Biomarkers and Surrogate Markers: An FDA Perspective

Russell Katz

Division of Neuropharmacological Drug Products, United States Food and Drug Administration, Rockville, Maryland 20852

Summary: Interest is increasing rapidly in the use of surrogate markers as primary measures of the effectiveness of investigational drugs in definitive drug trials. Many such surrogate markers have been proposed as potential candidates for use in definitive effectiveness trials of agents to treat neurologic or psychiatric disease, but as of this date, there are no such markers that have been adequately "validated," that is, shown to predict the effect of the treatment on the clinical outcome of interest. While the current law and regulations permit the United States Food and Drug Administration to base the approval of a drug product on a deter-
mination the effect of the drug on an unvalidated surrogate marker (that is, one for which it is not known that an effect on the surrogate actually predicts the desired clinical benefit), there are a number of difficulties in interpreting trials that use surrogate markers as primary measures of drug effect. In this article, the relevant regulatory context will be discussed, as well as the epistemological problems related to the interpretation of clinical trials in which unvalidated surrogate markers are used as primary outcomes. **Key Words:** Surrogate marker, accelerated approval, validation, biomarker, FDA.

INTRODUCTION

In considering the potential utility of, as well as the potential problems raised by, the use of surrogate markers as primary outcome measures in definitive effectiveness trials of new therapeutic agents, it is valuable to describe the regulatory context in which such use falls. Once this context is described, I will describe the uses for which these measures have been proposed, and then describe the various advantages and disadvantages of this approach.

It is worth defining the terms biomarker and surrogate marker. Typically, "biomarker" is defined as a laboratory measurement that reflects the activity of a disease process. There are many such markers identified for many diseases of the nervous system, for example, various magnetic resonance imaging (MRI) measures in multiple sclerosis and Alzheimer's disease treatments, positron emission tomographic (PET) scanning of dopamine transporters in Parkinson's disease, etc.[1–4] In essentially all cases, these markers quantitatively correlate (either directly or inversely) with disease progression. A "surrogate marker" can be defined as "...a laboratory mea-
surement or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful endpoint that is a direct measure of how a patient feels, functions, or survives and is expected to predict the effect of the therapy."[5] The primary difference between a biomarker and a surrogate marker is that a biomarker is a "candidate" surrogate marker, whereas a surrogate marker is a test used, and taken, as a measure of the effects of a specific treatment.

This discussion will focus on the use of surrogate markers as primary outcome measures in controlled trials designed to provide evidence of effectiveness of new treatments. The use of either biomarkers or surrogate markers in early phases of drug development is, from a regulatory perspective, noncontroversial. That is, the use of these markers to obtain information early in drug development that may be critical to further development seems self-evidently appropriate. For example, the use of various imaging modalities to determine if an investigational drug localizes to a particular brain region (or receptor subtype) thought to be important to the therapeutic effect(s) of the drug seems appropriate, as does the use of other markers to determine presumed effective doses. As a general matter, these uses of biomarkers or surrogate markers are encouraged; however, as will be discussed below in the context of the use of these markers in definitive effectiveness trials, it is possible that the changes seen in these markers early in drug development

---

may be misleading. Nonetheless, their use in these early stages seems reasonable.

REGULATORY CONTEXT

Before a drug may be approved for marketing in the United States, it must be shown to be effective. The legal standard of effectiveness that such a drug product must meet is defined in the Food, Drug, and Cosmetic Act (the Act) as "substantial evidence of effectiveness," which is further defined as "...evidence consisting of adequate and well-controlled investigations, including clinical investigations, ...that the drug has the effect it purports or is represented to have under the condition of use prescribed ...in the labeling or proposed labeling thereof." Although a strict reading of this definition might permit the approval of a drug product that has an effect on some measurement that reflects absolutely no clinical benefit to the patient (as long as it can be adequately described in the label), in practice, drugs are permitted to be marketed only if they can be shown to confer a clinical benefit to patients (although, of course, they need not produce a beneficial effect in all patients). Typically, drugs are approved on the basis of adequate and well-controlled clinical trials that permit the conclusion that the drug has a beneficial effect on an outcome that is directly and obviously related to the patient's clinical status. For example, a drug proposed as a treatment for epilepsy must be shown to decrease the frequency of seizures, a drug proposed as a treatment for Major Depressive Disorder (MDD) must demonstrate, in adequate clinical trials, a decrease in the depressive symptoms of patients with MDD, usually assessed by ratings on validated scales of depressive symptoms, etc. In other words, ordinarily, drugs are approved on the basis of their manifest effects on the clinical signs and/or symptoms of the disease for which they are intended as treatments.

However, in some cases, the proposed clinical benefit (for example, an effect on survival) might not be detectable in trials of reasonable duration or size; that is, in these cases, the detection of such an effect might require enormously large studies of considerable (i.e., many years) duration. In those cases in which the predicted effect is considered clinically very important (such as mortality), and there are few if any other available treatments available for the given condition, delaying approval of products until such studies are completed might be considered inappropriate.

These considerations became critical during the search for effective treatments of Acquired Immune Deficiency Syndrome (AIDS) and HIV-related disease in the early 1990s. As a result of the interest of numerous parties in rapidly bringing treatments with important benefits to patients with AIDS, the Agency adopted new regulations designed to hasten approval of important new therapies. Found at Subpart H of 21 CFR 314.500 (Code of Federal Regulations), and known as the Accelerated Approval provisions, the Agency for the first time articulated an explicit standard for drug approval based on the effect of a drug on a surrogate marker, and not a clinical outcome. The relevant portion of the regulation is as follows:

The United States Food and Drug Administration (FDA) may grant marketing approval for a new drug product on the basis of adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit, or on the basis of an effect on a clinical endpoint other than survival or irreversible morbidity.

It is critical to recognize that this regulation applies to the use, in definitive effectiveness trials, of what can be called "unvalidated" surrogate markers. Validated surrogate markers are those for which evidence has established that a drug-induced effect on the surrogate predicts (results in) the desired effect on the clinical outcome of interest. An unvalidated surrogate, on the other hand, is a surrogate that is, as the regulation describes, "reasonably likely" to predict the clinical benefit of interest, but for which there is insufficient evidence to establish that such an effect, does, in fact, result in the desired clinical outcome.

The preamble to these regulations[6] (a discussion of the regulations printed in the *Federal Register*, which, while not a part of the regulations, helps in their interpretation), makes clear a number of important points.

First, the Agency rejected the objections of those who believed that the adoption of this new standard ran counter to the specific requirements of the Act. Since the new regulations required that the effect on the surrogate marker need only be "reasonably likely" to predict the clinical benefit, some felt that this was inconsistent with the requirement of the Act that the drug be shown to have the effect claimed for it in labeling. The Agency's view is that a drug may be approved under Subpart H only if there is substantial evidence that it has the effect claimed for it on the surrogate marker, and that, therefore, the new regulations are consistent with the provisions of the Act.

Further, the preamble explicitly states that when a clinical benefit can clearly be shown in reasonable clinical trials, the Agency will not rely on an effect on a surrogate to support approval. The regulations impose numerous requirements on sponsors of drug products approved under these provisions, e.g., the regulations apply only to proposed treatments for serious and life-threatening conditions that "...provide meaningful therapeutic benefit... over existing treatments...," they require that studies be done after marketing to demonstrate the link between the effect on the surrogate and the predicted

clinical benefit, and they require that the Agency evaluate the advertising materials of a sponsor before their use. Underlying these requirements is the recognition that approval based on the effect of a drug on an unvalidated surrogate marker introduces uncertainty about the drug's true clinical benefit that is ordinarily not present at the time of approval under more typical circumstances. The regulations are meant to apply only in those clinical situations in which this level of uncertainty is acceptable but, even then, only temporarily; that is, while attempts are ongoing to "validate" the surrogate. If the attempt to validate the surrogate fails (or no reasonable attempt is made), the drug may be removed from the market.

Although these (and all) regulations have the force of law, they were adopted via the process of rule-making (as are all regulations), which does not require the action of Congress. As a result, Congress, in 1997, as part of the FDA Modernization Act (FDAMA), incorporated the following provision into Section 506(b) of the Act, which permits the Agency to approve a drug product:

"...upon a determination that the product has an effect on a clinical endpoint or on a surrogate endpoint that is reasonably likely to predict clinical benefit."

In fact, the Agency has approved drug products on the basis of their effects on surrogate markers for many years before the adoption of these regulations (examples include anti-hypertensives, cholesterol-lowering agents, and treatments for glaucoma). In each of these cases, drugs have been shown to have a beneficial effect on the surrogate of interest (blood pressure, serum cholesterol, and intra-ocular pressure, respectively) without a requirement that they be shown to have an effect on the relevant, desired clinical outcomes (heart attacks, strokes, death in the first two cases, and visual loss in the last). What distinguishes these cases from the cases to which the new regulations and law are intended to apply is that in the former cases, there is (or presumed to be) evidence that the effect on the surrogate does, in fact, predict the desired clinical outcome; that is, these are considered "validated" surrogate markers.

It is worth noting that the FDA does not have a monolithic position on the approval of drugs under the Accelerated Approval regulations. Clearly, there are units in the Agency in which such approvals are routinely granted; for example, some anti-cancer treatments are approved under the Accelerated Approval regulations.[7] In these cases, drugs tested in patients refractory to available treatments are approved on the basis of their effects on tumor size, as assessed by imaging. In these cases, however, the drugs are considered to be reasonably likely to predict the desired clinical benefit, and the labeling adopted at the time of the approval indicates the treatment for patients with the clinical condition studied, with accompanying language describing that the approval was based on the effects of the treatment on tumor size, and

that there is no evidence of a clinical effect. The labeling is changed to include a description of the documented effect on survival only when the subsequent studies are done that demonstrate the effect on the clinical outcome. To date, however, no treatments for neurologic or psychiatric diseases have been approved on the basis of an effect on a surrogate marker, validated or unvalidated.

Most, if not all, recently approved treatments for neurologic and psychiatric disease are considered to have symptomatic effects; that is, there is no (clinical) evidence that these treatments have any effect on the underlying pathologic substrate of the disease under study. Although, of course, some of them may have such an effect on the underlying pathology, the clinical studies performed to support the approval of these treatments are not capable of detecting such an effect. In the typical study conducted to support approval, patients are randomized to treatment or control (usually placebo), observed throughout the time course (depending upon the condition under study, the duration of the study may be 3 weeks to 2 years), and assessed on a relevant clinical outcome, as described above. Effectiveness is determined by the difference in outcome between patients who received the drug and those who received the control treatment. This design is incapable of detecting whether or not the effect is symptomatic or structural and enduring. Typically, indications for such treatments in approved product labeling are silent about the nature of the effect (for example, a drug approved on the basis of its effect in reducing, relative to control, the frequency of partial seizures, will be indicated as a treatment for partial seizures), but no implication that the effect is structural (either in labeling or advertising) is permitted. Clearly, for the treatments approved to date, practical study designs exist (in terms of duration, size, and available clinical measures), and are capable of detecting treatment effects. It is explicitly in these circumstances that the use of surrogate markers as primary outcome measures is unwarranted, as discussed in the preamble to the regulations, and described above.

It is important to recognize this, because the advantages often ascribed to reliance on the effect of a drug on a surrogate marker are that trials can be considerably shorter in duration and smaller in size (number of patients enrolled) if the primary outcome is a surrogate rather than the clinical outcome of interest (e.g., mortality). Indeed, authors have estimated the "savings" in time and duration that would accrue in various clinical settings if a surrogate, instead of a clinical outcome, were chosen as the primary measure of drug effect.[8] As noted above, though, for most treatments for neurologic disease, studies of reasonable duration and size can detect an effect on a clinical outcome.

There are now a number of pharmaceutical sponsors that have begun to develop treatments that they believe

might have an effect on the progression of the disease (i.e., they are presumed to have an underlying structural, enduring effect). For example, treatments designed to inhibit the formation of amyloid plaques in patients with Alzheimer's disease or mild cognitive impairment, treatments designed to retard the loss of dopaminergic neurons in patients with Parkinson's disease, and treatments designed to minimize the lesion burden in patients with multiple sclerosis are all actively being investigated. It is possible to imagine, for example, that for a particular treatment that produced an effect on the underlying progression of an illness, studies of reasonable duration or size might not be practical to conduct. That is, it might not be possible to detect a difference between the treatment and control on an obvious clinical outcome in a trial of reasonable duration or size. It is in the context, then, of attempting to demonstrate the effect of a drug on the underlying progression of the disease when a clinical effect would be unexpected to be easily detectable that there has begun to be serious discussion about the use of surrogate markers as primary measures of drug effect.

Clinical trial designs have been proposed, however, that are capable of detecting the effect of a drug on the underlying progression of disease, and these trials are unambiguously interpretable.[9] In these trials, so-called randomized withdrawal designs, patients are randomized to drug or placebo and observed for an appropriate duration, as in typical clinical trials. However, in this design, patients initially treated with active drug are withdrawn from treatment (patients initially treated with placebo are continued on placebo). In this withdrawal period, if patients initially treated with active drug approach the status of patients initially treated (and continuing to be treated) with placebo, this is taken as evidence that the active treatment had only a symptomatic effect (that is, when the treatment is removed, patients approach where they would have been had they never received treatment). If, on the other hand, patients initially treated with active drug do not approach the (continuing) placebo patients (that is, the slope of their decline is at least parallel to that of the placebo patients) this implies that the treatment had a structural effect (that is, when the treatment is removed, there is a persistent difference compared to how they would have fared had they not been treated).

Although these designs appear to be able to provide clear evidence for either a symptomatic or structural effect (that is, they provide unambiguous evidence for either conclusion), they are difficult to perform adequately, for at least two reasons. It is typically not clear how long the withdrawal observation period should be. It would be inappropriate to conclude that the post-treatment slopes of the two groups were parallel, for example, if the patients were followed for too short a period to adequately assess the slope. In addition, the goal of such a study (assuming the goal is to demonstrate a structural effect) is to demonstrate parallel slopes. This leads to a situation in which equivalence of slopes is the goal; demonstration of equivalence of outcomes ordinarily requires large numbers of patients. Perhaps for these reasons, these trials are rarely performed.

Given that the clinical trials designed to detect a progression effect are difficult, and given the appearance of many new investigational treatments designed to affect the underlying pathology of various diseases, there has been renewed interest in relying on the effect of a drug on surrogate markers to support approval. However, there are a number of important issues that need to be addressed before such use can be contemplated.

As discussed above, the Agency recognizes that basing an approval on the effect of a drug on an unvalidated surrogate introduces additional uncertainty into the approval process. Clearly, this uncertainty is considered acceptable under certain circumstances. As noted earlier, however, current and past treatments for neurologic and psychiatric drugs have all been approved on the basis of relatively simple clinical trials. In a real sense, therefore, approvals on the basis of an effect on a surrogate marker have been "unnecessary." By this I mean that, while it may be that (some of) these previously approved drugs do, in fact, have an effect on the underlying pathology of the disease, a manifest clinical benefit has been demonstrable in clinical trials that are relatively easy to perform. In such cases, relying on an effect on a surrogate might, perhaps, document an additional effect on the underlying progression; however, such an effect would only support a particular claim in labeling (for an effect on progression) but would not be necessary for actual approval (because the approval could be based on a direct clinical outcome). It is, of course, important to know if a drug has an effect on the underlying progression of an illness, but it may not be necessary for initial approval. Therefore, the use of surrogates as primary outcomes can be considered only in those cases in which an effect on a clinical outcome (whether or not it represents a symptomatic or structural effect) cannot be detected in a practically performable clinical trial.

However, the reliance on the effect of a drug on a surrogate measure is fraught with interpretative problems. Most of these problems have been clearly described, perhaps most clearly by Fleming and DeMets.[10] As these authors point out, there is a natural inclination to accept as a useful surrogate those measures that correlate with disease progression. That is, many consider that a marker that is known to progress with the disease in the untreated state would be, *a priori*, a reasonable choice as a surrogate marker. For example, many imaging markers of Alzheimer's disease have been proposed as useful surrogate markers (e.g., total brain volume, hippocampal volume, etc.). All of these markers corre-

late very well with the progression of Alzheimer's disease. Similarly, various dopamine transporters, as visualized via PET scanning and taken to be a measure of functioning dopaminergic neurons, also correlate well with progression of Parkinson's disease, and have been proposed as useful surrogate markers. Many other such markers in various other diseases of the CNS have also been proposed as useful surrogate markers.

However, as Fleming and DeMets[10] have observed, the correlation of a measure with clinical progression does not automatically confer "surrogate" status on that measure. It may be a necessary condition, but it is not sufficient. It is critical to recognize that this relationship (the correlation between the surrogate and the clinical outcome of interest) may not persist under drug treatment. For example, it is possible that the drug may have the desired effect on the clinical outcome, with no effect (or an effect in the wrong direction) on the surrogate, Alternatively, the drug may have the effect desired on the surrogate marker, but no effect (or a deleterious one) on the clinical outcome (this is the outcome most troubling to the Agency, because approval of a drug based on such an effect would result in the availability of a treatment that is at best worthless and at worst only harmful).

Before we could be sure that the drug-induced effect on the surrogate predicts the desired clinical benefit; therefore, we have to understand the (biological) relationship between the surrogate and the clinical outcome. In other words, we need to know if and how the surrogate and the clinical presentation are intrinsically biologically related to one another; they need not be. As Fleming and DeMets[10] describe, there may be many pathophysiologic pathways that contribute to disease manifestations, and the surrogate may be "in" some, any, or none of these pathways. Furthermore, beyond knowing the biologic relationship between the surrogate and the clinical disease, we need to know (all) of the actions of the drug being investigated to conclude that the effect on the surrogate will translate into the beneficial clinical outcome desired. This is critical because, although the drug may have the desired effect on the surrogate, it might also have other properties that counteract the beneficial effects, and so therefore may result in no clinical effect, or even a harmful one. Indeed, there are many examples of the disjunction of the responses on a surrogate marker and the desired clinical outcome.[11,12]

A simple hypothetical example may serve to illustrate the points discussed above.

Consider the effects of a treatment for Alzheimer's disease on total brain volume as assessed by MRI. Before one could interpret any difference between treatment and control on this measurement as being beneficial to patients, several questions would need to be satisfactorily addressed.

First, one would need to be assured that there is no interaction between the test system (MRI) and the treatment, giving rise to a completely spurious relationship between treatment and apparent benefit. That is, such an interaction could exist (similar to the interactions known to occur when the presence of one drug interferes with the assay for a second drug in plasma). Assuming that such an interaction can be excluded, the question of how to interpret any actual structural change as seen on MRI must be addressed.

It is possible that, in fact, the drug does induce a structural change in the brain, as seen on MRI, but that the change is clearly irrelevant to any presumed beneficial effect of the treatment. For example, an increase in extracellular fluid might be interpreted as a decrease in brain atrophy (or an increase in brain volume), but this would clearly not be a structural change of interest.

More difficult to assess, however, would be the occurrence of the desired effect on brain structure, as imaged by MRI. For example, the drug might, in reality, "preserve" neurons, which might be seen as less brain atrophy on MRI. However, it would not be easy to determine if these preserved neurons were (normally) functional. Alternatively, a treatment might actually decrease amyloid plaque formation, and this might be visualized on imaging. The interpretation of this as a clinical benefit presupposes that amyloid plaque formation is, in fact, fundamental to the genesis of dementia, and not, perhaps, itself a result of the dementing process.

These two considerations [knowledge of the biological relationship of the surrogate and the clinical outcome and knowledge of (all) of the mechanisms of drug action-beneficial and deleterious] pose serious regulatory issues.

As discussed earlier, drugs are typically approved on the basis of a beneficial effect seen in randomized controlled trials. This is a distinctly empirical approach. That is, it relies little on an understanding of the explanation of the results. This may be surprising to some. After all, ordinarily, a great deal is known about the pharmacological actions of any drug before it is approved, and often there is considerable knowledge about the biological mechanisms producing the signs and symptoms of the disease under study. What is important to recognize is that we rarely have knowledge (in the sense of knowing the truth) about which, if any, of the identified pharmacologic actions of a drug contribute to its clinical effects (and, of course, we have no knowledge about any unidentified actions), and we typically have even less knowledge about the mechanisms of disease, and little to no knowledge (although there are usually theories) about the interaction of the two. The history of drug development is replete with examples of drugs that have been shown to have, only years after approval, presumably relevant pharmacologic mechanisms of action, and only when further (incomplete) knowledge has been obtained about the pathways involved in disease production (it is

typically only these latter findings that motivate a search for the relevant mechanisms in the drug). Similarly, mechanisms of drug action believed to be relevant to the clinical effect of a drug are not uncommonly found to be irrelevant, based on new information about the relevant pathways producing disease.

Because our knowledge of the relevant pharmacologic and biologic events is always imperfect and incomplete, drugs are typically approved on the simple finding of a beneficial effect in adequate clinical trials (and on adequate safety data), and this is ideal, not only because it is a direct clinical benefit that is obviously desired by the patient, but also because waiting to fully understand the relevant biologic events before drugs are approved is an undesirable strategy.

This is why approval of a drug on the basis of an effect on an unvalidated surrogate marker represents such a fundamental departure from the typical course of action in drug approval. That is, approval of a drug on the basis of such an effect presupposes knowledge of events that is normally not only absent, but, in a sense, irrelevant. Because we do not ever have all of this knowledge, approval based on effects on surrogate markers will invariably involve a level of uncertainty not typical of the more standard route to drug approval.

There are other problems associated with effects on surrogates that relate to the duration of studies and size of treatment effect, which are also worth making explicit. Although typical drug studies vary considerably in duration (again, depending upon the disease in question and ethical and other considerations), for drugs to be administered chronically, study durations can typically be 3- to 6-months long, or longer. The purpose of longer duration studies is to provide some assurance that any effects seen will persist for at least some reasonable amount of time (again, the duration of these trials always represents a compromise between trials that can document years-long benefit and ethical and public policy considerations).

The situation, however, may be very different with surrogates. It is possible that an effect on a surrogate may be detectable very early in the course of treatment. Because one of the advantages claimed for surrogates is that their use may shorten clinical trials, there could be motivation to perform very short trials when relying on surrogates as primary outcomes. In this case, we might have little assurance that any effect seen early (and on which it would be concluded that the drug was effective) would persist, even to the relatively short duration typically assessed in clinical trials (for example, 6 months for a clinical trial of a treatment for Alzheimer's disease).

In addition, there is usually little attention paid to the size of a treatment effect in trials of neurologic agents. That is, a difference between drug and control in the mean measurement of an appropriately chosen clinical outcome measure is ordinarily considered sufficient to support a conclusion that the treatment is effective (for example, a mean decrease in the monthly seizure frequency, compared to placebo, is considered a face-valid, clinically useful outcome). Indeed, such a mean difference (however small) is ordinarily reflective of a complete shift in the distribution of responses on drug compared to placebo (that is, more patients on drug than placebo have a given degree of improvement over the entire possible range of improvement).

The situation can be very different with surrogates. Again, one of the advantages claimed for surrogates is that they can be very sensitive to treatment effects, and that, therefore, they can markedly diminish the size of a clinical trial. However, given that we do not know the relationship between drug-induced changes on the surrogate and the clinical outcome (recall that we are discussing unvalidated surrogates), we would have no assurance that any (small) effect seen on the surrogate would translate into a clinically meaningful benefit, which must be the basis for any drug approval.

As we have been discussing above, an effect on a surrogate marker can be proposed as predicting the desired clinical benefit of a specific treatment. Also as noted above, in these cases, the Subpart H regulations require that sponsors "validate" the marker after approval has been granted. Typically, this "validation" takes the form of demonstrating that there is, in fact, a strong correlation between the effect on the surrogate and the clinical outcome of interest (e.g., decreased mortality). Such a showing is ordinarily taken to provide evidence that the effect on the surrogate predicted the effect on the clinical outcome.

However, surrogate markers are most useful when they can be more generically "validated." That is, instead of validating the surrogate for a specific drug, the ideal approach would be for a sponsor (more realistically, this would require the cooperation of the entire therapeutic community) to demonstrate that the effect on the surrogate and the effect on the clinical outcome are quantitatively similar across many drugs (and especially across many drug classes designed to treat a given clinical indication).[13] In this way, one could be confident (although still not absolutely certain) that an effect seen on the surrogate for the next (as yet untested) proposed treatment will have the expected and desired clinical effect. In other words, such an approach is the only reliable method of demonstrating (although, again, not with certainty) that the effect seen on the surrogate, regardless of the treatment applied, will translate into the clinical benefit. This is the case currently for anti-hypertensives and cholesterol-lowering agents.

While the FDA is working with sponsors and experts to address the issues raised by the use of surrogate markers discussed above, there appears to be general consensus at this time across numerous therapeutic communi-

ties that surrogate markers are not yet sufficiently understood (and none are validated in the sense described earlier) to be used as sole primary measures of effectiveness in definitive trials of investigational drugs. Indeed, at a recent meeting of the Peripheral and Central Nervous Systems Drugs Advisory Committee (a standing panel of academic experts that advises the FDA on numerous drug development issues) held in November, 2002, panel members and other invited experts agreed that there were no validated MRI markers with which to assess drug effect in patients with Alzheimer's disease, and that no such marker should be used as a primary measure of drug effect. Attempts are underway in the Alzheimer's community, as well as in other disease-specific areas, to validate various surrogates. Although current regulations permit drug approval on the basis of an effect on an unvalidated surrogate marker, a sponsor who wishes to obtain approval on the basis of the effect of a drug on such an unvalidated marker will need to adequately address the numerous thorny issues attendant to demonstrating that any such effect will be "reasonably likely" to predict the desired clinical effect. The evidence that will support such a showing, which may include animal data (both clinical and histopathological), human data (pathology as well as clinical) as well as require further investigation into the pathophysiology of the condition under study and into the pharmacology of the drug(s) under study, remains to be determined.

REFERENCES

1. Rovaris M, Comi G, Ladkani D, Wolinsky JS, Filippi M, and the European/Canadian Glatiramer Acetate Study Group. Short-term correlations between clinical and MR imaging findings in relapsing-remitting multiple sclerosis. *AJNR Am J Neuroradiol* 24:75–81, 2003.
2. Stoof JC, Winogradzka A, van Muiswinkel FL, Wolters EC, Voorn P, Groenewegen HJ et al. Leads for the development of neuroprotective treatment in Parkinson's disease and brain imaging methods for estimating treatment efficacy. *Eur J Pharmacol* 375:75–86, 1999.
3. The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association and the National Institute on Aging Working Group. Consensus report of the Working Group on: molecular and biochemical markers of Alzheimer's disease. *Neurobiol Aging* 19:109–116, 1998.
4. Morrish PK, Sawle GV, Brooks DJ. An ^{18}F dopa-PET and clinical study of the rate of progression in Parkinson's Disease. *Brain* 119:585–591, 1996.
5. Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* 282:790–795, 1999.
6. New drug, antibiotic, and biological drug product regulations; accelerated approval—FDA. Final rule. *Fed Regist* 57:58942–58960, 1992.
7. Johnson JR, Williams G, Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J Clin Oncol* 21:1404–1411, 2003.
8. Fox NC, Cousens S, Scahill R, Harvey RJ, Rossor MN. Using serial registered brain magnetic imaging to measure disease progression in Alzheimer's disease: power calculations and estimates of sample size to detect treatment effects. *Arch Neurol* 57:339–344, 2000.
9. Leber P. Slowing the progression of Alzheimer disease: methodologic issues. *Alzheimer Dis Assoc Disord* 11(Suppl 5):S10–S21, 1997.
10. Fleming TR, DeMets DL. Surrogate endpoints in clinical trials: are we being misled? *Ann Intern Med* 125:605–613, 1996.
11. Riggs BL, Hodgson SF, O'Fallon WM, Chao EYS, Wahner HW, Muhs JM et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 322:802–809, 1990.
12. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 324:781–788, 1991.
13. Hughes MD. Evaluating surrogate endpoints. *Control Clin Trials* 23:703–707, 2002.