Food and Drug Administration (FDA)'s impact on laboratory performance: FDA's perspective

STEVEN GUTMAN

The Division of Clinical Laboratory Devices is responsible for the premarket review of in vitro diagnostic devices (laboratory tests). We currently process >1000 diverse applications per year. New versions of old devices are handled as premarket notifications, so-called 510(k) submissions. The review objective is to establish that the new product is "substantially equivalent" to its predicate. Fundamentally new devices are handled as premarket applications. The review objective is to establish de novo that the product is "safe and effective." A central regulatory issue over the past several years has been the development of a standardized model for scientific review. The Food and Drug Administration contributes to the quality of in vitro diagnostic devices by providing oversight and objective review, by setting thresholds for product safety and effectiveness, and by ensuring that organized data and appropriate labeling is present in support of a device's intended use.

INDEXING TERMS: premarket review • in vitro diagnostic devices • premarket application • premarket notification • 510(k) submissions

THE REGULATORY PROCESS

The Division of Clinical Laboratory Devices (DCLLD) is currently composed of 50 scientists ranging from medical technologists to clinical pathologists who are collectively involved in the premarket review of in vitro diagnostic devices.1 In general we process two types of submissions. New versions of old devices are "cleared" (the specific term used is "cleared") as premarket notifications [1]. Because the regulation cited is the 510(k) regulation in the Code of Federal Regulations, these are referred to as 510(k) submissions. Fundamentally new devices are "approved" (the specific term used is "approved") as premarket applications (PMAs) [2].

510(k)s

The largest volume of review activity for FDA involves 510(k) submissions. DCLLD currently handles ~1000 510(k)s per year. The operative term in 510(k) review is "substantially equivalent." As the term implies, evaluation is directed at determining if the new device is equivalent to its identified predicate. Review of most 510(k) submissions is straightforward and based on an analysis of the fundamental operating principles of a test including accuracy, precision, analytical sensitivity, and analytical specificity. Limitations in the review process are also quite straightforward. 510(k) review is entirely a paper review; FDA does not perform direct laboratory evaluation on any of these devices and the agency therefore has no hands-on experience with the vast majority of products under review. In addition, the agency is continually challenged by the need to determine appropriate minimum standards for the substantial equivalence decision. The FDA goal for 510(k) submissions is to process these with a total review time of 90 days. Reviews are generally conducted internally, although occasionally a product will be sent to outside experts employed as "special government employees" for outside grounding and input as a so-called "home-work" assignment. In very rare instances a formal public panel meeting may be held to deal with an unusual or problematic 510(k) submission.

PMAs

DCLLD reviews generally from 1 to 2 dozen new PMAs in the course of a year. The operative term for a PMA review is "safety and effectiveness." Since no predicate can be defined, it is necessary de novo to establish that the product is "safe and effective." Since passage of the Safe Medical and Devices Act of 1990, FDA has in fact taken a broader interest in the safety and effectiveness of all devices. We now require even with 510(k) submissions either a summary of safety and effectiveness or a statement that such information about the submission will be

1 Nonstandard abbreviations: DCLLD, Division of Clinical Laboratory Devices; PMA, premarket applications; FDA, Food and Drug Administration; and GAO, General Accounting Office.

Received November 29, 1995; accepted January 15, 1996.
made available upon request. As a result of this amendment, for all PMAs and for many 510(k)s, FDA since 1990 has increased data requirements, which include information not only on the analytical performance of a device but on clinical performance as well, including clinical or diagnostic sensitivity, clinical or diagnostic specificity, and in some cases information on the expected predictive values of testing. Again, the limitations in the review process are quite obvious. In evaluating new products there is often a lack of a laboratory or clinical “gold standard” against which to judge performance, and adjustments must be made in performance yardsticks. Data collection to establish safety and effectiveness is often complicated by either overt or latent bias. Finally, as with 510(k) submissions, determining acceptable limits for minimum performance can be difficult and challenging. The FDA goal for PMAs is a total review time of 180 days. Most new PMAs are subject to external review by outside experts. As PMAs become more familiar through repeat submissions of the same type, review intensity can be decreased; panel input is, for example, not required for the third or fourth of a kind of identical PMA. Regulations are being developed to provide for increased flexibility for reducing review of these products.

LABELING OF IN VITRO DEVICES

In vitro devices are unique in the area of device regulation in that they have their own labeling regulations for the product insert—21 CFR 809.10. These regulations clearly indicate the information required to support device labeling and submissions. The labeling regulation is divided into 15 separate components, which include: (a) the proprietary name and established name of the product, (b) intended use, (c) summary and explanation of the test, (d) the principle of the procedure, (e) information on reagents when appropriate, (f) information on instruments when appropriate, (g) information on specimen collection and preparation when appropriate, (h) procedures, (i) results, (j) limitations of the procedures, (k) expected values, (l) specific performance characteristics, (m) bibliography, (n) name and place of business, and (o) date of the package insert.

Of these various elements the most important is evaluation of the intended use and the related indications for use. The intended use and indications for use of a product will determine the type of review, the questions likely to be raised, and the data requirements that will be required in the course of review.

DEVELOPMENT OF A SCIENTIFIC REVIEW MODEL

A central concern of the Office of Device Evaluation over the past several years has been development of a strong but pragmatic scientific model to frame our review. DCLD believes that while there is not one path to truth in terms of the development of information to support an in vitro diagnostic product, there are several basic tenets for good science. These include the need for:

1) Up-front study design. All submissions, whether simple or complex, require an established up-front design. In some cases, all that is needed is referencing a standard evaluative methodology or set of voluntary standards; in others, there may be a need for development of extensive and complex protocols with carefully formulated hypotheses.
2) Careful and meticulous collection and reporting of data.
3) Interpretation of results with sound, preferably referenceable, statistical techniques.

There is little excuse for poor-quality submissions for in vitro diagnostic devices. In the design of devices and preparation of supporting data to be provided FDA, sponsors can draw from a rich scientific literature, numerous voluntary standards, and a growing number of FDA guidances, all of which serve as guideposts to well-grounded and high-quality submissions.

Individual product review obviously varies by product line, type, and intended use.

QUANTITATIVE TESTS

For a quantitative test 510(k) review, minimum requirements are usually targeted at: (a) information on bias or, if possible, accuracy, comparing the new method by linear regression with a reference and (or) a predicate method; (b) information on precision, ideally studied with an ANOVA to allow comprehensive assessment of components of variation; and (c) experiments designed to evaluate analytical specificity and, if appropriate, sensitivity.

QUALITATIVE TESTS

For a qualitative test, 510(k) review usually at a minimum requires information on cutoff points established and discrimination or equivocal zones present in the test system.

For submissions involving new analytes, matrices, and (or) novel technologies, clinical as well as analytical data are required to allow test performance to be analyzed within a clinical framework.

FDA review has historically not required: (a) outcome data: The presumption that clinical information is useful usually suffices to support the review process; and (b) prospective clinical studies: In only a handful of cases are prospective clinical studies required; usually concurrent sample analysis—sometimes framed in demographic or clinical data—will support product review and timely clearance or approval.

FDA review does require meticulous attention to detail in data collection and presentation; the better this is done by a sponsor, the more likely there will be timely review and approval.

In short, general DCLD review caveats remain: (a) the need for up-front clinical design, (b) the need for meticulous collection of data, (c) the need for use of valid statistics, and (d) the requirement for truth in labeling.

HISTORY OF FDA REVIEW

FDA review of in vitro diagnostic devices has been in place for >20 years and predates implementation of the Safe Medical Device Amendments (1976). Scientists in the Centers for Drugs and Biologics as long ago as the early 1970s were concerned with the safety and effectiveness of in vitro diagnostic devices. By using creative and imaginative schemes for introducing regulation, they made efforts to classify these products as drugs or biologics. The result for a small cadre of tests including tumor
markers, antimicrobial susceptibility testing, and pregnancy tests was efforts to undertake regulation.

The seminal work leading to the inclusion of in vitro diagnostic devices as a component of medical device review was a not particularly well-known device study published in 1975 [3]. This study, commissioned by the General Accounting Office (GAO), was performed by the Centers for Disease Control (CDC) and involved assessment of >44 kits representing 8 common analytes. Failure to meet one or more laboratory test or labeling standard was reported in 32 of the studied kits.

The Medical Device Amendments to the Food, Drug, and Cosmetics Act were passed in 1976. Extensive regulations were promulgated and a comprehensive program developed to provide for premarket review, manufacturing review, and postmarketing surveillance for medical devices.

VALUE OF FDA REVIEW

The bottom line for in vitro diagnostic devices is that DCLD currently reviews all commercially marketed laboratory devices with the clear goal to ensure product safety and effectiveness. We believe there is value added as a result of our review in at least three ways: (a) we provide for oversight and objective review of new laboratory tests, (b) we set and maintain minimum thresholds for product safety and effectiveness, and (c) we ensure that organized data and appropriate labeling are provided to the users in support of a device's intended use.

In a perfect world perhaps this review might not be necessary. In the highly competitive world of in vitro diagnostic devices, this degree of oversight is a protection to laboratories and the patients they serve in helping to ensure quality devices from (a) new companies unfamiliar with good science and regulatory requirements and (b) old companies that for various reasons exhibit quality drifts.

The outcome of FDA oversight is difficult to measure. The GAO report has not been replicated. The main mechanism FDA has for tracking product successes and failures in the general marketplace is through the proficiency testing available. Certainly there is no evidence in 1995 of parallels to the 75% failure rate reported by CDC in 1975, but, contrary to popular belief, problems are identified in the use of in vitro diagnostic devices, ranging from reported and unreported product recalls to medical device reports of both morbidity and mortality. Last year 2000 medical device malfunction reports for DCLD products were filed. Although most centered on two product lines, home glucose and pregnancy testing, many other problem products were identified. The medical literature has a small but interesting and growing number of reports of product failures. Last April a problem with positive IgM rubeola tests was reported in the New England Journal of Medicine [4]. The authors took FDA to task for failing in its job. They were obviously unaware that for historic reasons unclear to us, IgM rubeola tests are exempt, i.e., not subject to FDA review. FDA does its part in identifying and trying to head off product problems. For example, in October of 1995, FDA issued a product recall and consumer alert because of failures in test strips for home glucose monitors.

DOCUMENTING RESULTS OF THE REVIEW PROCESS

Part of the FDA problem in documenting the good work that I believe FDA does is the fact that much of this work is invisible. It is estimated that only ~2% of products submitted to FDA are found to be unsatisfactory and are either rejected by the agency or voluntarily withdrawn. Products that have been rejected or withdrawn are considered proprietary, and information about them is not available through the usual freedom-of-information reporting systems. The dazzling array of device design and performance failures that lead to rejection are, from the standpoint of the clinical laboratory and the public, simply hidden.

Also, the impact of FDA review is largely undocumented for cleared products. Over 15% of 510(k) submissions processed raise sufficient concerns to be issued deficiency letters. This means two or more review cycles before clearance. There are three common types of deficiencies identified: inadequate performance data, incorrect or incomplete labeling, and fundamental design problems. FDA reviewers, through the questions and suggestions raised in their deficiency letters, attempt to improve and refine products in these three areas.

FDA has never performed an audit of the impact of this review because frankly we did not appreciate that appropriate methodology for such a study existed. We now realize that this is not the case. In July of 1994 the editors of the Annals of Internal Medicine published a fascinating article entitled "Manuscript Quality Before and After Peer Review and Editing at Annals of Internal Medicine [5]." They reported significant improvement in 34 quality items as a result of the peer review process. Of particular interest to us at the FDA were the marked improvements and changes in the study limitations, generalizations, and conclusions reached. These represent areas of review that some would suggest we routinely obsess over.

As scientists engaged full time in this type of review activity, we believe that our review contribution is comparable with that reported in the above-mentioned article, albeit documentation of this process remains to be performed. Whereas the public health benefits of FDA review of performance and labeling are difficult to assess, the up-front requirement for minimum standards in instrumentation and laboratory reagents is consistent with the modern pursuit of a controlled laboratory environment to assist in total quality management of laboratories.

THE FDA BACKLOG—PROBLEMS IN THE APPLICATION OF GOOD SCIENCE TO FDA REVIEWS

The application of a heightened scientific review model coupled with the increasing complexity of products developed and submitted to FDA by industry has been intellectually exciting but administratively problematic. We have over the past several years experienced an imbalance between work force and work load. The result has been a backlog in products.

We take this backlog seriously. FDA recognizes that it has a twofold responsibility: (a) to keep poor products out of the commercial marketplace, and (b) to assist manufacturers in bringing innovative and good products to the marketplace in a timely fashion.

We have been struggling over the past year to bring the total review times within the FDA goal of 90 days for routine 510(k)
submissions, with some success. The good news is that our product backlog—defined by FDA as products under active FDA review for >90 days (either on initial evaluation or during a reevaluation in response to FDA questions raised)—has dropped from 90 products as recently as 1 year ago to zero. The bad news is that during this period, total review time has remained fairly constant at ~135 days. Part of this lingering delay represents the residual challenge of dealing with a large work load with a limited work force. It is worth noting that in product review, quality submissions do pay off. In a small administrative pilot study conducted by DCLD earlier this year, we isolated 10 submissions that closely followed well-established guidance documents. Average review time for these was only 70 days.

Although our review times have markedly improved, given our current resources and the increasing number and complexity of premarket submissions, it is unlikely that a total review time of 90 days can be achieved. As a result we are looking for tools to help bring balance to our program.

A number of office-wide and DCLD-specific programs are being explored, some of which cut across review of all medical devices and others which are specific to in vitro diagnostic products. FDA will be seeking outside input from both healthcare professionals and manufacturers as it seeks to improve and refocus the way it conducts premarket review. Although we expect that the results of these efforts will be changes in the traditional manner in which we do handle submissions, it is our clear goal to maintain sound science as a basis for future regulatory changes.

References