**Review Criteria for In Vitro Diagnostic Devices for the Assessment of Thyroid Autoantibodies using Indirect Immunofluorescence Assay (IFA), Indirect Hemagglutination Assay (IHA), Radioimmunoasay (RIA), and Enzyme Linked Immunosorbent Assay (ELISA)**

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.

REVIEW CRITERIA FOR IN VITRO DIAGNOSTIC DEVICES FOR THE ASSESSMENT OF THYROID

AUTOANTIBODIES USING INDIRECT IMMUNOFLUORESCENCE ASSAY (IFA), INDIRECT

HEMAGGLUTINATION ASSAY (IHA), RADIOIMMUNOASSAY (RIA), AND ENZYME LINKED

IMMUNOSORBENT ASSAY (ELISA).

This is a flexible document representing the current concerns and suggestions

regarding thyroid autoantibodies in vitro diagnostic devices employing IFA,

IHA, RIA, and ELISA methodologies. It is based on (1) current basic science,

(2) clinical experience, (3) the Safe Medical Devices Act of 1990 (SMDA) and

(4) FDA regulations in the Code of Federal Regulations (CFR). As advances are

made in science and medicine, these review criteria will be re-evaluated and

revised as necessary.

PURPOSE OF THE GUIDANCE DRAFT

The purpose of this document is to provide guidance and clarification on

information to present to the Food and Drug Administration (FDA) before a

device to detect, quantitate and/or semi-quantitate thyroid autoantibodies in

clinical specimens can be cleared for marketing.

A premarket notification 510(k) submission provides evidence that the device

is accurate, safe, effective and substantially equivalent to a predicate

device legally marketed in the United States.

DEFINITION

This generic type of device is intended for use in clinical laboratories or

physician's office laboratories\* as an in vitro diagnostic test for the

qualitative, quantitative and/or semi-quantitative measurement of thyroid

autoantibodies by IFA, IHA, RIA or ELISA.

\*Devices may be cleared for use in physician's office laboratories when

additional data are submitted to demonstrate performance in these settings.

PRODUCT CODES: JNL, DDC, DDJ, JZO

REGULATION NUMBER:

21 CFR 866.5870 Thyroid autoantibody immunological test system.

Identification. A thyroid autoantibody test system is a device that consists

of the reagents used to measure, by immunochemical techniques, thyroid

autoantibodies (antibodies produced against the bodies own tissues).

Measurement of thyroid autoantibodies may aid in the diagnosis of certain

thyroid disorders, such as Hashimoto's disease (chronic lymphocytic

thyroiditis), nontoxic goiter (enlargement of the thyroid gland), and Graves'

disease (enlargement of the thyroid gland with protrusion of the eyeballs).

CLASSIFICATION: CLASS II (Performance Standards)

PANEL: Immunology (82)

REVIEW REQUIRED: Premarket notification (510(k))

I. CLINICAL INDICATIONS/SIGNIFICANCE/INTENDED USE

A. INTRODUCTION

The principle role of the thyroid gland is the storage and synthesis of

thyroid hormones. The normal thyroid gland is composed of numerous follicles,

each of which consists of a single layer of epithelial cells surrounding a

central lumen containing colloid. Thyroglobulin is the major component of the

thyroid follicular colloid. It is produced by the thyroid epithelial cells

and is a water soluble glycoprotein with a molecular weight of 670,000

daltons. Small amounts of thyroglobulin are present in serum. Microsomal

antigen is a 110,000 dalton glycoprotein present in the cytoplasm and on the

apical membrane of the thyroid cell. Reports have shown that thyroid

peroxidase (TPO) is the major component of the thyroid microsomal antigen and

is the antigenic component recognized by autoantibodies directed against the

microsomal glycoprotein. 1

Autoimmune thyroid disease is organ specific and is defined by the presence of

circulating antibodies. The most common antibodies seen are antibodies to

thyroglobulin and microsomal or thyroid peroxidase (TPO) of the thyroid

epithelial cell lining.

Autoantibodies to these thyroid antigens are detected in Hashimoto's disease

and its variants, Graves' disease, myxedema, nontoxic goiter, and thyroid

carcinoma. Antithyroglobulin and antimicrosomal (TPO) antibodies are seen

most often and are in the highest titers in Hashimoto's disease. Patients

with Graves' disease can also demonstrate thyroid antibodies of relatively

high titer. These autoantibodies are also seen in low titers in other thyroid

diseases and in individuals without clinical evidence of disease. 2

Two other thyroid antibodies less commonly seen are antibodies to the second

antigen of the colloid or CA-2 and human thyroid-stimulating immunoglobulin

(TSI). CA-2 is a minor protein of colloid distinct from thyroglobulin and

antibodies to it are of uncertain significance. Positive CA-2 reactions are

seen in a low percentage of thyroiditis patients in the absence of other

antithyroid antibodies and in 5 to 10 per cent of patients with Graves'

disease and thyroid cancer. 2 CA-2 antibodies can be detected by IFA. TSI,

previously referred to as long-acting thyroid stimulator (LATS), is a

polyclonal gamma globulin which appears to bind to a receptor on thyroid cells

and stimulates thyroid activity. It is present in the sera of about 50 per

cent of those with Graves' disease and is absent or present only in a small

number of patients with nodular toxic goiter or other thyroid disorders. TSI

can be measured by a competitive inhibition assay using 125I labeled TSH and

thyrotropin (TSH) receptors.

Hashimoto's thyroiditis is an inflammatory condition occurring in about 1 to 2

per cent of the population, mainly in middle-aged women, and is characterized

by gland enlargement as a result of marked lymphocytic inflammatory changes.

The latter may consist of lymphoid follicles with active germinal centers in

which much of the antithyroglobulin antibody appears to be synthesized.

Normal thyroid glandular structures are adversely altered, and in prominent

cases progressive disease may lead to thyroid atrophy and myxedema (dry, waxy

type of swelling, with abnormal deposits of mucin in the skin and other

tissues associated with hypothyroidism). In thyrotoxicosis, the thyroid may

contain small areas of lymphoid infiltration as well as evidence of the

typical glandular hyperactivity. Graves' disease is a multi-systemic

disorder, particularly in young to middle-aged females, consisting of varying

degrees of (1) hyperthyroidism with diffuse hyperplasia of the thyroid (the

most common pattern seen with diffuse toxic goiter); (2) a myopathy; and (3)

an infiltrative ophthalmopathy, frequently leading to exophthalmos (protrusion

of the eye ball). 2

As with all the organ-oriented diseases associated with autoantibody

production, it is important to determine when the antibodies under discussion

are pathogenic or are reacting to antigens liberated as a result of tissue

damage due to non-immune causes. Immune reactivity may not be the primary

pathogenic event but, once present, causes further tissue damage. Evidence

against a primary pathogenic role for thyroid autoantibodies in Hashimoto's

thyroiditis and Graves' disease is (1) the lack of correlation between the

level of autoantibody and the severity of disease in individual cases, and (2)

the lack of development of thyroid disease in infants with high levels of

antithyroid antibodies because of placental transfer.

B. INSTRUCTIONS:

Provide a concise discussion to include the following as appropriate. Support

the discussion with key literature citations.

 1. Clinical indications, significance and intended use.

 2. Background description of the thyroid disease involved including the

 type of population affected (sex, age, etc.)

 3. Significance of a positive result (disease indication and follow up

 testing).

 4. Significance of false positive and false negative results.

 5. Salient concerns of the medical community including relevant medical

 issues that may impact the review process or possibly the development

 of public policy.

 6. A brief historical summary of all test methodologies used to detect

 the antibody(ies).

 7. Merits/advantages and limitations/disadvantages of the device

 methodology(ies) compared to other available methodologies.

 8. Matrices.

II. DEVICE DESCRIPTION:

The determination of substantial equivalence is based on the specific intended

use (what analyte is detected and the indications for use) and the

technology/methodology utilized in the device. Discuss the principles of the

device methodology and whether it is well-established or new and unproven.

III. CLINICAL AND NONCLINICAL LABORATORY STUDIES: SPECIFIC PERFORMANCE

 CHARACTERISTICS:

FDA requests different types and amounts of data and statistical analyses in

pre-market notification applications to market in vitro diagnostic devices.

The amount and type of data requested depends on the intended use,

technological characteristics of the new device, whether the test is

qualitative, quantitative or semi-quantitative and on certain claims made by

the manufacturer. The performance of the device can be established by

comparison to any legally marketed medical device (the predicate) with the

same intended use.

Prove all claims for substantial equivalence and specific performance

characteristics for using the device. Clearly document all protocols for in

vitro testing. Present test data results with analyses and conclusions.

Summarize results and include explanations for unexpected results and any

additional testing performed. Charts (scattergrams, histograms, etc.) may be

used as part of the analyses and conclusions when appropriate. Actual,

unprocessed laboratory data may be requested.

A. ANALYTICAL/LABORATORY/IN VITRO STUDIES

 1. Validation of the Cut-off

 Describe the rationale for determination of the assay cut-off(s).

 Furnish descriptive information and laboratory data to show how the

 cut-off point (distinction between positivity and negativity or

 medical decision limit) was determined by the assay.

 a. Define the population(s) used, including the following

 information:

 i. Number of samples in the normal population (used to

 determine initial screening dilution) with samples

 summarized according to gender and age groups. 4

 ii. Number of specimens included in each disease group

 summarized according to gender and age groups.

 iii. Geographical area(s) from which the population was derived.

 iv. Graphical (e.g., scattergrams, histograms, etc.)

 representation of population characteristics.

 b. Define the statistical method used to determine the cut-off

 point(s).

 c. Present a Receiver Operator Curve (ROC) analysis of cut-off point

 selection and other graphical representations as appropriate.

 d. Define the basis for the equivocal zone (if applicable).

 2. Reagent characterization

 a. Give a brief description of the antigen(s) and antibody(ies)

 used in the assay.

 b. If any recombinant technology was used in the preparation of the

 antigen(s), describe method used.

 3. Assay Specificity/Interfering Substances

 Any potentially cross reacting or interfering substances encountered

 in specific specimen types or conditions should be tested using the

 assay system, e.g., hemolysis, lipemia, microbial contamination,

 additional analytes or other autoantibodies present, and storage or

 freeze-thawing.

 a. Verify that recommended storage conditions are compatible with the

 assay. State the optimal conditions based on specimen storage

 stability studies. Both false positivity and negativity should be

 evaluated (if applicable).

 b. If the use of plasma is claimed, a study with each anticoagulant

 must be performed to show that each anticoagulant does not

 interfere with the assay.

 i. For each anticoagulant, test 10 matched serum and plasma

 specimens which are positive at the cut-off point.

 ii. For each anticoagulant, test 10 matched negative serum and

 plasma specimens.

 Provide an explanation if interference from the anticoagulant is

 not anticipated (e.g., high dilution factor).

 4. Performance Characteristics

 Include the following performance characteristics:

 a. Analytical Sensitivity (if applicable)

 The analytical sensitivity or detection limit is defined as the

 lowest quantity differentiated from zero (95% confidence intervals

 or 2 standard deviations (SD) above the mean of the Zero control

 are commonly used). 4,5 Run the Zero standard (Zero diluent) at

 least 20-25 times in the same run and calculate the mean of the

 Zero standard and 2 SD of the mean (counts, OD's, etc.). If

 levels of the analyte are not clinically significant,

 determination of the detection limit may be irrelevant.

 b. Relative Sensitivity and Specificity

 The relative sensitivity and specificity as determined by

 comparison to a legally marketed device or to a reference method

 should be determined and reported in the Performance

 Characteristics section of the package insert.

 c. Linear Range

 Validate the linear range of the assay with normal and abnormal

 specimens covering the entire reportable range of the assay. 6

 d. Reproducibility and Repeatability Studies 4,5,6,7,8

 The National Committee for Clinical Laboratory Standards (NCCLS)

 recommends an analysis of variance experiment testing two

 clinically significant levels near medical decision limits (normal

 or elevated) of an analyte, in this case thyroid autoantibodies. 8

 Use controls simulating patient samples or actual patient

 specimens 3 times in the same run and in two different runs each

 day for 20 days. This permits separate estimation of between-day,

 between-run and within-day standard deviations (SDs), as well as

 within-run and total SDs.

 i. Qualitative/Quantitative Tests:

 Calculate total, between- and within-day and between- and

 within-run means and coefficients of variation of

 imprecision for each set of values.

 ii. Semi-quantitative Tests:

 In devices with a titration format, e.g., immunofluorescence

 assays, demonstrate that intra-run reproducibility is within

 the commonly accepted limits of plus or minus one two-fold

 dilution.

 iii. Means, SDs, and Coefficients of Variation:

 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.

 e. Prozone or High-Dose Hook Effect Studies

 Test a sample with the highest titer available, serially diluted

 and undiluted. If prozone problems are encountered, state in the

 Performance Characteristics section of the package insert the

 titer at which prozone problems were detected and a procedure for

 the user to follow to correct the problem. Where appropriate,

 describe the appearance of a prozone reaction for subjective

 tests.

 f. Alternative Testing Sites

 Include reproducibility studies performed in these settings. In-site

 testing for new technologies should include at least three

 independent alternative testing sites. At each of the three

 sites, the precision and accuracy of the device should be

 evaluated. A statistically valid number of samples should be

 tested by the site personnel and by professional laboratory

 personnel, and the results compared, to show how the device

 performs in the hands of the lesser trained user.

 5. Comparison Studies

 Compare the new device to a legally marketed device. Include the

 package insert for the legally marketed device.

 It is recommended that a recognized reference method (if available)

 also be employed for comparison to enable a fair evaluation of the

 proposed device's performance characteristics, particularly if there

 are broad differences in methodology/ technology between the new

 device and the legally marketed device.

 a. Qualitative Tests:

 The studies should be performed on an adequate number of positive

 and negative specimens to support statistical significance. (An

 appropriate number may be suggested by a statistician.)

 b. Quantitative/Semi-quantitative Tests:

 An assay is considered quantitative only if a recognized reference

 material of known concentration is available for standardization

 of a calibrator or standard used in the assay to calculate

 results.

 If the same reference material is used in the new device as in the

 predicate device to substantiate the quantitative claim,

 comparison data should be presented to show correlation between

 the two assays when running the reference material as samples.

 Run the serially diluted reference material by the new device and

 the predicate device. The assays should show similar results.

 Compare results obtained using positive thyroid autoantibody

 samples free from interfering substances from 40-100 persons

 covering the whole assay range (from low to high levels of

 antibodies). 7,9

 Perform a linear regression analysis and report the slope,

 intercept, correlation coefficient, the assay range, and the

 nature and number of samples tested.

 c. Comparison Discrepancies:

 Equivocal results or discrepancies between the new device and the

 comparison method should be resolved using another method or

 clinical diagnosis.

 6. Specimen Collection and Handling Conditions

 State specimen collection, storage and handling conditions in the

 package insert and provide data or appropriate literature references

 in the submission to substantiate claims.

 7. Computer Controlled Medical Devices

 For information regarding computer assisted clinical laboratory

 devices, refer to "Review Guidance for Computer Controlled Medical

 Devices Undergoing 510(k) Review" available from the Division of

 Small Manufacturers Assistance (DSMA), 1-800-638-2041.

B. CLINICAL INVESTIGATIONS

 In certain instances it is necessary to require comparative clinical data

 to establish substantial equivalence, e.g., a new or unfamiliar methodology

 or technological feature is introduced in a device category in which

 clinical performance is claimed to be equivalent to a legally marketed

 device using "conventional" technology.

 For 510(k) submissions, perform a comparison of the device to a legally

 marketed device. Ideally this study should be done at independent clinical

 laboratory site(s). A minimum of two additional independent investigators

 at separate outside locations is recommended. The investigators should be

 identified by institutional name and address.

 1. Adequate Clinical Investigations

 a. Prove all claims for substantial equivalence and specific

 parameters for using the device.

 b. Describe all protocols for clinical studies and consistently

 adhere to the protocols.

 c. Determine the sample size, prior to beginning the study, that will

 be statistically sufficient to determine whether or not the device

 is safe and effective.

 d. Sampling Method:

 Describe sampling method used in the selection and exclusion of

 patients.

 i. Patient selection:

 Include samples from individuals with diseases or conditions

 that may cause false positive or false negative results with

 the device. Ideally, a prospective study is preferred.

 However, if a retrospective study is used, include all

 eligible patients who meet the patient selection criteria as

 specified in the protocol.

 ii. Account for all patients and samples. Insure that data

 points are included for every sample for every patient.

 2. Establishing Reference Ranges

 a. Normal individuals:

 Establish a normal reference range with a statistically sufficient

 number of samples from normal persons characterized by age, sex,

 geographical location and any other factors that would influence

 the values obtained. 3,10

 b. Patient groups:

 i. Confirm that the new device detects the percentage of

 positives generally expected for each disease for which the

 device is intended. Use a statistically sufficient number

 of patients characterized by age, sex, geographical

 location, any symptoms of disease, clinical presentation,

 and any other factors that would influence the values

 obtained. 3,10

 ii. False results:

 Patients positive for antinuclear antibodies (ANA) could

 give a false positive result in an IFA assay for the

 detection of thyroid autoantibodies.

 Radioisotopes administered to the patient for diagnostic or

 therapeutic purposes may interfere in some RIA assays.

 Provide reports, if any, of false positive and false

 negative results for each disease as appropriate.

 c. Sample Types Claimed:

 Investigate all sample type(s) claimed in the intended use

 statement unless other data proves that there is no difference

 between them.

IV. LABELING CONSIDERATIONS

The following are additional details for some of the points in the statute

[502(f)(1)] and regulations [21 CFR 809.10(b)].

Package Insert

Include the package insert for the new thyroid autoantibody device. Support

the statements throughout the document with key literature citations.

A. INTENDED USE

 Briefly describe the intended use based on the technology/methodology used

 in the device. Include the following information:

 1. Whether the assay is qualitative, quantitative, or semi-quantitative.

 In order to claim to provide a quantitative result, the calibrators

 or standards must be calibrated to a reference material of known,

 established value.

 2. Test methodology.

 3. Specimen type(s).

 4. Indicate if the device is for use in clinical laboratories and/or

 alternative care sites. The Limitations section should include any

 specific training required for test performance.

 A typical Intended Use statement would be:

 "ABC's \*\*\* test system is a device for the semi-quantitative measurement of

 anti-thyroglobulin antibodies by indirect immunofluorescence in human serum

 to aid in the diagnosis of certain thyroid disorders, such as Hashimoto's

 disease, nontoxic goiter, and Graves' disease."

B. QUALITY CONTROL (QC)

 The package insert should recommend levels of quality control samples and

 their number, matrix type, placement and interpretation to ensure that the

 system meets its performance claims. Include a statement that if controls

 do not behave as expected, assay results are considered invalid and should

 be repeated.

 Controls should be handled in the same manner as patient samples. For

 example, if the patient sample is diluted or titered the control material

 should also be diluted or titered using the same diluent.

C. RESULTS

 Give an adequate description of expected results and interpretation.

 1. IFA

 a. Give a description of the fluorescence for a positive and negative

 result.

 b. List possible staining patterns which may be found. Photographs

 or diagrams may be helpful.

 c. Give pattern description and interpretation as it relates to the

 particular antibodies.

 d. Give instructions for titering positive samples to end point.

 2. IHA

 a. Describe in detail the appearance of a positive and negative

 result.

 b. Give a definition and description of the end point for a positive

 result.

 c. Describe the appearance of a prozone reaction (antibody excess)

 and instruct the user what to do if prozoning is suspected.

 d. Explain the procedure for repeating samples which have results

 above the linearity of the assay.

 3. RIA

 a. Explain the procedure for manual (if applicable) calculation of

 percent bound for each sample and include a sample calculation.

 b. Give instructions for plotting percent bound versus concentration

 of the standard showing examples of typical results (numerical and

 graphical).

 c. Give a brief explanation of how automated calculations are

 performed, e.g., the type of data reduction program used.

 4. ELISA

 a. Explain the procedure for calculating the value of the unknown

 including a sample calculation.

 b. Explain the procedure for repeating samples which are above the

 linearity of the assay. Give instructions for dilution of samples

 including the dilution factor and type of diluent to be used.

D. LIMITATIONS OF THE PROCEDURE

 Include a statement of limitations of the procedure to include the

 following:

 1. A statement that the test result in and of itself is not diagnostic

 for thyroid disease and should be considered in conjunction with

 iodine uptake and other standard thyroid tests and the clinical

 presentation of the patient.

 2. IFA

 a. Explain possible variations between different types of

 fluorescent microscopes.

 b. Give warnings concerning distinguishing the thyroid specific

 cytoplasmic fluorescence from that obtained with

 mitochondrial antibody in primary biliary cirrhosis. If

 mitochondrial antibody is suspected, the distinction can be

 made by running more specific tests for antinuclear

 antibodies.

 c. Multiple antibodies may be present and complicate the

 staining interpretation. Serially diluting the patient

 sample will often aid in distinguishing multiple patterns.

 d. A prozone reaction can appear as a doubtful positive or

 negative because of a small amount of antigen in relation to

 the large amount of antibody present. If prozone is

 suspected, the patient sample should be serially diluted.

 3. IHA

 Give warnings concerning heterophile antibodies and a possible

 prozone reaction

 4. RIA

 a. Give warnings concerning possible interference from

 radioisotopes administered to the patient for diagnostic or

 therapeutic purposes.

 b. Supply instructions for proper disposal of radioactive

 materials.

E. EXPECTED VALUES

 1. The expected value in the normal population is negative. However,

 apparently healthy, asymptomatic individuals (5-10%) may test

 positive for thyroid autoantibodies. The incidence of these

 antibodies increases with increasing age beginning in the seventh

 decade for women and the eighth decade for men.

 2. Thyroid autoantibodies may be present in non-thyroid disorders such

 as pernicious anemia, diabetes mellitus, Addison's disease, and

 Sjogren's syndrome.

 3. Present information showing the incidence or prevalence of each type

 of thyroid autoantibody for each disease state.

From: Division of Clinical Laboratory Devices

 Immunology Branch

Prepared by: Deborah M. Moore, Scientific Reviewer

 February, 1994

V. BIBLIOGRAPHY

 1. Czarnocka B,et al. Purification of the Human Thyroid Peroxidase and

 its Identification as the Microsomal Antigen involved in Autoimmune

 Thyroid Disease. FEBS 190:147 (1985).

 2. Henry J, Immunology and Immunopathology. Clinical Diagnosis and

 Management by Laboratory Methods, 18th ed 1985; WB Saunders Co.,

 Philadelphia, PA.

 3. National Committee for Clinical Laboratory Standards. How to define,

 determine, and utilize reference intervals in the clinical

 laboratory; proposed guideline. Villanova, PA 1991. Order code C28-P.

 4. Vadlamudi SK, Stewart WD, Fugate KJ, Tsakeris TM. Performance

 characteristics for an immunoassay. Scand J Clin Lab Invest

 1991;51:134-138.

 5. Peters T, Westgard JO. Evaluation of methods, Chapter 7 in: Tiets

 NW, editor. Fundamentals of Clinical Chemistry, 3rd ed, 1987: 225-37

 Philadelphia, PA; WB Saunders Co.

 6. National Committee for Clinical Laboratory Standards. Evaluation of

 the linearity of quantitative methods; proposed guideline. 1986

 Order code EP6-P.

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

 8. National Committee for Clinical Laboratory Standards. Evaluation of

 precision performance of clinical chemistry devices - 2nd ed;

 tentative guideline. 1991:1-56. Order code EP5-T2.

 9. National Committee for Clinical Laboratory Standards. User

 comparison of quantitative clinical laboratory methods using patient

 samples; proposed guideline. 1985; 6(1). Order code EP9-P.

 10. Ash KO. Reference Intervals (Normal Ranges): A Challenge to

 Laboratories. Am J. Med Tech 1980; 46:504-11.