilirubin, etc.[53 54 56 58]

4. Reproducibility and repeatability[53 54 56 57 59 80]

a. Laboratory Use.

The National Committee for Clinical Laboratory Standards (NCCLS) recommends[59] an analysis of variance experiment testing two clinically significant levels near medical decision limits (subnormal, normal, or elevated) of an analyte, in this case. Use controls simulating patient samples or actual patient samples or actual patient specimens 2 times in the same run and in two different runs in the same day for 20 days. This permits separate estimation of between-day, between-run and within-day standard deviations (SD), as well as within-run and total SDs. Acceptable alternatives that include only one run per day are also discussed in the cited document[59]. The three important assumptions (homogeneity of error variances, additivity, and normality) of analysis of variance need to be demonstrated to validate the above results.

i. Quantitative tests:

Calculate total, between-day and within-day, and between-run and within-day, and between-run and within-run means and coefficients of variation of imprecision for each set of values.[60]

Report in the performance characteristics section of the package insert the appropriate means, SDs and/or coefficients of variation with confidence levels according to number of times the sample is repeated. Report the number of runs per day.

a). Provide between-laboratory assay variability from three different laboratories.

b). Provide lot-to-lot assay variability.

5. Describe all protocols for testing.

a. Statistical Methods for Expressing MSAFP Results

i) Multiple of the Median (MOM):

The United Kingdom Collaborative Studies [6] [7] express individual AFP values as a MOM AFP value, calculated from unaffected pregnancies of the same gestational week. Manufacturers and sponsors of AFP devices should narratively explain to the clinical investigators and their own clinical trials coordinators that this convention was developed originally as a method of normalizing data from testing centers with differing AFP standards, due to lack of common calibrator. It subsequently became apparent that the MOM automatically compensates for gestational age-related changes in MSAFP and amniotic fluid AFP values, thereby providing a common currency for interpreting test results. The MOM value can also be readily adjusted to take into account the other variables that affect the interpretation of AFP results (e.g., maternal weight, insulin dependent diabetes, race). The MOM value may also be used to calculate individual patient-specific risks. The College of American Pathologists Laboratory Accreditation Program for Prenatal Screening with AFP requires the following with regard to the MOM: "The sponsor's clinical trials coordinator should have documentation that the laboratory established its own normal median AFP values and verified that the manufacturer's package insert or other sources for the population being tested are updated at least annually."

(ii) Arithmetic Units:

Other methods of normalizing AFP test results, such as means and standard deviations or percentiles, are much less robust and are not recommended because the skewed distributions of AFP values require special attention to defining percentiles of normal versus the total population.

b. Statistical data:

Provide data and statistical analyses determined with the device to support performance parameters specific to and important for operating the device (e.g., reproducibility).

c. Justification of statistical methods:

Avoid sole reliance on statistical hypothesis testing (such as use of **p** values) which fails to give important quantitative information. Give the working data, statistical methods used with justified assumptions, test statistic results and the corresponding **p**-values. The use of specific statistical methods must be fully justified (e.g., parametric vs. non-parametric).

d. Statistical methods:

Discuss the statistical methods for the type of data submitted (e.g., quantitative continuous data, qualitative discrete data, etc.) and the distribution of data (normal vs. non-normal type of data (paired vs. independent).

e. Statistical references:

Use references for study design and statistical methods that are from standard texts and/or refereed biomedical journals.

f. Test data:

Present test data with estimates of error, the estimated parameters and the confidence intervals, analyses for inference, and conclusions.

g. Electronic media (floppy disks)

If possible, provide copies of raw data and the proposed data "lay-out" and intermediate clinical information and clinical data in specified format on electronic media (floppy disks). The Summary of Safety and Effectiveness document is to be processed with Microsoft's WordPerfect™ version 5.1 without right justified margins. For additional information contact Office of Information Systems (OIS), phone (301-594-4550).

Additional performance characteristics may be required if clinical categorization of patients or microprocessor-controlled diagnostic interpretations (databases of patient information) are

performed and presented to the end-user with different or selected (mathematical adjustments to AFP MOM values to accommodate different maternal weights, racial distribution within a particular geographic locale) graphical and numerical report forms. In these submissions, manufacturers must adhere to the requirements listed for High Level of Concern in Reviewer Guidance for Computer Controlled Medical Devices Undergoing 510(k) Review. This guideline is available from the FDA Division of Small Manufacturers Assistance (DSMA), 1-800-638-2041.

The impact of Hospital or Institutional Committee on Human Investigation upon the duration of the study should be discussed if aspects of patient testing involve placing the patient at risk (e.g., amniocentesis performed to obtain amniotic fluid sample).

Reportable Dynamic Range:[3 4 6 7 9] Assay Optimization: the working standard curve for an AFP immunoassay should be broad enough to include most normal and pathological specimens at the time in pregnancy when testing is being carried out [FDA recommended (15 through 20 weeks of gestation)]. In practice, this requires a curve which spans the range from 0.5 to 8.0 multiples of the median for unaffected pregnancies, corresponding to values from approximately 10 to 200 IU/mL. Assays should, therefore, be linear over this range. Specimens found to be more concentrated than the upper range can be diluted for reassay.

The proportions of total AFP population variance which are attributable to assay imprecision for between-assay coefficient of variation (CV) of 20%, 10%, 5%, and 3% are 24%, 6%, 1.5% and 0.54% respectively.[32 80] Imprecision is thus a minor source of variability in MSAFP testing when between-assay CVs are 10% or lower in the range of values falling on and around the clinically important decision points (2.0 to 2.5 MOM).

6. Quality Control: All working curves (calibration), quality controls and patient samples are traditionally run in duplicate. Each assay run should contain at least three randomly placed quality control samples, processed in an identical manner to patient specimens. These three quality control samples should have concentrations near the normal median at 17 gestational weeks and also near 2.5 and near 1.0 times this normal median. Control limits for the working curve parameters and quality control samples (usually +/- 2 standard deviations from the mean) should be recorded, and individual patient samples should be rerun if the statistical calculations and graphic results from standards and controls fall outside acceptable limits.

IV. CLINICAL STUDIES: SPECIFIC REQUIREMENTS

A. Plan clinical studies to prove clinical utility and safety and effectiveness.

1. Prove all diagnostic claims and specific parameters important for the clinical utility of the device

2. Each manufacturer must have three testing sites and educate each of the investigators to the dangers in the use of incorrect reference data. This is a common cause of incorrect interpretation of MSAFP testing measurements by physicians and laboratorians, it is essential that the manufacturer monitor each of its testing laboratories to establish the testing laboratories' own reference data and/or demonstrate that reference data obtained from another testing laboratory are valid for the population being tested. Reference values in MSAFP testing consist of a set of median values (pooled from the three testing sites) calculated for each week of gestation using the pooled laboratories MSAFP assay values, measured preferably on the population to be tested with the device. Individual MSAFP test results are then expressed as multiples of the unaffected population median (MOM), which is obtained by dividing each individual MSAFP value by the median value for the relevant gestational week. Strategies are outlined below for establishing median values for both maternal serum and amniotic fluid.

3. Number of investigators:

   Use at least three independent investigators at separate (geographically diverse) sites with at least one in the United States.

4. Sample size:

   Plan a sample size that will be statistically sufficient to determine whether or not the device is safe and effective prior to beginning the clinical trial.

5. Sample type:

   Include data to support use of the test with all claimed specimen matrices (clinical and analytical). Due to interference of anticoagulants with assay results, plasma is not a recommended specimen.

6. Sensitivity and Specificity:

   Regardless of the sample size, present the device's diagnostic sensitivity (true positives) and its specificity [25] (true negatives) and their 95% confidence intervals in the Performance Characteristic section of the product package insert.

7. Sampling method:

   Describe sampling method used in the selection and exclusion of patients. All statistical analysis is based on the "random sample" assumption (e.g., probability sampling).

8. Pooling of investigator's data:

   Present clinical information and data with analyses and conclusions by each investigator. Additionally, present pooled data from each of the investigators, if statistically and clinically justified.

9. Describe statistical methods used and provide confidence intervals.

10. Representative data:

    Present the clinical information and data from the targeted population which must represent the gestational ages for which the device is intended.

11. Establishing reference median values in maternal serum:

    Present validated median values to the population being recommended for testing because of factors such as variability in reliability of gestational dating. Proficiency testing programs find that MSAFP measurements vary by as much as 15% between laboratories, even when results are expressed in IU/mL.[80] The major contributing factor to explain these differences is a bias among manufactured AFP devices. The use of median values obtained from published sources, as reference data is therefore, contraindicated.

12. Establishing reference median values in amniotic fluid:

    Obtaining sufficient numbers of amniotic fluid samples may be difficult for manufacturers. For example, a manufacturer may test 3,000 women per year, identifying only 2 to 4% of these women as candidates for amniocentesis. Amniotic fluid samples sent for AFP

analysis by cytogenetic laboratories may be used to supplement those obtained via the AFP testing program, since nearly all such specimens will be from unaffected pregnancies. FDA recommends that for each gestational week for which the kit is recommended, 50 amniotic fluid samples be used in calculating the median used for interpreting amniotic fluid samples from a particular maternal patient.

Amniotic fluid AFP values are expressed as multiples of the median in the same way as described for maternal serum. The relationship of the median values and gestational age is also log-linear for gestational weeks 15-21 (gestational weeks recommended for testing by FDA), but amniotic fluid AFP median values decline rather than increase with gestational age.

B. Monitor Clinical Trials Testing

1. It is important for manufacturers who sponsor clinical investigations to monitor the clinical performance of their testing programs by tracking the number of fetal malformations detected and missed. Affected pregnancies detected must be followed to the birth of the child or elective termination of the pregnancy. Monitoring the initial percentage of women with elevated (positive test results) AFP levels is a portion of FDA required clinical studies (3 separate testing sites of 1000 maternal patients each at geographically diverse testing sites). FDA's rationale for this epidemiologic surveillance is that the initial percentage of women with positive test results affords the clinical trials investigator an opportunity to properly offer genetic counseling prior to the offering of amniocentesis. FDA recommends that successive diagnostic modalities of ultrasonography dating of the fetus, correction of the AFP MOM interpretation (positive test result or negative), including interpretation of AFP levels from amniotic fluid samples and confirmatory testing for the presence of neural acetylcholinesterase must be accomplished with genetic counseling of the maternal patient. As an example, a clinical investigator's testing site laboratory whose MSAFP testing cut-off is at 2.0 MOM would have an initial positive test result rate of 3 to 5%. If the cut-off is at 2.5 MOM as recommended by FDA, the initial positive test result rate should be 1 to 3%. Manufacturers must followup all positive maternal patients with elevated AFP tests with genetic counseling and other maternal care modalities at each of the three testing sites.

2. The manufacturer should be aware that the initial positive test result is very sensitive to changes in precision and accuracy of the AFP assay, long term assay drift, and inappropriate normative reference data (AFP median values).[76] From AFP Post-Approval Studies, FDA found that epidemiologic

monitoring was a powerful addition to traditional quality control procedures and should be an integral part of the manufacturer's clinical testing program. A preliminary report of these studies was presented by Hybritech, Inc., Division of Eli Lilly Company.[79]

For manufacturers and sponsors of AFP test kits, the testing protocol previously referred to and recommended by FDA was published in the November 7,1980 FEDERAL REGISTER[38]. A clarification of the testing protocol and the importance of adequate followup to positive serum AFP tests was published in Clinical Chemistry News, September 1986[42]. The FDA-recommended protocol seeks two consecutive MSAFP measurements, sampled at least 1 week apart, before further diagnostic modalities are suggested. Ultrasonography is performed as the next step to corroborate gestational date, check for twins, evaluate the viability of the fetus, and possibly detect anencephaly. If the elevated MSAFP value remains unexplained, level 2 ultrasonography or gray-scale sonography is performed to search for other fetal malformations, most prominently open spina bifida and open wall defects. Amniocentesis for the purpose of measuring amniotic fluid AFP (confirmatory testing of amniotic fluid samples determined to have elevated levels of AFP are performed for the presence of neural derived acetyl-cholinesterase) is also offered, following genetic counseling of the maternal patient with the attending physician.[36] Generally speaking, the rate of amniocentesis should not exceed 3 percent of the tested population.[62 63]

3. The reason for the determination of the performance characteristics discussed below (for in-vitro devices) are two-fold; to assess the influence of disease prevalence upon the clinical laboratory test kit's assessment of a patient's condition and to assess the clinical laboratory procedures by which the poor detecting power of an in-vitro device for relatively low prevalence disease may be improved. These assessments include clinical sensitivity (CSE) and clinical specificity (CSP). Four parameters help FDA assess the probability of a correct in-vitro device result: sensitivity, specificity, prevalence of disease condition and efficiency of the test results in detecting diseased individuals. Predictive value of a positive test and predictive value of a negative test are secondary performance characteristics which are functions of CSE, CSP. Disease prevalence can and should be calculated for the range of expected disease prevalence. Interpretations of clinical in-vitro devices, aside from clinical considerations, are based on the probability that the test result will be within a given normal range of analyte values; in this instance, the range of elevated values for AFP in maternal specimens derived from pregnancies with neural tube defective fetuses.

Definitions:

Additional definitions are needed to allow interpretive reporting[21 22 23 24 25] of clinical laboratory AFP device results.

False-positive Rate: FDA accepts two definitions of false-positive rate (FPR). In evaluating the overall performance of clinical investigators, a manufacturer's clinical studies manager is interested in the testing laboratories' proportion of disease-free persons tested who test positive (FP) compared to that portion of disease-free persons tested who test negative (TN); for the testing laboratory, the false-positive (FP) rate is

$$FP/(FP + TN)$$

On the other hand, a clinical investigator, and particularly the physician's patient who has already tested positive, is much more concerned with the ratio between false-positive test results and all positive test results, because this ratio relates to the pregnancy outcome of the particular case in question. In this approach, the false-positive (FP) rate is:

$$FP/(FP + TP)$$

This is equivalent to 1-Predictive Value of Positive Test (1-PVPT). In practical situations employing tests for low prevalence diseases or conditions, this latter (detection) definition of false positive rate gives rates that are much higher. It is this definition of the false positive rate which most manufacturer's Managers of Clinical Testing (MSAFP testing programs and Clinical Investigators) apply when communicating with the clinician and with FDA.

4. Comparison Studies:

   a. Compare the device to at least one device for which there is an approved PMA.

   b. Provide data using three different lots tested in one laboratory.

      Compare results obtained using AFP samples free from interfering substances from 40-100 maternal patients covering the whole assay range (from low to high levels of AFP.[7 10] Analyze the data using linear regression methods [60] (the X axis is the independent variable or comparison test; the Y axis is the dependent variable or new test[54 56]). Linear-regression analysis is often most useful for estimating the differences or errors between two analytical methods, because the errors can be calculated at any medically important concentration

within the range studied; furthermore, the slope and intercept may give some indication of the type of systematic error, which may aid in reducing the analytical error, which may then aid in reducing the analytical errors. Because the reliability of the estimates of slope and intercept can be affected by nonlinearity in the data set, outliers, a narrow range, and variability of the comparison method, it is preferable that samples cover the complete range of concentrations that might be encountered[51]

5. Prozone or High-Dose Hook Effect Studies:

Immunoradiometric (IRMA) and similar type assays:

Test a sample with a very high concentration of AFP, diluted and undiluted.[53] If the test result is not erroneously low, state in the Performance Characteristics section of the product package insert the quantitative level below which no high dose hook effect was observed.

6. Stability Studies

a. Stability study requirements are outlined in 21 CFR 809.10 (a) and (b), 21 CFR 211.166. These studies are performed and conducted with the above specifications and performance evaluations. Real time stability studies should be the basis for estimating expiration dating of reagents. Data from three different manufactured lots are required [53]. Data from accelerated stability studies are acceptable only as interim data.

b. Shipping conditions that must be addressed. Include test results that show that the reagents are stable under variable shipping temperatures or state why the device would not be affected by shipping temperatures, e.g., the device is shipped only on dry ice by guaranteed overnight delivery service.

Labeling Considerations:

Provide the slope, intercept, and their estimated standard errors, correlation coefficient, the standard error of the estimate of the working curve; the assay range, and nature and size of samples tested should be reported in the Performance Characteristics section of the package insert. Copies of all proposed labeling for the device, including any physician and patient brochures, literature, or advertising that constitutes labeling under section 201(m) of the act (21 U.S.C. 321 (m)) must

be provided. Patient brochures must be provided to women with clearly written descriptions of the characteristics of the fetal neural tube conditions. The information presented can obviously not be exhaustive, and it should suggest that further questions be discussed with the physician. The patient brochure should also contain information concerning the procedures available to make a definitive diagnosis and should urge the mother to wait for additional test results before taking further action.

If she is to find the explanation of the follow-up procedures meaningful, the mother must clearly understand the need for them. The most important follow-up procedures, notably amniocentesis and the ultrasonic scan, should be explained in terms of how they help provide the physician with a method of interpreting test results. The general labeling requirements for medical devices are contained in 21 CFR Part 809.10. These regulations specify the minimum requirements for all in vitro devices. Additional guidance regarding device labeling can be obtained from FDA's publication "Labeling: Regulatory Requirements for Medical Devices", and from the Office of Device Evaluation's "Device Labeling Guidance"; both documents are available upon request from the Division of Small Manufacturers Assistance (HFZ-220), Center for Devices and Radiological Health, Food and Drug Administration, 5600 Fishers Lane, Rockville, MD 20857.

The package insert should be evaluated carefully because the information contained in these seven sections is important for correct interpretation of AFP test results. These seven sections are: (1) intended uses, with particular attention paid to the sponsor's attempts to include "inferred intended use" or expanded applicability based upon "literature citations" with little or no clinical data to support this expanded use, (2) conditions for use, (3) principles of procedure or operation, (4) reagents, (5) limitation(s) of the procedure, (6) interpretation of test results and (7) performance characteristics.

Manufacturers who choose to move the production facilities of their AFP device must re-establish AFP median values by testing 100 serum samples for each gestational week (15 through 20) for calculating median values. Three lots of manufactured reagents should be used to verify performance characteristics of the newly calculated median values. Manufacturers should be sensitive to the fact that clinical laboratories may be tempted to use the median values provided in package inserts as a source of reference data. Such median values have been documented to be widely in error for some manufacturer's kits, resulting in large numbers of false-positive or false-negative test results. FDA requires that manufacturers submit PMA supplements of clinical data to corroborate changes in package insert median values, since use of outdated median values when changes in AFP test kit

reagents (e.g., radioactive iodination of the AFP reagent) may result in large numbers of false-positive or false-negative test result interpretations.

Absolute compliance with the in vitro labeling regulations (21 CFR 809.10) would include patient labeling to give prospective patients (or their parents/guardian) realistic expectations of the benefits and risks of AFP testing. Such information should be written and formatted so as to be easily read and understood by most patients and should be provided to patients prior to scheduling AFP measurement so that each patient has sufficient time to review the information and discuss it with her physician(s). Technical terms should be kept to a minimum and should be defined if they must be used. Patient information labeling should be designed to meet the seventh grade reading comprehension level.

Kearby J. Fugate, Ph.D.

V. BIBLIOGRAPHY

1. Brock DJH, Sutcliffe RG: Alpha-fetoprotein in the antenatal diagnosis of anencephaly and spina bifida. Lancet; ii:197-199, 1972.

2. Leek AE, Ruoss CF, Kitau MJ, Chard T: Raised alpha-fetoprotein in maternal serum with anencephalic pregnancy. Lancet; ii:385, 1973.

3. Brock DJH, Bolton AE, Monaghan JM: Prenatal diagnosis of anencephaly through maternal serum alpha-fetoprotein measurements. Lancet; ii:923-924, 1973.

4. Wald NJ, Brock DJH, Bonnar, J: Prenatal diagnosis of spina bifida and anencephaly by maternal serum alpha-fetoprotein measurement. Lancet; i:765-766, 1974.

5. Brock DJH, Scrimgeour JB: Early prenatal diagnosis of anencephaly. Lancet; ii:1252, 1972.

6. Maternal Serum Alpha-fetoprotein Measurement in Antenatal Screening for Anencephaly and Spina Bifida in Early Pregnancy. Report of the U.K. Collaborative Study in AFP in Relation to Neural Tube Defects. Lancet; i: 1323-32, 1977.

7. 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 AFP in relation to Neural Tube Defects. Lancet; ii:651-662, 1979.

8. Haddow JE, Macri JN (eds). Proceedings of the First Scarborough Conference: Screening for Neural Tube Defects in the United States, Foundation for Blood Research, Scarborough, Maine, 1977.

9. Haddow JE, Macri JN (eds). Proceedings of the Second Scarborough Conference: Alpha-fetoprotein Serum Screening in Pregnancy, Foundation for Blood Research, Scarborough, Maine, 1978.

10. Haddow JE, Macri JN, Wald NJ (eds). Proceedings of the Third Scarborough Conference: The Regional Application of Alpha-fetoprotein Serum Screening and Ultrasonography in Mid-Pregnancy. Foundation for Blood Research, Scarborough, Maine, 1980.

11. Haddow JE, Macri JN: Prenatal screening for neural tube defects. J Am Med Assoc; 242:515-516, 1979.

12. Wald NJ, Cuckle H: Alpha-fetoprotein in the antenatal diagnosis of open neural tube defects. Br J Hosp Med; 23:473-489, 1980.

13. Kolata GB: Prenatal diagnosis of neural tube defects. Science; 209:1216-1218, 1980.

14. Proceeding, FDA/NICHHD National Conference on Maternal Serum AFP: Issues in the prenatal screening and diagnosis of neural tube defects. July 1980, Washington, D.C., B Gastell,JE Haddow, JC Fletcher, A Neale (eds), U.S. Government Printing Office, 1980.

15. Brock DJH, Barron, L, Jelen P, Watt M, Scrimgeour JB: Maternal serum alpha-fetoprotein measurements as an early indicator of low birth weight. Lancet; ii:267. 1978.

16. Wald N, Cuckle H, Stirrat GM, Bennett MJ, Turnbull AC: Maternal serum alpha-fetoprotein measurements as an early indicator of low birth weight. Lancet; i:268-270, 1977.

17. Macri JN, Weiss RR, Libster B, Cagen MA: Maternal serum alpha-fetoprotein and low birth weight. Lancet; i:600, 1978.

18. Wald NJ, Cuckle H, Stirrat GM, Turnbull AC: Maternal serum alpha-fetoprotein and birth weight in twin pregnancies.Br J Obstet Gynecol; 85:582-584, 1978.

19. Wald NJ, Cuckle H, Stirrat GM: Maternal serum alpha-fetoprotein levels in triplet and quadruple pregnancy. Br J Obstet Gynecol; 85:124-126, 1978.

20. National Committee for Clinical Laboratory Standards. Assessing the quality of radioimmunoassay systems; Revised Guideline. NCCLS document LA1-A. NCCLS, 771 East Lancaster Avenue, Villanova, Pennsylvania, 19085, 1992.

21. Speicher CE, Smith JW:Interpretive reporting in clinical pathology. J Am Med Assoc; 243:1556-1560, 1980.

22. Lundberg, GD: The reporting of laboratory data interpretations: To omit or commit? J Am Med Assoc; 243:1554-1555, 1980.

23. Sohn D: The clinician-laboratory connection, the vital link: Comments regarding the NCCLS proposed standard for clinical laboratory requisition forms. Ther Drug Monit; 1: 1979.

24. Pippenger CE: Editorial. Ther Drug Monit; 1:451-452, 1979.

25. Galen RS, Gambino SR: Beyond normality: The predictive value and efficiency of medical diagnosis. John Wiley and Sons, New York, 1975.

26. Sizaret P, Breslow N, Anderson, SG, and twelve other participants: Collaborative study of a preparation of human cord serum for use as a reference in the assay of alpha-fetoprotein. J Biol Stand; 3:201-223, 1975.

27. Reimer CB, Smith J, Wells TW: The U.S. National Reference Preparation for Alpha-fetoprotein in Mid-Pregnancy Serum. Clin Chem; 28:709-716, 1982.

28. Sizaret P, Anderson SG: The international reference preparation for alpha-fetoprotein. J Biol Stand; 4:149, 1979.

29. Muenz LR, Sizaret P, Bernard C, et al: Results of the second international study on the WHO alpha-fetoprotein standard. J Biol Stand; 6:187-199, 1978.

30. NCCLS Approved Standard: C2-A. Calibration, reference materials and control materials in clinical chemistry. National Committee for Clinical Laboratory Standards, Villanova, Penn. 19085, 1974.

31. Wald NJ, Cuckle H, Boreham J, Stirrat G: Small biparietal diameter of fetuses with spina bifida: Implications for antenatal screening. Br J Obstet Gynecol; 87:20-21, 1980.

32. Weiss RR, Macri JN, Elligers KW: Origin of amniotic fluid alpha-fetoprotein in normal and defective pregnancies. Obstet Gynecol; 47:69 1976.

33. Hook EB: Down syndrome frequency in human populations and factors pertinent to variations in rates. Trisomy 21 (Down Syndrome); Research Perspectives of the National Institute of Child Health and Human Development. DeLaCruz FF, Gerald PS (eds.). Baltimore, University Park Press, 1980.

34. Kjessler B, Johansson SG, Lidbjork G, Sherman MS: Alpha-fetoprotein (AFP) in early pregnancy. Acta Obstet Gynecol Scand Suppl; 69:1-94, 1977.

35. Wald NJ, Cuckle HS, Haddow JE: Should ultrasound be used to estimate gestational age in the screening and antenatal diagnosis of open tube defects in early pregnancy? Lancet; ii:690, 1980.

36. Smith AD, Wald NJ, Cuckle HS, Stirrat GM, Bobrow M, Lagercrantz H: Amniotic fluid acetylcholinesterase as a possible diagnostic test for neural tube defects in early pregnancy. Lancet; i:684-688, 1979.

37. Amniotic Fluid Acetylcholinesterase Measurement in the Prenatal Diagnosis of ONTD. Second Report of the Collaborative Acetylcholinesterase Study. Prenat Diagn; 9:813-829, 1989.

38. 21 CFR Parts 16,20,899 [Docket No. 80N-0002]. FEDERAL REGISTER Vol. 45, No.218, Friday, November 7, 1980, Proposed Rules, AFP Test Kits; Proposed Restrictions and Proposed Additional Quality Control and Testing Requirements. Testing Protocol and Genetic Counseling Flow Chart, page 74167.

39. 21 CFR Parts 16,20,899 [Docket No. 80N-0002]. FEDERAL REGISTER Vol. 48, No.118, Friday, June 17, 1983, Alpha-Fetoprotein Test Kits; Withdrawal of Proposed Rule. Quality Control, page 277781.

40. Haddow, JE and Milunsky, A: Deregulation of Screening For Alpha-fetoprotein In Pregnancy. N Eng J Med; 310:1669, 1984.

41. Burton, BK, Dillard, RG, Clark, EN: The psychological impact of false positive elevations of maternal serum alpha-fetoprotein. Am J Obstet Gynecol; 151:77-82, 1985.

42. Vadlamudi, SK, Fugate, KJ, Appell, RN, and Dierksheide, WD: AFP Testing Protocol Questioned. Letter to Editor: Clinical Chemistry News; 12:no. 9, September 1986.

43. The Quality Control of AFP Reagent and Assay for the Antenatal Screening and Diagnosis of Open Neural-Tube Defects: Report of a (1978) Workshop Sponsored by the U.S. National Institute of Child Health and Human Development, Bethesda, MD. Clin Chim Acta; 105:9-24, 1980.

44. Kallner A, Magid E, and Albert A., eds. Improvement of comparability and compatibility of laboratory assay results in life sciences. Vadlamudi, et al., Third Bergmeyer Conference on Immunoassay Standardization; Lenggries, Germany, Scan J Clin Lab Invest; 51:(Suppl 205):1-143, 1990.

45. Noorgaard-Pedersen B, Toftager-Larsen JP, Hindersson P: Concanavalin-A reactivity pattern of human amniotic fluid AFP examined by crossed affino-immunoelectrophoresis: A definite test for neural tube defect? Clin Genet; 17:1-8, 1980.

46. Ganrot PO: Crossed immunoelectrophoresis. Scand J Clin Lab Invest; 29:39-47, 1972.

47. Dyer SN, Burton BK, Nelson LH: Elevated maternal serum AFP levels and oligohydramnios: poor prognosis for pregnancy outcome. Am J Obstet Gynecol; 31(2):336-369, 1987.

48. Haddow JE, Knight GJ, Kloza EM, Palomaki G: AFP, vaginal bleeding, and pregnancy risk. Br J Obstet Gynoecol; 93:589-593, 1986.

49. Burton BK: Elevated maternal serum AFP (MSAFP)interpretation and follow-up. Clin Obstet Gynecol; 31 (2): 293-305; 1988.

50. Cowan LS, Phelps-Sandall B, Hanson FW, Peterson AG, Tennant L: Prenatal diagnostic center's first year experience with the California AFP screening program. Am J Obstet Gynecol; 160:1496-1504, 1989.

51. Palomaki GE, Hill LE, Knight GJ, Haddow JE, Carpenter M: Second trimester maternal serum AFP levels in pregnancies associated with gastroschisis and omphalocele. Obstet Gynecol; 71:906-909, 1988.

52. Bock JL: Current issues in maternal serum AFP screening. Am J Clin Pathol; 97:541-554,1992.

53. Vadlamudi SK, Stewart WD, Fugate KJ, Tsakeris TM: Performance characteristics for an immunoassay. Scand J Clin Lab Invest; 51:134-138, 1991

54. Peters T, Westgard JO: Evaluation of methods, chapter 7 in: Tietz NW, editor. Fundamentals of clinical chemistry, Third Edition, Philadelphia: Sanders. 225-37, 1987.

55. National Committee for Clinical Laboratory Standards. Evaluation of the linearity of quantitative analytical methods; proposed guideline. Order code EP6-P. 1986.

56. Westgard JO, de Vos DJ, Hunt MR, Quam EF, Carey RN, Garber CC: Method evaluation. American Society of Medical Technology, @@@ Washington. 1978.

57. Information for authors: Clin Chem; 37:1-3. 1991

58. National Committee for Clinical Laboratory Standards. Interference Testing in clinical chemistry; proposed guideline. Order code EP7-p, 1986.

59. National Committee for Clinical Laboratory Standards. Evaluation of precision performance of clinical chemistry devices - second edition; tentative guidelines. 1-56. Order code EP9-p, 1991.

60. National Committee for Clinical Laboratory standards. User Comparison of quantitative clinical laboratory methods using patient samples, proposed guideline, 6 (1). Order code EP9-p, 1985.

61. Bayes Reverend Thomas: An essay toward solving a problem in the doctrine of chance. Philo Trans Roy Soc; 53:370-418. 1763.

62. Adams MJ, Windham GC, James LM, Greenberg F, Clayton-Hopkins JA, Reimer CB, Oakley GP: Clinical interpretation of maternal serum AFP concentrations. Am J Obstet Gynecol; 148:241-254, 1984.

63. Crandall BF, Lebherz TB, Schroth PC, Matsumoto M: Alpha - fetoprotein concentrates in maternal serum: Relation to race and body weight. Clin Chem; 29:531, 1983.

64. Adams MJ Jr, Windham GC, James LM, Greenberg F, Clayton-Hopkins JA, Reimer CB, Oakley GP: Risk reporting of maternal serum AFP (AFP) concentrations. Atlanta: U.S. Department of Health and Human Services, 1985.

65. Wald NJ, Cuckle H: Recent Advances in Screening for Neural Tube Defects and Down's Syndrome, Bailliere's Clinical Obstetrics and Gynaecology, Rodeck C (ed), Vol. I; 3:649-676, 1987.

66. Crandall BF, Robertson RD, Lebherz TB, King W, Schroth PC: Maternal serum alpha-fetoprotein screening for the detection of neural tube defects. Western J Med; 138:524-530, 1983.

67. Macri JN, Kasturi RV, Krantz DA, Koch KE: Maternal Serum AFP Screening, Maternal Weight, and Detection Efficiency, 'Am J Obstet Gynec; 155:758-760, 1986.

68. Soloway HB, Sohn AP, Morris C, Slaughter RJ: Operational Considerations in Maternal Serum AFP Screening, MLO; December 37-40, 1988.

69. Milunsky A, Alpert E, Kitzmiller JL, Younger MD, Neff RK: The Importance of Serum AFP Screening in Diabetic Pregnant Women, Am J Obstet Gynec; 142;1030-1032, 1982.

70. Wald N, Cuckle H, Boreham J, Terzion E, Redman C: The Effect of Maternal Weight on Maternal Serum AFP Levels, Br J Obstet Gynec; 88:1094, 1981.

71. Cowchock FS, Jackson LG: An analysis of pregnancies with elevated alpha-fetoprotein levels in maternal serum and/or amniotic fluid samples, Birth Defects; XV:75-85, 1979.

72. Smithells RW, Seller MJ, Harris R, Fielding DW, Schorah CJ, Nevin NC, Sheppard S, Read AP, Walker S, Wild J: Further experience of vitamin supplementation for prevention of neural tube defect recurrences. Lancet; ii:1027-1031, 1983.

73. Hibbard ED, Smithells RW: Folic acid metabolism and human embryopathy. Lancet; i: 1254, 1965.

74. Laurence, KM, James N, Miller MH, Tennant GB, Campbell H: Double blind randomized controlled trial of folate treatment before conception to prevent recurrence of neural-tube defects. B M J; 282:1509-11, 1981.

75. Czeizel AE, Dudas I: Prevention of the first occurence of neural-tube defects by periconceptional vitamin supplementation. N Engl J Med; 327:1832-5, 1992.

76. Bishop JC, Dunstan FDJ, Nix BJ, Reynolds TM, Swift A: All MoMs are not equal: some statistical properties associated with reporting results in the form of multiples of the median. Am J Hum Genet; 52:425-430, 1993.

77. Fugate KJ: Post-Approval studies for monitoring the performance of alpha-fetoprotein test kits as an aid in the detection of ONTD. Immunology Device Panel Meeting. Hubert H. Humphrey Building, Washington, D.C. June 22, 1988.

78. 21 CFR Parts 16,20,899 [Docket No. 80N-0002]. FEDERAL REGISTER Vol. 48, No.118, Friday, June 17, 1983, Alpha-Fetoprotein Test Kits: Withdrawal of Proposed Rule. Pages 277780-277782.

79. Felder RA, Butts W, Bradley L, King P, MacMahon W, Wians F,Jr., Dev J: A multi-center study of serum AFP as an aid in the detection of fetal open neural tube defects (NTD). Clin Chem; 36: Abstract; Annual Meeting, July 25, 1990.

80. Knight GJ, Palomaki GE, Haddow JE: Assessing reliability of AFP test kits. Contem Ob-Gyn Tech; October 1987.