Impact of Pharmacogenomics on Prescription Drug Labeling

Lawrence J. Lesko, Ph.D.
Director, Office of Clinical Pharmacology and Biopharmaceutics
Center for Drug Evaluation and Research
Food and Drug Administration

DR. McCABE: In addition to requesting presentations on the FDA's role in the regulation of genetic technologies and its efforts to enhance oversight of genetic tests, we also requested a briefing on how the agency is addressing pharmacogenomics and is -- can we please have the doors closed in the back and have the conversation cease? We also requested a briefing on how the agency is addressing pharmacogenomics and its potential to enhance drug development.

Dr. Lawrence Lesko is head of the Office of Clinical Pharmacology and Biopharmaceutics at the FDA's Center for Drug Evaluation and Research.

Dr. Lesko, thank you very much for being with us today and please proceed.

DR. LESKO: Thank you, Dr. McCabe, and good morning, everyone. I'm going to try to give you a perspective on pharmacogenomics and the FDA drug review process, and in particular some of the impact that this science has had on product labeling in the recent years.

The perspective is based upon what we've seen at the Center in terms of submissions either of INDs or NDAs. That's not to say there isn't a lot more going on in drug development that we're not aware of because companies have either not submitted it to us or they consider it to be some sort of exploratory experiment and they're waiting to see the outcomes of that information and so on.

But let me try to give you a sense of what is going on. Let me start off by saying when we use the terminology, I'll be talking about pharmacogenomics in a very broad way and I'll be referring to hereditary differences in gene expression profiles at the RNA level, although this could be at the mRNA level or protein level, or in nucleotide sequences at the DNA level. The purpose is to better understand variability in disease phenotypes, disease progression, and dose response. The data itself we feel can be used to either select a drug for a particular patient or select a dose for a particular patient.

Now, the two types of data that we have experience with in submissions include microarrays, which basically I think of as quantitative gene expression profiling at the RNA level using target tissue or surrogate tissue. As you're well aware, this could be from a host, a patient, it can also be from a tumor, or it could be from a pathogen. In fact, most of the information we've seen is from the latter two categories, as opposed to the patient.

The goals that companies have in their protocols when they're using gene expression are any one of a number of the ones I've listed on the slide. But basically the ultimate goal is to identify a panel of biomarkers that can be used in a predictive way. That can be, for example, to diagnose a disease or a subtype of a disease on a molecular basis, to monitor disease progression as in a carcinoma, assess severity, predict clinical outcome a priori to see if a patient is going to be a responder, to give a drug and then look at gene expression changes dynamically as a function of drug response, and the ultimate goal if all works out is to develop a diagnostic or drug response predictive test.

The other technology that we see quite often is genotyping, and this comes in several varieties. It comes in whole genome scans when there's attempt to develop a hypothesis that's related to genotyping and clinical outcome, or it may be a candidate sequence that a company might look at using blood or some tissue sample as a surrogate. Basically, the goal of this research is to identify one or more single nucleotide polymorphisms or alleles, sometimes haplotypes, and this category includes what I would call pharmacogenetic tests, the commonly validated variants of drug metabolism genes which have been around for a good 40 years.

Sometimes it's a custom set of SNPs that's related to a specific issue in safety or efficacy, such as hypersensitivity to abacavir. Oftentimes this technique is used in population analysis to see if one could distinguish between responders or non-responders retrospectively. Sometimes this technology is used to include or exclude patients from treatment based upon what's known between the genotype and phenotype association. Sometimes it's used to guide dose selection a priori to improve the risk/benefit of a particular drug.

Now, one of the problems with pharmacogenomics has been the regulatory pathway for sponsors to submit the data, and at least the sponsors, from what they've told us, this pathway is unclear. Beginning about two years ago, we frequently would get questions from sponsors about what data needs to be submitted to the FDA. Because most of this data is exploratory and not suitable for regulatory decisions -- it's novel, it's new -- companies were unclear whether we want to see that data or not. The exception was the drug metabolism genotypes which are well established in the scientific community and there's a lot of public information on it.

Another question comes up, what formats can be used for submission? This is all evolving. Standardization of assays, standardization of reports, the level of detail that comes in is not clear. There are not standards that have been evolved. Then finally the companies would say what are you going to do with this data in decisionmaking? Are you going to hold up a clinical trial? Are you going to ask for more data? Are you going to interpret it incorrectly? There's this kind of uncertainty, and that of course is all dependent on the validity of what they're submitting. So there are a lot of questions related to pharmacogenomics.

Part of the problem was the regulations that were written 30 years ago did not think about pharmacogenomics, and as a result the regulations require interpretation in light of the new science. This is the current regulations that relate to the submission of genomic data during the IND phase, and you can see what I've highlighted in italics there. "Data should be submitted on the basis of which the sponsor has concluded that it's reasonably safe to conduct a proposed clinical investigation." What does that mean? What does "on the basis of which the sponsor has concluded" if much of this information is exploratory, so it's a gray area?

When you move from INDs to NDAs, submission of pharmacogenomic data, again there's regulations that dictate what should be done, but they're not very clear. As this quote indicates, pertinent to the evaluation of the safety and efficacy information in application. So "pertinent" is open to interpretation on the part of a sponsor. How pertinent do we mean? Very pertinent, or what?

It was the regulations that created some problems in interpretation, so we had a public workshop on this question, on those three questions actually, back in May of 2002, and at that workshop the agency made a proposal to develop a regulatory pathway within the construct of the current regulations to facilitate the advancement of pharmacogenomics in drug development. We called it at the time a safe harbor, and the safe harbor was intended to get sponsors to submit exploratory genomic data to the agency for the purposes of increasing our understanding of the data, looking

at the data, learning from it, and using that information to develop good regulatory standards.

That workshop led to an initiative over the past year whereby we developed a guidance for industry we call "Pharmacogenomic Data Submissions." It will be out in the public domain at the end of this month, and later on in November we're going to have a public workshop to get comments on this guidance primarily from the pharmaceutical industry and others that would be interested in it. So this guidance, then, we hope will clarify the situation and address some of the questions that companies had. So we hope it will address questions like this: When is a sponsor required to submit genomic data to the FDA?

We have three general principles that address the question. Whenever a sponsor uses genomic data in decisionmaking in animal studies or human trials. For example, they may use this data to include certain people or exclude certain people. Maybe this data will be used to support a claim by the sponsor related to safety or efficacy or dosing, use a different dose in a poor metabolizer genotype. That data we need to see. It supports a claim. Finally, if the exploratory data ends up providing information or recommended uses of genomics in product labels, we need to see it and evaluate it.

So these are the general principles of this guidance when it comes out, and hopefully this kind of framework will enable a sponsor to make the decision about what ought to be submitted.

We do talk about level of validity of biomarkers, which are a byproduct of genomics in this guidance. We talk about valid biomarkers as being those we're most interested in reviewing, and a valid biomarker by definition is one measured in an analytical system with well-established performance characteristics and described within the framework that establishes its significance either in toxicogenomics or clinical pharmacogenomics.

Now, let me turn to what we see at the IND and NDA level at FDA and show you the growth of applications that have come in. This is an incomplete picture. It's actually an informal survey that our review staff has conducted looking at INDs and NDAs and basically doing head-counting as to how many of them have proposed to collect samples for genomic testing. You can see the increase, and I think that increase is real. If we had everyone captured, I think the trend would be the same. But it shows an increasing interest in this science within the drug development process.

Now, the types of data in that survey, in those over 100 INDs and NDAs, include microarrays. When we see these data, we think about it as relatively new technology. There's a lot of heterogeneity in the techniques and test procedures across applications, frequently not really well validated because the intent of the sponsor is to use this information in an exploratory way.

When you go to interpret the data, it's often not confirmatory, but rather hypothesis generating. The studies are oftentimes small, whether they be animal or human, so that extrapolation of the findings are tenuous, and in the overall count from what I showed you on that trend, we have relatively few examples that involve microarrays. So most of those examples that I showed from the IND/NDA world are from genotyping experiments.

We've had more recently many informal meetings with companies where they talk about hypotheticals: "What if I had a drug for this carcinoma, and what if I had this microarray?" These types of meetings are helpful and we encourage them because it allows us to have a dialogue about what might be coming down the pike later on.

The majority of those INDs and NDAs contain genotyping information. I think this is reflective

of the maturity of the techniques and the test procedures, and the fact that there are several well-established biomarkers, particularly in the pharmacogenetic or enzyme activity category.

On the other hand, frequently the interpretation of these data is unclear if it isn't something we already know a lot about, trying to link a genotype with a clinical phenotype such as an adverse event or the absence of efficacy. The hypothesis for that link is often unclear. So these again are hypothesis generating, for the most part.

For the same reason, we have difficulty with what we've seen so far, extrapolating findings from these studies across patient populations. It's not uncommon to see a study with 95 percent Caucasians, and we know that certain alleles distribute differently among the different race groups, and thus to extrapolate this information on a global scale -- for example, into a product label -- is a challenge.

So most of the examples in the survey fall into this category. In fact, what I showed you there, that over 100 count of INDs and NDAs, 75 percent of those involved genotyping of the cytochrome P450 enzymes, this isn't bad. It just shows where companies are starting to focus on more intensively. The pie chart illustrates the distribution of the cytochrome enzymes, and the ones that are most interesting, the ones people feel are most associated with variability and dose-response are the 2C9, 2C19 and 2D6. They all have standard nomenclature, as you may know -- family, for example, of CYP2. There's a subfamily 2D6, and there's a gene indicated as *3. When companies do, for example, 2D6, they generally look at six to eight alleles of 2D6 and then categorize patients into extensive, intermediate or poor metabolizers for the purposes of interpreting those response data.

These are important enzymes because they account for such a large percentage of the metabolism of drugs in the marketplace, so it's not trivial by any means, and we're just beginning to get data on understanding the association between these genotypes of drug metabolism enzymes and clinical outcome.

This little cartoon illustrates the question that we're asking of sponsors in drug development. We're interested in dose-response, dose exposure as I call it here, and many drugs are metabolized through the liver by the polymorphic enzymes, and this shows what happens when one has a single change in a nucleotide in the sequence of the gene that encodes for enzyme activity. Imagine this enzyme is 2D6, so 90 percent of the population would have a sequence that would encode for high levels of enzyme activity. They'd be called an extensive metabolizer. If I give them the same dose as the other folks, they have a certain degree of exposure to the drug in their systemic circulation.

A change in one of the nucleotides in that sequence creates a poor metabolizer. The patient may have no enzyme to metabolize, or perhaps an intermediate level of enzyme to metabolize. We would refer to those as poor metabolizers. They obviously, given the same dose, would have a much different degree of exposure. So exposure leading to response, then, is influenced by the genetic state of the enzyme activity.

Let me just pick one example that's been in the news recently. It's serotonin reuptake inhibitors for childhood depression. The agency has approved two of these this year. There's a lot of clinical trials being reported in the Wall Street Journal and so on. Most of these drugs are 2D6 substrates. The drugs have a relatively narrow therapeutic index which I've indicated with the two lines there, indicating therapeutic response and toxicity.

So imagine I have a fixed dose for all comers who are going to receive this drug. If I have an extensive metabolizer, they're going to have a certain level of exposure. I've indicated with that green, the little mark, and they're going to be happy. They're in the therapeutic window. Their efficacy endpoints, their clinical endpoints will be an improvement in their childhood depression rating scale, and this is composed of nine different symptoms.

Now let's imagine I give the same dose to another child that happens to be a poor metabolizer. Well, their higher exposure pushes them further out in their plasma drug concentrations. They're now on a plateau of the efficacy curve, so they don't get much more benefit from that drug in terms of efficacy. But as you can see from the relationship between safety and efficacy, those children may be at risk for CNS difficulties, insomnia, irritability, and there's some concern now about long-term growth and suicide potential.

So getting the dose right is an important factor in drug development and therapeutics, and knowing the genotype for a drug with a relatively narrow therapeutic index like this can be useful in optimizing the risk-to-benefit ratio, and that's why the agency is interested in it.

I'll share with you three examples to illustrate how this information on drug metabolism has impacted labels. The fact of the matter is we do not have a lot of examples where genomic information has been included in labels. We of course have Herceptin with the FSH test to identify those patients that are candidates for Herceptin. We have resistance testing genotyping in labels for some of the AIDS drugs. We have, for a drug like tamoxifen, in the label some information about receptor-positive or -negative nature.

But let me focus on some recent things dealing with the enzymes, and I'll go through this quickly to illustrate how it works. Voriconazole is an antifungal we approved in 2002. We knew that 2C19 is a major metabolic enzyme that controls clearacin exposure. We also knew from the clinical trials that visual disturbances and potential adverse events were of concern on the safety side. An obvious question would be if I give somebody a dose of this drug and they're a poor metabolizer, does that predispose them to these adverse events?

Well, typical in drug development is to look at genotype early on in drug development in the Phase I studies, and this is some data that shows the difference in exposure depending upon which genotype you happen to be. From a poor metabolizer I have a plasma concentration of 4. If I'm extensive, I'm on the other side and my plasma concentration is 1. So there's this four-fold difference in exposure that could have a bearing on the risk/benefit of this drug in therapeutics. We don't know that in the case of this drug because genotyping, while it was done in Phase I, was not done in Phase III, so there was no way to associate an adverse event with genotype from the clinical studies that were done to demonstrate safety and efficacy.

What often happens in drug labels is that we put information in the section called "Special Populations," where we look at covariates that affect exposure and dose adjustments. For this drug, there were several covariates that warranted a reduction in dose. As you can see, low body weight. If we didn't adjust the dose, they'd have a two-fold higher exposure. Hepatic impairment. People with hepatic impairment had a three-fold higher exposure to the drug, so the sponsor recommended reduction.

But as I showed you that data on 2C19 genotype, people had as much as a four-fold increase in exposure, and we did not in that label recommend a dose reduction, partly because of the fact that if we were to do that, we have to be assured that the test is available to physicians that want to use it, and it's available to patients at a reasonable cost as well.

The other problem is racial differences in phenotypes and genotypes. When a sponsor does this type of testing in drug development, they often will identify in advance what alleles of the enzymes they're going to consider to be poor metabolizers. If they happen to only pick only two of the alleles and not four, that means that whenever the patient doesn't have the two genotypes that they picked, they're going to be classified as the other genotype extensive metabolizer, and that's going to be a mistake.

So it's important in monitoring these studies that we look at the alleles that are included in the study to make sure that the alleles are relevant to all of the major demographic groups that are likely to take the drug if it were to be approved, and this shows an example of that with the 2C19 where the prevalence of poor metabolizers is greater in Asians than Caucasians, and also the actual allele -- for example, the *3 -- actually doesn't appear in Caucasians but it appears in Asians. So you really have to have this information straight when you're reviewing this information.

Now, that was a case -- and I'll turn to a different case to give you a variety of the types of things being done. That case was where a sponsor looked at genotyping in the early phase studies and then did not do anymore in the clinical phases of Phase III. This was a different story. This was atomoxetine for attention deficit disorder. We approved it in January. It's a 2D6 substrate, so it has the same sort of situation as the 2C19, and you can see that a certain percent of Caucasians and African Americans are poor metabolizers. If they get the usual dose of the drug, their exposure based on this area under the curve goes up about 10-fold. Their clearance half-life is about five-fold longer.

This particular drug development program looked at the genotype in the Phase III trials. The numbers are on there. The prevalence of poor metabolizers was right on as to what we would expect, 7 percent in this population. That, in fact, is one way we assess the quality of these tests, is the prevalence consistent with what we already know.

Then when you looked at the clinical outcome data, what they did here was really have everyone double-blinded, and they looked at patients who discontinued therapy and then analyzed the genotype into either poor or extensive metabolizers. The data is interesting because clearly you can see there's a double-edged sword when it comes to genotyping. On the adverse event side, poor metabolizers had a higher rate of adverse events that caused them to discontinue the drug. That's because their exposure was too high. On the other hand, the extensive metabolizers had a different degree of efficacy. I'm not clear why that occurred. It could be related to an active metabolite or something like that.

So you really have to be careful about this type of data, but it does sort of raise the question if I had a patient going on this drug, would testing for their genotype in advance and perhaps lowering the dose proportionately have any benefit in improving the risk/benefit ratio? We don't have data to support that, so there was nothing in the label on that.

What we did put in the label is basically truth in labeling. We put in the label what we know and the evidence that's backed up by credible studies. We put information actually in seven different sections of the label in a hierarchical fashion. The most extensive information went where it was most important, and we had information in the four sections of the label that you see there, including laboratory tests to let the physicians and patients know that a test for this poor metabolizer genotype is available.

Just to mention that this is consistent with labeling regs, the label should describe the evidence and identify specific tests needed for the selection or monitoring of patients who need a drug, and we feel this kind of information is similar to liver enzyme function tests that are used to monitor drugs or blood levels of drug for TDM.

I'm going to finish with a third example. The first two were really talking about new drugs, new drug development programs. But what about drugs in the marketplace, drugs that were approved years ago before genomics was here? Why should they not be considered?

So we began to look at drugs that have narrow therapeutic indices and ask the question can the use of these drugs in therapeutics be improved on the basis of a genomic test? So we looked at 6-mercaptopurine as our example of this idea, and we took it to a pediatric oncology subcommittee in July, as you can see.

Very quickly as a little backgrounder, 6MP is the drug of choice for childhood leukemia. The important part about using this drug is dose titration is critical. Getting the dose right not only affects long-term survival, event-free survival, but it also affects myelosuppression. Too much and you have a problem that requires a patient to go off the drug for three months. Going off the drug for three months affects event-free survival going out years. So getting the dose right is really critical.

It's metabolized by two pharmacologically active nucleotides by the enzyme TPMT. That's the key to this example. Now, TPMT genetic polymorphism is a well-established and well-documented situation, and through the literature and through ongoing research protocols there's a strong link between TPMT polymorphism and clinical effects, and in particular toxicity. On the left you can see the gene frequency of patients that have either normal to high, and then at the other extreme low to absent. 0.33 percent, one out of 300 patients, have no enzyme activity, and if they get the usual dose of this drug, 100 percent of them get toxic with myelosuppression and require hospitalization.

It follows, then, would the availability of a TPMT genotype test be useful in steering the physician and the patient to the correct dose that will have the appropriate risk/benefit? That was the question that we asked the advisory committee of experts.

PG tests for TPMT are fairly widely available commercially and in academic centers. Places like Mayo Clinic and St. Jude's have used them routinely for years to guide therapy, and more recently commercial laboratories as well. These are the three alleles that are measured showing the distribution. There is no difference between ethnic or racial groups, and all the commercial labs that I'm aware of operate under CLIA certification, and some of them operate under, in addition, GLP conditions. Academic labs operate under, in addition, research protocols for using this test.

So it follows, then, that there's a likelihood that the agency will look at product labels of approved drugs and advise those labels appropriately on the evidence that is out there.

So here's a summary. We think the technology in biomarkers from genomics is new, but the fundamental concept of using these markers to enrich populations or to exclude patients from studies or guide dosing, that's not a new concept. We've done it before with phenotypical markers, and all we have now is more precision in terms of a molecular basis for those phenotypes.

For co-development of a test and a drug that is intended for simultaneous marketing, let's say, I'm sure the FDA would recommend submission of complete information on both. The EGs are just examples. IDE shouldn't be the only way to think about this. It depends on the intended use and results of the test. As far as the analytical validity of these tests go, we're comfortable and we rely on the internal QC programs that you've heard about already this morning under the CLIA certification. We're also familiar with laboratories doing these studies that operate under GLP conditions, and this refers again to all the things that you heard about sample handling, the integrity of the incoming RNA and DNA.

Most labs put positive and negative controls. They do things in duplicate. We also look at the outcome in terms of the percentage of alleles being reported for a population to see if that's consistent with what we know, and then some of the laboratories are engaged in the voluntary proficiency testing.

So we tend to review this information like we review pharmacokinetic information or drug blood level information and bring a lot of the bioanalytical standards to this field that we've been applying to therapeutic drug monitoring and so on.

Thanks.

DR. McCABE: Thank you very much, Dr. Lesko. That was very informative. Please join us at the table here for the roundtable.