

2006 Cardiovascular Biomarkers And Surrogate Endpoint Symposium

September 18-20, 2006

Bethesda North Marriott Hotel & Conference Center
Bethesda, Maryland

The Use of Biomarkers for FDA Regulatory Decision Making – Cardio-Renal

Norman Stockbridge, M.D.

Division of Cardio-Renal Drug Products
Center for Drug Evaluation and Research
US Food and Drug Administration
Silver Spring, MD

It is important to draw some distinction between the biomarkers that are true measures of organ function and the molecular biomarkers like lipids. I think that biomarkers that are reflections of organ function are at least somewhat more likely to make it into the surrogate endpoint status and we have several examples of organ function kinds of endpoints that we deal with in the cardio-renal division.

We will consider first the situation with respect to blood pressure as a surrogate endpoint. Mountains of evidence from epidemiological studies and from interventional studies involving a wide variety of drug products and drug classes, established the principle that clinical benefit was derived from lowering blood pressure.

It is worth thinking about what a development program for an antihypertensive is. I assert that a very small number of subjects could be put into a carefully crafted clinical trial and you could probably work out pretty much all of the pharmacodynamic properties of the new antihypertensive in a very small number of subjects, maybe a dozen or so. The fact is that development programs for new antihypertensives tend include approximately 1500 patients for monotherapy submissions and approximately 2500 patients for programs that evaluate a combination product in tandem with the monotherapy. Occasionally, things show up in the development program that resulted in a much larger development program than even one of those.

Why does this happen? Why is there such a discrepancy between the amount you need to establish the surrogate endpoint and what is necessary to support approval for this very well validated surrogate endpoint? The answer is clearly that not everything a drug does can be predicted from its mechanism of action or even its pharmacodynamic properties. Therefore, the bulk of the information that is collected in that development program is in fact intended to support the safety of the drug to deal with the anticipated discrepancies and the ability to predict what the true benefits are going to be based on looking at the pharmacodynamics. Of course, it remains the case that the studies within the development program tend, almost exclusively, to target effectiveness and so the safety data become the subject of interpretations that are not in the frequentist realm. So I think this is an example of one of the best possible cases for a surrogate endpoint. It is fully validated by an enormous body of evidence and relates to organ function and yet, the

development programs will always remain very large. A corollary of that is that the size of the effect on blood pressure matters quite a bit, but for practical reasons, people don't tend to bring small effect sizes to us because you will have trouble reaching goal and a product with a very small treatment effect is apt to be more of a nuisance than anything else. Second of all, you are going to need a correspondingly larger safety database in order to support a small treatment effect.

Let's next consider myocardial infarction, which as it is treated, in terms of it being a clinical event, in fact, has components of it, which are defined by biomarkers. The endpoint is often engineered to establish myocardial necrosis that is large enough to lead to later clinical events or to predict real clinical implications after the resolution of the acute event. This magnitude of event is assured by paying attention to elements that are true clinical symptoms, but also things that relate to ECG changes of some magnitude and enzymatic changes of some magnitude. The process of defining an MI, that is an MI in a clinical event sense, is entirely empirical and may be subject to redefinition depending on what other events are going on at the same time.

You might still ask why a little bit of myocardial necrosis isn't bad enough to warrant attention and of course, there are several parts to this answer. One is again to recognize that drugs have unintended consequences, even when you think you understand mechanisms very well.

We have been burned, if you will, in the cardiovascular area by antiarrhythmic drugs, which are well understood in terms of what their mechanism of action is and are effective in sense of suppressing arrhythmias in a predictable fashion and yet, have been shown not to have the net clinical benefit in terms of mortality that you might expect.

Another example of the same kind of phenomenon has been the experience with the use of inotropes in heart failure, where at least some of the phosphodiesterase inhibitors, in particular, certainly do have beneficial effects on hemodynamics, but do not have beneficial effects when looking at long-term consequences.

Again, these are all circumstances where there are symptomatic and organ function related things that manifested exactly what they were supposed to do and these are not based on markers per se, which presumably takes you a little bit further from outcomes. There remains a need to understand the relationship between changes in the biomarker and the clinical events, if you are going to make some sense or how much of a change in a marker is enough to matter. You always have to worry, at least with drugs, about unexpected or unintended consequences of drug action, nonspecific actions, and that these are going to be a particular problem for a new product if there has already been a product that has been approved based on real clinical endpoint.

I don't have any problems saying that I would like to keep my troponin inside my myocytes and I would like for my coronary arteries to stay open, all other things being equal, but it is difficult to pin down what it takes to show that all other things in fact remain equal. I think, for those reasons, it is always going to be true that claims based on

markers such as ejection fraction or other measures of hemodynamics will be difficult. Some anatomical things, I think fall into the same category and of course the same thing is going to be true for soluble biomarkers.

In addition to having biomarkers play a role however in components of a definition of a clinical endpoint, biomarkers do show up as components of composite endpoints, that is, things that are considered to be events by themselves. One example of that is the use of a rise in serum creatinine as a measure of a clinical event relating to renal disease progression. We have accepted that as a clinical endpoint. It requires again, setting some magnitude of effect that is large enough to be considered very likely to be followed soon by a clinical event and the trials that have utilized endpoints, composite endpoints like this have had considerable amount of internal data generated that supported the idea that the softest part of the endpoint was indeed followed by harder clinical events.

As sort of an intermediary step in establishing surrogates, there is Subpart H, and I will mention to you that for biological compounds, there is a Subpart E, which is much analogous to Subpart H. Another speaker mentioned the notion that a putative surrogate endpoint needs to be reasonably likely to provide clinical benefit. "Reasonably likely" is the term that does not have any clear definition. It is a matter of judgment and reference was made to the idea that if you got approved this way, you undertook a commitment to provide supportive data and clinical outcome data in the post-marketing setting.

Is this in fact a reasonable middle ground for promising candidates for surrogacy? There are, I think now, about as many review divisions as the total number of Subpart H CDER approvals over this four-year period, so it is on the order of about one approval per review division per four years, so this is not commonly done.

Cardio-renal last had one of these in 2002, a case that requires a footnote because it was not in fact based on a surrogate in the sense that we have been talking about it; it was in fact a creative use of subpart H in a setting where a drug had failed in its primary analyses, which were based on a clinical endpoint. The sponsor came in with a completely retrospective analysis that combined several clinical benefit endpoints and we elected this mechanism (Subpart H) to support approval. It was subsequently followed up with a trial that had a more straightforward interpretation.

The previous time the cardio-renal division was involved in a Subpart H approval was in 1996, ten years ago and we are still awaiting the confirmatory clinical trials there demonstrating clinical benefit.

So insuring that a phase IV study gets completed in the setting of one of these approvals is a big problem. Probably the best chance someone would have in approaching us today about getting a Subpart H approval would be a setting in which the phase IV study was underway or preferably completely enrolled by the time the approval was granted under Subpart H, and we would only have to trust that the sponsor completed the followup of the ongoing study. Theoretically, this is a mechanism for getting accelerated approvals, but it is not in fact a pathway that is going to be successful very often.

For markers for risk, there is clearly a lesser degree of validation required. As many people have eluded to, the idea is that things need to have biological plausibility and need to have some relationship to pathogenesis. But you don't absolutely have to be able to show that the clinical effects follow from the marker; you just have to show that they do some reasonable relationship. Then of course how we interpret changes in a marker depends on which direction a marker moves. If a marker moves to a small extent in a favorable direction, we are apt to say, "well, why does that matter?", whereas if it moves in an adverse reaction, we are apt to consider it a cause for concern.

I guess I can't talk about markers in cardiovascular area without talking about proarrhythmic risk and QT. Of course the big problem in understanding proarrhythmia is despite the fact that we have a good understanding at the receptor and single channel level of how the electrical properties of the heart originate, and at the macroscopic level how the electrical activities in the heart manifest in a more global sense, in fact, it is exceedingly difficult to put those pieces together in a way that is adequately predictive of the effects of drugs. The result of that is that we pay attention to a fairly imperfect marker, the surface manifestations of electrical activity in the form of the QT interval, and take a very conservative stand with respect to what magnitude of an effect there is so small as to be insensibly involved in risk. This is however, of course a fertile area for additional preclinical or clinical tools to assess risk and FDA has put effort to facilitate in that regard.

In summary, a biomarker for benefit is generally pretty hard to establish and even if you are successful, you are going to get stuck with some fairly large sample sizes. Don't think that getting a surrogate approved results in small trials. It does not. Much more fruitful area to think about is the role of biomarkers in the assessment of risks, which is something in many cases is sorted out by the sponsor before drugs get to FDA, and in identifying promising targets and promising compounds. It has to do with dose ranging and it has to do with figuring out where you think you can best establish the effectiveness in your phase III program. One concern, I guess, you should at least think about as you pick out biomarkers to enrich a trial is that you run some risk that we will decide that the benefits may only apply in the setting of the risk factor. You may end up with narrower labeling than you might wish.

Just a couple of final thoughts: I do not think that it is an appropriate idea to think about validating a surrogate endpoint in the context of a single development program. A single development program is one piece of a much larger database that is necessary in order to validate a biomarker as a surrogate. But you should continue to work just as hard as you ever did in developing biomarkers. It is the only game in town. You have got to figure out whom to treat and what the basis for pathology is and that is just where everything has to lead.