THE REAL WORLD - REGULATIONS AND ETHICS

Morton K. Schwartz PhD

February 10, 1992

DR. FIDELINO: Now, The Real World: Regulations and Ethics and here's Dr. Schwartz.

DR. SCHWARTZ: Thank you very much. It's a delight to look out on the audience, after six intensive hours, and see that there are practically no seats that have been vacated.

My assignment is one that I'm not sure I'm prepared to undertake, but I'm going to try to talk to you about where we really are with tumor markers, and what we can do and what we can't do, and perhaps even more particularly, what we should do. May I have the first slide please?

You all know the potential clinical uses of tumor markers in screening, diagnosis, evaluating prognosis, tumor staging, detecting tumor recurrence, and monitoring therapy, and you also have been told about sensitivity, specificity, the relationship to tumor burden, and the need for simple, reliable methods that are cost-effective.

Now, I'd like to talk about the regulations involved with performing these tests, the ethics that may or may not be involved in doing them, and our relationships with patients, with the government, with the insurance industry, and with all those who are involved in evaluating the tests and applying them in clinical practice.

The Medical Devices Amendment was published in 1976, when the Food & Drug Administration regulations were amended to include devices, basically to protect the public. I'm also going to talk about CLIA 1988 (Clinical Laboratory Improvement Act), which was passed by Congress in 1988 and regulates laboratory operation. It really has not yet been implemented, although it took effect on January 1st of this year. Last, we must include Medicare regulations, which relate to the laboratory. During our discussion, you may bring up other regulations. I was thinking of medical necessity programs, which recommend which test array is reimbursable for a particular disease for health insurance purposes.

We all have our own definitions of ethics. Ethics, in a sense, are in the eyes of the beholder. What is ethical to you may not be ethical to me, but the pressures of ethics are put upon us by at least two groups. All of our professional societies are deeply involved in ethics, and have developed codes of ethics, related to what we do. And then, consumer organizations certainly propose ethics. We read what consumer-related organizations think we should do, and what they think we should charge, and what they think we should provide. They are very concerned with ethics.

I have not been inclusive. This is the kind of listing one thinks of when one sits on a Sunday afternoon and tries to put his thoughts together. There is no doubt that laboratory medicine is the most highly regulated of all the areas of medicine. There is no place else where an inspector can arrive with specimens in the middle of the night, provide them to the technologists, ask them to perform an assay and, based on that assay, can certify or decertify that laboratory. There is no place else that I know of where regulations for personnel have been described for every single level of staff, from technicians through technologists through supervisors and on up through the professional directors. So we're highly regulated, and we still have not reached perfection.

The FDA has recently achieved a certain degree of notoriety. The FDA was mandated, through the Medical Devices Act, to evaluate medical devices, diagnostic tests, equipment, and reagents to achieve safety and protect the public. When the Medical Devices Act came into being, there were many unreliable test kits. I served on the initial FDA panels, and was chairman of the Clinical Chemistry/Hematology Panel for many years, and I understood and sympathized with the FDA's mandate.

In its evaluation of tumor marker tests the FDA was required to take a different approach than for most analytes. It was stated that tumor markers must be
considered as class 3. (Assays were described in the regulations as class 1, class 2 and class 3, with class 3 the most stringent of the categories, requiring large clinical trials, similar to drugs, before they could be put on the market.) The FDA has taken the stance that it is not only the test reliability that is subject to their control, but also the cost-effectiveness of the test.

Over this past weekend, I read the legislation in as great detail as one can read a 60-page legalistic document, and I could not the statement that tumor markers are to be considered in a special category. But in any case, on July 9th of this past year the FDA issued a cease & desist order against one of the large biotechnology companies, namely, Centocor, for marketing several of the mucinous antigens without FDA approval, most importantly, CA15-3 and CA19-9, and also others like P-glycoprotein, which someone may be involved in multidrug resistance. All of a sudden reagents for these tests dried up because they were sold for investigational use only or for research. On October 19th of this past year, the FDA redefined "investigational or research only." They said that clinical use of unapproved diagnostic tests "may expose unknowing patients to questionable diagnosis." And, secondly, research use only "will not be used to determine the safety or effectiveness of a test for clinical practice or for clinical diagnostic purposes." Research use apparently has been eliminated, to evaluate a tumor marker. Investigational use is subject to strict controls. We can only evaluate a marker if we have institutional review board (IRB) approval, and if the patients are on a true clinical protocol with informed consent for that assay. The Clinical Improvement Act (CLIA) of 1988 states that laboratories may only use FDA approved tests. So, many of the markers that we are talking about are no longer available. CLIA regulations include insurance laboratories. Any laboratory that is dealing with human material is subject to these regulations. Therefore, it'll be very difficult for us to use non-FDA approved tumor marker tests.

Very few tumor marker tests have been approved by the FDA. CEA has been approved; alpha-fetoprotein has been approved; PSA has been approved; CA125 has been approved, but only for limited use. FDA approval of CA125 is only to monitor the second look surgery that Dr. Bates and Dr. Fritsche spoke of. You might be interested to know that chorionic gonadotropin had never been approved by the FDA as a tumor marker; it's approved for pregnancy testing, and that is all.

Any test that was available and in use prior to 1976 is grandfathered. We have the problem that tests used successfully in Asia and in Europe are not available to us in the United States, even though they may have some practical value. What are the ethics of this mandate preventing us from testing? I tried to convince some of the manufacturers that they should take the lead in discussing this with the FDA, but it seemed to me that most of the diagnostic companies are now part of large pharmaceutical companies, most of whom had FDA applications for therapeutic drugs which were worth hundreds of millions of dollars in value, whereas the diagnostic tests are only a market of several million dollars. They weren't about to go and shake the boat.

The FDA has also taken the stance in evaluating tests that a test must not only be safe, it also must be cost-effective. That is the area of concern which has led to the difficulty of certain tests in getting through the FDA. The definition of cost-effective is in the eyes of the beholder. There can be four definitions of cost-effective. One, is cost-effective when it's truly cost-effective. It's the least costly of several alternatives, and I suppose that's the bureaucratic definition.

The one extreme is effective with no consideration of monetary cost, and that's the one to use when you're treating me for my disease. I don't care what you spend, as long as you help me.

What the FDA has really wanted to see is cost savings with equal or better health outcome. So, in order for a manufacturer to make the point that a tumor marker test should be accepted, they've had to try to make the point that that test will have cost savings with equal or better outcomes, or there will be a slight additional cost but there will be great additional benefit. This is the definition that will require judgement on what society will pay to achieve that added benefit. Is dialysis in end-stage renal disease a cost-effective process? Certainly to the individual with end-stage renal disease it is, but will our society pay to achieve that? If any of the markers that we spoke of today can replace some expensive technique, it would be acceptable. That was the basis for the acceptance of CA125. Dr. Fritsche made the case that second-look surgery could be eliminated if CA125 was used. If you look at the literature, and put all the cases together, of about 200 cases that are reported there is only one patient who had an elevated CA125 in whom cancer was not found. The other side of that coin is, resectability of pancreatic cancer and CA19-9 elevations. Using CA19-9, CEA, and another marker called CA195, we were able to predict which patients are resectable. However, we were never able to convince surgeons that they should even consider this approach, because their answer was, "Most patients with pancreatic cancer always need 'palliative surgery.'" Now, whether they need palliative or economic
surgery, I don’t know. But cost-effectiveness becomes a very important consideration in how the FDA views a tumor marker.

I ran across this and I thought you’d be interested in sharing it with me. This is the perception of the public on who pays for the high cost of medical care. This study is called "Pointing the Finger" was a survey of 800 consumers: laboratory directors, providers of care, insurers, etc., asking them three questions as to where they thought responsibility for the cost of medical care resided. Physicians never said they were responsible for the cost of medical care. There were a lot of other reasons listed, but 30% of physicians said the cost is related to new technology and equipment, and 14% say that it’s due to the government. Hospitals, again, never said the physicians were responsible. They said the government was responsible for the cost of medical care, and 15% said new technology. I would estimate 20-30% of hospital costs are due to the bureaucracy — the quality assurance needs, the regulatory needs, completing all the many insurance forms, lead to a large percentage of total cost. When the question was asked of employers and consumers, none said it was new technology and equipment; all of them blamed the doctors and the government. So, you see, it’s like the three blind men feeling the elephant’s leg and defining what it is. It’s all in the eyes of the viewer, as to why costs are high. There’s no doubt costs are high. I saw a figure recently that we spend something like 8% of GNP on health costs, but only 6% on education. The time has come for us to reverse those figures. In any case, 86% of consumers and 79% of employers stated that hospitals charge more than costs.

So, as we and the FDA evaluate markers, we must keep in mind that clinical utility must go into the equation. Clinical utility is the unique and profitable use of a technique or procedure in the management of the patient.

We talked tangentially I think, about the future in genetic testing. Genetic testing is going to give us a great deal of concern as we measure risk factors. There are many problems. But we don’t know where these tests will be performed. We don’t know whether a new area of laboratory medicine will evolve. We don’t know where, how and by whom the quality assurance programs will be developed. We don’t know if proficiency testing is possible, and if it is, who will do it and where will the material come from? These are all very serious problems. Should we in the laboratory just stand aloof and let all this go around us, or should we be involved in the ethical and legal implications of genetic testing?

If we’re going to determine a genetic risk factor for colon cancer 20-30 years before it happens, what are the legal and ethical implications of that? Genetic testing is broader than just the cancer field. Our professional societies should all come to grips with these problems. We don’t want to wait until the bureaucratic organizations tell us how it will be done.

As we consider genetic testing, we should consider the National Genome Project and what will happen when the whole genome is worked out and we know the genetic aspects for every single disease. I thought you might like to read comments made by Nancy Wexler, Assoc. Prof. Clinical Neuropsychology at Columbia, at the 1989 Human Genome Project Conference. I’m quoting her as precisely as I can, but this is secondhand because it was in an article: "Can insurance companies require that genetic testing be done in order for a person to get coverage?" "Can they charge exorbitant fees for high risk individuals?" "Is legislation needed to ensure confidentiality?" These are the kinds of questions that are being made as the National Genome Project is being developed, and I think they’re very relevant to the people in this room. Obviously there is fear that once these tests are available, you will take the bridle and run with it. I have more faith than that. I think that these things will be worked out well.

CLIA, the Clinical Laboratory Improvement Act of 1988, is the legislation for all laboratories that are dealing with material for human testing. Basically, this legislation establishes quality assurance for all assays. It prescribes how to do it, what proficiency limits are, and, for the first time, there is a laboratory program is punitive as well as educational. The Proficiency Program will send material for proficiency testing and if you fail on two occasions, you can be legally precluded from performing that assay. This has awesome and frightening implications.

Quality assurance for tumor markers is really no different from quality assurance for any other group of analytes. It’s the step necessary to insure the reliability, validity and efficacy of tumor marker tests and use of the data in the overall care and management of the patient. That’s it, in a nutshell, but quality assurance is a full cycle, with the patient at the top of the loop. It goes from test ordering, availability of the test, how the patient is prepared for the test, what the collection and delivery of the specimen is, what the turnaround time of a result is, how the data is handled, how the results are interpreted, and what the clinical action is. A full circle, and, believe it or not, CLIA addresses all of these.
I want to talk a little about what we don’t know about most tumor markers. There are pre-analytical problems: that’s the ordering of tests, the collecting and the delivery of the specimen, and the preparation of the specimen. What tests should be ordered? You’ve heard a lot about that today. I can only propose that we begin to decide what tests are needed before we’re buried under an avalanche of new tests. We must try to develop precise and definite research programs that are not overwhelming, that are not too expensive, that allow us to answer these questions rapidly. We must determine how we collect the specimen, how we deliver it, and how we prepare the specimen for analysis.

I looked through the brochures of every single tumor marker I have, and they all have the same statement on the collection and delivery of the specimen. They say, approximately, that the specimen may be maintained at 4 for 24 hours, and then it must be immediately frozen. (This is serum. Please don’t freeze red cells or white cells.) Every one of them has the same boilerplate statement that they must be frozen if they’re not to be analyzed within 24 hours.

Well, the fact is that most of these are glycoproteins, and are very stable. You can do most anything to them and they remain stable. CEA is stable for a very long time. For PSA, there is data indicating that from 4-5 days of room temperature has no effect on the PSA. There has not been a very thorough study over a long period of time because in a laboratory such as ours and Dr. Fritsch’s, we freeze. We have no reason to keep them four or five days. But I do think that all of these things are sufficiently stable, except for one caveat. If there’s a lot of bacteria in the specimen, obviously, you’ll begin to have a problem with it. That can be the problem in urine specimens. Years ago urinary lactic dehydrogenase was proposed as the very best test for urinary tract cancer. We found that urinary lactic dehydrogenase was really just a very expensive white cell assay in urine specimens.

Urine specimens need to be made sterile if they’re going to be shipped, or if they’re infectious. That requires a little bit of work, to see if one of the preservative agents could not maintain the specimens for shipment and not interfere with the assay. Why is the conservative statement in the brochure? Because, under FDA regulations the brochure of a kit is the label of the kit, and the manufacturer is responsible for every single statement in that label. He must hold himself accountable for any statements that are made that are not held up in fact. He obviously is going to take the way out that presents the least liability for him. And, that is to have a statement requiring minimal storage at room temperature before freezing. There are actually some analytes that can be destroyed by freezing. If LDH isoenzymes are your marker, and you freeze them and thaw them, you will destroy them. Under the FDA regulations, and JCAH and HCFA regulations, if your laboratory changes any statement in the brochure supplied with the kit, you must document the change, and you must have that documentation in a separate notebook where an inspector can come and examine it.

Certainly storage is important, but most of the markers we have discussed can be kept for a number of days. They cannot be kept for very long on a radiator. I had experience once with fecal occult blood testing in a public program in Arizona. Specimens were sent by mail, and the temperature in a mailbox in Phoenix was able to destroy the specimen.

An important consideration is which assay method? I told you that for one analyte there may be two or three methods. Can we standardize them and calibrate to give similar values? Dr. Rose showed the tortuous path to purification. We have very little standardized material for use in these assays. We can use what is known as a sandwich immunoassay, in which we don’t need purified material. In the United Kingdom labs are now beginning to do proficiency testing on a World Health Organization standard for CEA. The standard is material obtained by scooping metastatic colon cancer from the liver of an autopsy patient. After simple purification 2 liquets were put in small vials and frozen. They are using the material and have declared that the contents of every vial is 100 international units of CEA. That is the World Health Organization standard for CEA. So, you can see the problems we have.

HCG should be the most standardized procedure of all. We’ve been using it now for many, many years. Probably the first screen for testicular cancer was done by Zodeck in 1928, when he was evaluating his rabbit test for pregnancy. He collected urine from all of his technologists, and found a male technologist who had an elevation. The technologist who did the assays was castigated. How can a man have a positive pregnancy test? In his memoirs, many years later, Zodeck apologized to that young man, who died of testicular cancer. For the HCG kits available on the market there are four World Health Organization standards used, all approved by the FDA. The factors for conversion range from 5 miu/nanogram to 16 miu/nanogram. That means if you have your assay done with different kits, you’re going to get different numbers. If you’re monitoring a patient, and today that patient is in New York and tomorrow in Washington, and two different assays are used, you’re going to get different numbers.
There is a company called Biorad that markets excellent control materials for tumor markers. In their brochure they list what you should get for each assay for each method. Just for HCG there are about 18 companies listed who sell an FDA-approved HCG kit. Results are 2½ with this one, 11 with this one, 14 with this one. One can understand why there is so little confidence in the laboratories' ability to do things right.

CLIA requires us to do proficiency testing. There has been practically no proficiency testing of any tumor marker in the United States. Our Italian colleagues are far ahead of us in this regard, for better or for worse. This is data on 4 tumor markers, none of which we really discussed in any depth. One is a general marker, tissue polypeptide antigen. It's one of the most extensively studied general marker. It's widely used in Europe, but not used at all in the United States. These 50 Italian laboratories reported these results over a period of 2 years. The coefficients of variation ranged from 23½% to 38½%, etc., for the various laboratories. CA125 values ranged from 10½ to 65. TPA from 27 to 426, etc. If you're monitoring a patient and you're getting this kind of variation, it makes the test almost worthless. It is essential for regulatory agencies or professional societies to introduce proficiency testing. This is another Italian study of five laboratories which did CA125 assays for 17-32 months, with 38 to 104 assays. The coefficients of variation vary greatly from laboratory to laboratory, as well as within each laboratory.

You might say, "Well you guys can't do these tests correctly." Why are we so concerned about them? I think some of us can do them right, and I think our responsibility is to insure that everyone does them right. The question is, "What will the FDA do about this? And what will the CLIA do?"

In CLIA, targets are given for proficiency of an analyte. CLIA mandates that you meet an accuracy target with a standard deviation. We have examined several PSA assays and compared them to the Hybertech PSA method, which is our standard method. They all give different results. The Diagnostic Products Company product has a bias against Hybertech of about 27%. The Procheck assay can give a value that's twice as high as a Hybertech number: twice as high.

Hybertech has been approved by the FDA, as has the Imex method from Abbott. Imex is an automated instrument. The methods have a correlation coefficient of 0.99 and you obtain almost the same number. Abbott has made the claim that their method can detect lower amounts of PSA and is better for determining whether all the tumor has been removed. There is no clinical data that I know of that supports the contention that the lower numbers are clinically superior. There is no doubt that in removing the tumor, you would like to reach a level that is "female." Theoretically, PSA is not found in female serum and you would like to reach that level.

I think Dr. Bates mentioned neuron-specific enolase. No one else has talked about it. If you're interested in neuroblastoma or small cell carcinoma of the lung, it is an excellent marker. There are two methods, Pharmacia and Roche; values obtained with one are double the other. One gives you a normal value of 12 and the other gives you a value of 4½. We hope that regulation will bring about some control of those things.

I really have great difficulty with a discussion of ethics. I think that ethics will come when we understand more about the significance of tumor markers. It seems to me we have a responsibility to the patient, to the physician who is involved with that patient, as well as to the industry, to make sure that we report abnormal results that in the proper way. Even if the abnormality is found as part of an insurance application, an individual must be treated for what may be a devastating or life-threatening disease. On the other hand, we must insure that we have not destroyed this individual from a psychological point of view by reporting an abnormal value that cannot be interpreted.

I would like to recommend that anyone who is involved in tumor markers report results as ratios to the normal value. In other words, your value to what is the normal value in your laboratory, and propose a ratio. That would eliminate differences between methods. Just report a ratio, and 1.00 is the upper limit of normal. Thank you. (applause)