Guidance for Industry and for FDA Reviewers/Staff

Guidance on Labeling for Laboratory Tests

Draft Guidance – Not for Implementation

This guidance document is being distributed for comment purposes only.
Draft released for comment on June 24, 1999

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Division of Clinical Laboratory Devices
Office of Device Evaluation

Preface

Public Comment:

For 90 days following the date of publication in the Federal Register of the notice announcing the availability of this guidance, comments and suggestions regarding this document should be submitted to the Docket No. assigned to that notice, Dockets Management Branch, Division of Management Systems and Policy, Office of Human Resources and Management Services, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD 20852.

Additional Copies:

World Wide Web/CDRH home page at http://www.fda.gov/cdrh/ode/1352.pdf or CDRH Facts on Demand at 1-800-899-0381 or 301-827-0111, specify number 1352 when prompted for the document shelf number.

Guidance on Labeling for Laboratory Tests

INTRODUCTION

The evaluation of laboratory test performance should compare a new product's test results to some appropriate and relevant diagnostic benchmark that can be used to correlate results from the new test with the clinical status or condition of individuals/patients for whom the test is intended to be used. Determination of the clinical status of patients whose specimens are used in an evaluation may be based on laboratory and/or clinical endpoints. As a result of some 20 years of review experience, the Division of Clinical Laboratory Devices recognizes two major categories of endpoints for assessing diagnostic performance of new "in vitro diagnostic" assays based on the case definitions used to define the performance of a new test. These categories are (1) Operational Truth or (2) Laboratory Equivalence. Characterization of test performance is important to allow labeling that will clarify the performance of the device for both laboratories and health care givers.

1. **Operational Truth (the "True" diagnostic state as determined by patient clinical status or condition):** Test performance is characterized in terms of direct comparison to the relevant clinical condition or status of individuals/patients evaluated. Definition of the clinical state of a patient may be established using the results of autopsy studies, outcome studies, and/or the use of well-defined diagnostic algorithms. Examples include the diagnosis of myocardial infarction using the WHO standards, the diagnosis of lupus or rheumatoid arthritis using American Rheumatology guidelines, or the diagnosis of H. pylori infections by use of combinations of culture, histology, and urease testing. The case definitions being used as the reference point in determining performance should be clearly referenced and explained either in performance tables and/or supporting text. Sponsors are encouraged to consult with FDA to ensure case definition being chosen for study will meet agency requirements. Ideally, studies of "true" diagnostic state test performance would be performed in the intended use population that includes patients with relevant confounding medical conditions to help determine test performance.

2. **Laboratory Equivalence:** Performance of the new test is characterized in terms of a comparison to a predicate. In the absence of acceptable operational truth, the relevant objective is to demonstrate the ability of the new test to agree with the predicate.

PROPOSED LABELING

All package inserts for laboratory tests should clearly explain how performance has been deduced or determined.

1. In situations where a candidate device is being compared to a "true" diagnostic state, the case definition against which new test results are being compared should be clearly stated and both strengths and limitations likely to result from the selection of this definition should be discussed.

    - Estimates of sensitivity, specificity, and ROC (Receiver Operating Characteristic) Curves[1], along with confidence intervals are appropriate measures of performance and may be presented in labeling.

    - Ideally, these studies would be performed in a population representative of the proposed intended use, which would include individuals/patients with relevant confounding medical conditions to help characterize the robustness of test performance. If disease prevalence precludes naturalistic sampling, it may be desirable to include disproportionately more case positive individuals in the study population than exist in the intended use population.

    - If specimens have been preselected for testing (for example, if case defined positive and negative samples are obtained from archived collections and retrospectively tested with the candidate test), sensitivity and specificity claims may still be appropriate. The labeling should indicate the nature of the specimens studied and the limitations introduced through selective sampling.

    - The effects of prevalence on the usefulness and reliability of the test should be discussed, when appropriate, with performance estimates translated into the hypothetical predictive values for documented frequency of the condition/disease in patient populations that would be tested.

2. In cases where a candidate device is being compared to a predicate, the predicate and conditions under which it is performed should be defined. Conditions of use include operator experience, clinical laboratory facility or other test setting, controls applied, specimen acceptance criteria, etc.

    - A test that has been characterized to a predicate but has not been compared to "true" diagnostic states should be labeled WITHOUT sensitivity or specificity claims. Relative performance may be described in terms of agreement, co-positivity and co-negativity, or using other similar terms.

3. The case definition of positive and negative results being used to characterize a new test should be independent of results derived for testing with the new device and be uniformly applied to all samples.

4. Discrepancy analysis may be performed and described in the submission and package insert if statistically valid techniques are employed. FDA suggests sponsors consult with FDA scientific reviewers or statisticians about appropriate techniques to apply to this analysis.

Reference

1. NCCLS Approved Guideline. GP10-A Assessment of the Clinical Accuracy of Laboratory Tests Using Receiver Operating Characteristic (ROC) Plots. NCCLS, 940 West Valley Road, Suite 1400, Wayne, PA 19087.