Regulatory Aspects in Using Surrogate Markers in Clinical Trials

Aloka Chakravarty

3.1 Introduction and Motivation

Surrogate marker plays an important role in the regulatory decision processes in drug approval. The possibility of reduced sample size or trial duration when a distal clinical endpoint is replaced by a more proximal one hold real benefit in terms of reaching the intended patient population faster, cheaper, and safer as well as a better characterization of the efficacy profile. In situations where endpoint measurements have competing risks or are invasive in nature, certain latitude in measurement error can be accepted by deliberately choosing an alternate endpoint in compensation for a better quality of life or for ease of measurement.

3.1.1 Definitions and Their Regulatory Ramifications

Over the years, many authors have given various definitions for a surrogate marker. Some of the operational ramifications of these definitions will be examined in their relationship to drug development.

Wittes, Lakatos, and Probstfield (1989) defined surrogate endpoint simply as “an endpoint measured in lieu of some so-called ‘true’ endpoint.” While it provides the core, this definition does not provide any operational motivation. Ellenberg and Hamilton (1989) provides this basis by stating: “investigators use surrogate endpoints when the endpoint of interest is too difficult and/or expensive to measure routinely and when they can define some other, more readily measurable endpoint, which is sufficiently well correlated with the first to justify its use as a substitute.” This paved the way to a statistical definition of a surrogate endpoint by Prentice (1989):
“a response variable for which a test of null hypothesis of no relationship to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint.” This definition, also known as the Prentice Criteria, is often very hard to verify in real-life clinical trials. An operating definition given by Temple (1999) states: “a laboratory 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 that is expected to predict the effect of the therapy.” This definition has been used as the operational definition of surrogate endpoints in a regulatory setting.

International Conference on Harmonization (ICH) document E8 states: “a validated surrogate endpoint is an endpoint which allows prediction of a clinically important outcome but in itself does not measure a clinical benefit. When appropriate, surrogate outcomes may be used as primary endpoints.” It further states that the “methods used to make the measurements of the endpoints, both subjective and objective, should meet accepted standards for accuracy, precision, reproducibility, reliability, validity and responsiveness (sensitivity to change over time),” thus providing a valid premise to use it in multinational trials.

A well-validated surrogate will predict the clinical benefit of an intervention both quantitatively and qualitatively with consistent results in several settings. According to Temple (1999), a surrogate endpoint is a laboratory measurement or physical sign used in therapeutic trials as a substitute for a clinically meaningful endpoint that is expected to predict the effect of the therapy. The U.S. Food and Drug Administration (FDA) is able to rely on validated surrogates for accelerated approval of drugs that provide meaningful benefit over existing therapies for serious or life-threatening illnesses (e.g., acquired immunodeficiency syndrome). In these cases, the surrogates should be reasonably likely to predict clinical benefit based on epidemiological, therapeutic, pathophysiologic, or other scientific evidence. However, in general, trials examining surrogate endpoints, even where the endpoint is well correlated with a clinical outcome, surrogates will be unable to evaluate clinically relevant effects of the drug not related to the surrogate, whether these are beneficial or adverse.

3.1.2 Support for Surrogates

Next, we examine what are the motivations for using a surrogate endpoint. The motivation to use a surrogate can be judged by its biological plausibility, its expected success in clinical trials, and its risk-benefit ratio or public health considerations. We summarize these issues in Table 3.1.
### 3. Regulatory Aspects in Using Surrogate Markers in Clinical Trials

| Factor | Favors surrogates | Does not favor surrogates |
|--------|-------------------|--------------------------|
| Biological plausibility | Epidemiological evidence extensive and consistent | Inconsistent epidemiology |
|  | Quantitative epidemiological relationship | No quantitative epidemiological relationship |
|  | Credible animal model shows drug response | No animal model |
|  | Well-understood disease pathogenesis | Pathogenesis not clear |
|  | Drug mechanism of action well-understood | Novel actions not previously studied |
| Success in clinical trials | Effect of surrogate has predicted outcome with other drugs of same pharmacological class | A negative outcome without clear explanation |
|  | Effect on surrogate had predicted outcome in several classes | Inconsistent results across classes |
| Risk-benefit, public health considerations | Serious or life-threatening illness and no alternate therapy | Disease not life-threatening and alternate therapy with different pharmacological action known to affect outcome |
|  | Large safety database | Little known about safety |
|  | Short term use | Long-term use |
|  | Difficulty in studying clinical endpoint (rare, delayed) | Easy to study clinical endpoint |
|  | | Long-delayed, small effect in healthy people |

Thus, a surrogate to be useful has to have unequivocal biological plausibility, be expected to perform consistently in a clinical trial, and possess superior public health benefits.

### 3.1.3 Criteria for Surrogate Markers To Be Used in Drug Development

In epidemiological studies, a useful surrogate marker is a causal factor for the disease of interest, not merely a correlated factor. As Fleming (1996) stated, “a correlate does not a surrogate make.” The higher the level of explanatory evidence the surrogate is able to carry, the better it is to explain the disease process. Table 3.2 summarizes the relationships surrogate endpoints (SEP) can have with the “true” clinical endpoints (CE).

Then, sensitivity \((SE)\) of the surrogate endpoint for the clinical endpoint
TABLE 3.2. Relationship of surrogate endpoints with the clinical endpoints.

|       | T good | T poor | Total |
|-------|--------|--------|-------|
| S good| a      | b      | a+b   |
| S poor| c      | d      | c+d   |
| Total | a+c    | b+d    | N     |

\[a = \text{number of patients where both } S \text{ and } T \text{ provide good disease characterization.}\]

\[b = \text{number of patients where } S \text{ is good but } T \text{ provide poor disease characterization.}\]

\[c = \text{number of patients where } S \text{ is poor but } T \text{ provide good disease characterization.}\]

\[d = \text{number of patients where both } S \text{ and } T \text{ provide poor disease characterization.}\]

is defined by

\[SE = \frac{a}{a+c}. \tag{3.1}\]

Specificity (SP) of the surrogate endpoint for the clinical endpoint is defined by

\[SP = \frac{d}{b+d}. \tag{3.2}\]

For the surrogate to be useful, both sensitivity and specificity have to be numerically close to 1.

The relative risk (RR) is defined as

\[RR = \frac{a(c+d)}{c(a+b)}. \tag{3.3}\]

and the attributable proportion (AP) as

\[AP = \frac{SE}{1 - \frac{1}{\text{RR}}}. \tag{3.4}\]

For a surrogate marker to be a successful one, AP has to be numerically close to 1. Schatzkin, Freedman, and colleagues proposed strategies for determining whether a biomarker is a valid surrogate for a disease of interest, for instance whether human papillomavirus infection is a valid surrogate for cervical dysplasia. The attributable proportion is a useful measure of association between the surrogate endpoint and the clinical endpoint, but establishing the causality of the relationship between the surrogate and the clinical endpoint require data from either observational studies or, preferably, intervention studies. Intervention studies would focus on the triplet intervention / biomarker / disease in much the same way as a clinical trial would focus on the triplet treatment / surrogate endpoint / clinical endpoint.
3. Regulatory Aspects in Using Surrogate Markers in Clinical Trials

3.1.4 Surrogate Markers and Biomarkers

Definitions and Differences

*Biological marker* or biomarker, as more commonly known, refers to a variety of physiologic, pathologic, or anatomic measurements that are thought to relate to some aspect of normal or pathological biologic processes (Temple 1995, Lesko and Atkinson 2001). These biomarkers include measurements that suggest the etiology of, the susceptibility to, or the progress of disease; measurements related to the mechanism of response to treatments; and actual clinical responses to therapeutic interventions. Biomarkers differ in their closeness to the intended therapeutic response or clinical benefit endpoints, classified as follows:

1. biomarkers thought to be valid surrogates for clinical benefit (e.g., blood pressure, cholesterol, viral load);

2. biomarkers thought to reflect the pathologic process and be at least candidate surrogates (e.g., brain appearance in Alzheimer’s disease, brain infarct size, various radiographic/isotopic function tests);

3. biomarkers reflecting drug action but of uncertain relation to clinical outcome (e.g., inhibition of ADP-dependent platelet aggregation, ACE inhibition);

4. biomarkers that are still more remote from the clinical benefit endpoint (e.g., degree of binding to a receptor or inhibition of an agonist).

From a regulatory perspective, a biomarker is not considered an acceptable surrogate endpoint for a determination of efficacy of a new drug unless it has been empirically shown to function, as a valid indicator of clinical benefit (i.e., a valid surrogate). Theoretical justification alone does not meet the evidentiary standards for market access. Many biomarkers will never undergo the rigorous statistical evaluation that would establish their value as a surrogate endpoint to determine efficacy or safety, but they can still have use in earlier drug development process. Changes in biomarkers typically exhibit a time course that is different from changes in clinical endpoints and often are more directly related to the time course of plasma drug concentrations, possibly with a measurable delay. For this reason, exposure-response relationships based on biomarkers may help establish the dose range for clinical trials intended to establish efficacy that will then be studied more formally, indicate how soon dose titration should occur, examine potential pharmaco-dynamic interactions, and give insight into potential adverse effects.
TABLE 3.3. Biomarkers as surrogate endpoints – possible relationships.

| Type of relationship                                                                 | Value of biomarker                                      | Example                                                                 |
|-------------------------------------------------------------------------------------|--------------------------------------------------------|------------------------------------------------------------------------|
| Unreliable interaction between biomarker and the treatment intervention              | Biomarker is of no value as a surrogate endpoint        | Prostate-specific antigen (PSA) is a useful biomarker for prostate cancer detection but unreliable as an indicator of treatment response |
| The full effect of the intervention is observed through the biomarker assessment     | Biomarker is an ideal surrogate endpoint                | None known at present                                                  |
| Intervention affects the endpoint and the biomarker independently; only a proportion of the treatment effect is captured by the surrogate endpoint | Biomarker has value as a surrogate endpoint but explains only a part of the treatment effect | Most established surrogate endpoints (e.g., development of opportunistic infections with HIV anti-viral and mortality) |
| Intervention affects favorably on the biomarker but unfavorably on the well-state and disease | Biomarker is of little practical use as a surrogate endpoint but may have utility in exploratory studies | Suppression of ventricular ectopy as a biomarker of fatal arrhythmia following myocardial infarctions (CAST trial) |

Relationship Between Biomarkers and Surrogate Markers

While all surrogate markers are biomarkers, it is likely that only a few single biomarkers will qualify as surrogate endpoints in therapeutic intervention trials, or as surrogate markers in natural history or epidemiological studies. For ease of reference, we use the terms surrogate “markers” and surrogate “endpoints” interchangeably, although we acknowledge that some surrogate endpoints (such as patient self-assessment scales) are not biomarkers. For the concept of a surrogate endpoint to be useful, one must specify the clinical endpoint, class of intervention, and population in which the substitution of the biomarker for a clinical endpoint is considered reasonable. Table 3.3 summarizes the various possible relationships that can exist between a surrogate marker and a biomarker.

Figure 3.1 gives a schematic description of the conceptual model for surrogate endpoints and biomarkers.

It shows that only a small proportion of biomarkers will be useful to be considered as a surrogate endpoint, which will then have to be subjected to a rigorous set of risk-benefit considerations to eventually arrive as an instrument of global intervention assessment.
3.2 Surrogate Markers in Regulatory Setting

The U.S. Food and Drug Administration has supported the use of surrogate markers when clinically appropriate to bring therapeutic agents through the approval process faster and in a more efficient way. If a surrogate endpoint can be measured more easily or efficiently or with higher precision, then it translates into faster treatment access for the patients. If a surrogate endpoint is less affected by other treatment modalities, then the precision of the trial can also be expected to increase. The FDA has responded to faster approval of promising through various specific regulatory mechanisms, which we will discuss now.

3.2.1 Fast Track Program – A Program for Accelerated Approval

Fast Track programs at the U.S. Food and Drug Administration are designed to facilitate the development and expedite the review of new drugs that meet two criteria: (1) are intended to treat serious or life-threatening conditions and (2) demonstrate the potential to address unmet medical needs for the condition. Whether a condition is serious or not is a matter of judgment, but is generally based on its impact on such factors such as survival, day-to-day functioning, or the likelihood that the disease if left untreated would progress from a less severe condition to a more serious one. When focusing on morbidity, consideration is given to its persistence or recurrence if it is not irreversible (57 Federal Register 13234 dated April 15, 1992). Whether a therapeutic agent is intended to treat a serious condition
is determined by the following criteria: (1) a therapy directed at serious symptoms or serious manifestations of the condition; (2) a diagnostic evaluated for the impact on a serious aspect of the condition; (3) a preventive intended to prevent a serious aspect; (4) a product that could ameliorate serious side effects of other treatments. Now let us discuss the second criterion. For an agent to demonstrate potential to address unmet medical needs, the following conditions have to be considered: (1) there is no existing therapy for the condition; (2) the new therapy is better; (3) the new therapy is for the patients intolerant or unresponsive to existing therapy; (4) the new therapy is less toxic, but preserves similar benefit; (5) the new therapy improves compliance which is shown to improve effects on serious conditions. Details on the designation, development and application review of a fast track therapy can be found at http://www.fda.gov/cder/guidance.

Figure 3.2 summarizes the criteria in a schematic format.

Fast Track emphasizes the critical nature of close early communication between the FDA and the sponsor. It highlights the procedures such as pre-trial (before Investigational New Drug (IND) is initiated) and End-of-Phase I meetings as methods to improve the efficiency of pre-clinical and clinical development. It focuses on efforts by the FDA and the sponsor to reach early agreement on the design and analysis of the major clinical efficacy studies that will be needed to support approval. The requests for Fast Track are expected to be resolved within a designated 60-day period from the initial request. As of June 30, 2002, 151 applications for Fast Track have been submitted. As seen from Table 3.4, there has been a marked increase in fast track designation requests in recent years and most requests have been acted upon within the 60-day period.
3. Regulatory Aspects in Using Surrogate Markers in Clinical Trials

TABLE 3.4. Responses to request for Fast Track designation.

| Fast track requests | Granted | Denied | Pending | Total (%) |
|---------------------|---------|--------|---------|-----------|
| Submissions 1998–2002 |         |        |         |           |
| Within goal of 60 days | 96      | 30     | 4       | 130 (79.3%) |
| Overdue (>60 days)    | 13      | 12     | 9       | 34 (20.7%)  |

| Submissions in fiscal year 2004 |       |        |         |           |
|-------------------------------|-------|--------|---------|-----------|
| Within goal of 60 days        | 12    | 7      | 7       | 26 (96.0%) |
| Overdue (>60 days)            | 1     | 0      | 0       | 1 (4.0%)   |

NOTE: These figures are for sponsor requests for ‘Fast Track’ designation for a specific drug product and indication, which is not necessarily the same as a product being granted Approval under Subpart H. Report updated through April 28, 2004.

3.2.2 Subpart H and Its Relevance to Surrogate Markers

So far, from the discussion of the process, it is indicated that Fast Track can be considered irrespective of whether surrogate endpoints were used or not. For drug development programs specifically utilizing surrogate endpoints, a special regulatory mechanism called Subpart H (refers to the specific code of regulations governing it) is available. Under Subpart H, approval may be based on a surrogate endpoint or on an effect on a clinical endpoint other than survival or irreversible morbidity (“Surrogate”) [21 Code of Federal Register (CFR) 314.510 and 21 CFR 601.41], or a product may be approved with restrictions to assure safe use (“Restricted”) [21 CFR 314.520]. Note that Subpart H applications are usually candidates for Fast Track also, but not necessarily so.

The 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 epidemiological, 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. Approval under this section will be subject to the requirement that the applicant study the drug further, to verify and describe its clinical benefit, where there is uncertainty as to the relation of the surrogate endpoint to clinical benefit, or of the observed clinical benefit to ultimate outcome. Post-marketing studies would usually be studies al-
ready underway. When required to be conducted, such studies must also be adequate and well controlled. The applicant shall carry out any such studies with due diligence.”

Tables 3.5–3.7 summarize New Drug Applications (NDAs) that have been approved under Subpart H regulations.

Tables 3.8 and 3.9 summarize already approved drugs that were considered for a different disease indication using surrogate markers. These applications are known in regulatory parlance as NDA Supplements.

Fast Track policies are primarily designed to expedite drug development during the IND stage, whereas Approval Under Subpart H allows for marketing approval of an NDA based on an effect on a surrogate endpoint along with well-controlled post-marketing studies.

A post-approval study will not necessarily be required in the exact population for which approval was granted. For example, where a product was approved to treat patients with refractory malignancy, additional information from that population may not, for example, be as useful as randomized controlled trials in a previously untreated population. In many instances, additional studies would be already under way at the time the accelerated approval is granted. If such studies are adequate and well controlled (either utilizing proper historical controls or randomization), they may fulfill the accelerated approval requirements for post-approval studies. All required post-approval studies should be carried out with due diligence. Failure to do so would constitute grounds to withdraw approval of the product application (21 CFR 314.530(a) or 21 CFR 601.43(a)). FDA may also withdraw approval of the application if studies fail to demonstrate clinical benefit based on the traditional long-term endpoint.

Next, three therapeutic areas where surrogate markers have been used will be discussed – in anti-viral, anti-cancer, and cardiovascular drug products. It is not meant to be an exhaustive treatise; there are several other therapeutic classes where use of surrogate markers is being considered or done. However, the experience is most established in these three areas.
TABLE 3.5. NDAs approved under Subpart H based on surrogate endpoints. Part I.

| NDA   | Trade name | Generic name                  | Approval date | Indication for treatment                                                                 |
|-------|------------|-------------------------------|---------------|------------------------------------------------------------------------------------------|
| 20199 | Hivid      | Zalcitabine                   | 19-Jun-92     | Combination therapy with zidovudine in advanced HIV infection.                           |
| 50698 | Biaxin     | Clarithromycin (suspension)   | 23-Dec-93     | Disseminated mycobacterial infections due to Mycobacterium avium and Mycobacterium intracellular. |
| 20412 | Zerit      | Stavudine                     | 24-Jun-94     | Adults with advanced HIV infection – alternative therapy.                               |
| 20212 | Zinecard   | Dexrazoxane                   | 26-May-95     | To reduce the incidence and severity of cardiomyopathy associated with doxorubicin administration in certain breast cancer patients. |
| 20498 | Casodex    | Bicalutamide                  | 04-Oct-95     | Use in combination therapy with a Luteinizing-Hormone Releasing Hormone (LHRH) analogue for the treatment of advanced prostate cancer. |
| 20564 | Epivir     | Lamivudine                    | 17-Nov-95     | HIV infection in selected patients.                                                     |
| 50718 | Doxil      | Dexamethasone hydrochloride (liposomal formulation) | 17-Nov-95 | AIDS-related Kaposi’s sarcoma in patients with disease that has progressed on prior combination chemotherapy or in patients who are intolerant to such therapy. |
| 20628 | Invirase   | Saquinavir mesylate           | 06-Dec-95     | Advanced HIV infection in selected patients in combination with nucleoside analogues.   |
| 20659 | Norvir     | Ritonavir                     | 01-Mar-96     | In combination with nucleoside analogues or as monotherapy for the treatment of HIV infection. |
| 20685 | Crixivan   | Indinavir sulfate             | 13-Mar-96     | HIV infection in adults.                                                                |
| 20449 | Taxotere   | Docetaxel                     | 14-May-96     | Patients with locally advanced or metastatic breast cancer who have progressed or relapsed during anthracycline based therapy. |
| 20571 | Camptosar  | Irinotecan hydrochloride      | 14-Jun-96     | Refractory colorectal cancer.                                                          |
| 20636 | Viramune   | Nevirapine                    | 21-Jun-96     | Combination with nucleoside analogues for the treatment of HIV-1 infected adults who have experienced clinical and/or immunologic deterioration. |
| 20604 | Serostim   | Somatropin                    | 23-Aug-96     | AIDS wasting associated with catabolism loss or cachexia.                               |
| 19815 | ProAmatine | Midodrine hydrochloride       | 06-Sep-96     | Treatment of symptomatic orthostatic hypotension.                                       |
| 20778 | Viracept   | Nelfinavir mesylate           | 14-Mar-97     | HIV infection when therapy is warranted.                                                |
| 20705 | Rescriptor | Delavirdine mesylate          | 04-Apr-97     | HIV infection in combination with appropriate antiretroviral agents when therapy is warranted. |
| 20896 | Xeloda     | Capecitabine                  | 30-Apr-98     | Patients with metastatic breast cancer resistant to both paclitaxel and an anthracycline-containing chemotherapy regimen or resistant to paclitaxel and for whom further anthracycline therapy may be contraindicated. |

Updated through April 30, 2004.
| NDA     | Trade name    | Generic name        | Approval date | Indication for treatment                                                                 |
|---------|---------------|---------------------|---------------|-----------------------------------------------------------------------------------------|
| 19832   | Sulfamylon    | Mafenide acetate    | 05-Jun-98     | As an adjunctive topical antimicrobial agent to control bacterial infection when used under moist dressings over meshed autografts on excised burn wounds. |
| 21024   | Priftin       | Rifapentine         | 22-Jun-98     | Pulmonary tuberculosis (TB).                                                            |
| 20933   | Viramune      | Nevirapine          | 11-Sep-98     | Provides for an oral suspension, which is indicated for use in combination therapy with other antiretroviral agents for the treatment of HIV-1 infection. |
| 20972   | Sustiva       | Efavirenz           | 17-Sep-98     | In combination with other antiretroviral agents for the treatment of HIV-1 infection.   |
| 20977, 20978 | Ziagen     | Abacavir sulfate   | 17-Dec-98     | In combination with other antiretroviral agents, for the treatment of HIV-1 infection.  |
| 21041   | Depocyt       | Cytarabine liposomal injection | 01-Apr-99   | Intrathecal treatment of lymphomatous meningitis.                                      |
| 21029   | Temozolomide  | Temozolomide        | 11-Aug-99     | Adult patients with refractory anaplastic astrocytoma, i.e., patients at first relapse who have experienced disease progression on a drug regimen containing a nitrosourea and procarbazine. |
| 21007, 21039 | Agenerase   | Amprenavir          | 15-Apr-99     | In combination with other antiretroviral agents, for the treatment of HIV-1 infection.  |
| 50747   | Synercid      | Quinupristin/dalfopristin I.V. | 21-Sep-99 | Vancomycin-resistant Enterococcus faecium.                                              |
| 21174   | Mylotarg      | Gemtuzumab/ozogamicin | 17-May-00   | Patients with CD33 positive acute myeloid leukemia in first relapse who are 60 years of age or older and who are not considered candidates for cytotoxic chemotherapy. |
| 21226, 21251 | Kaletra     | Lopinavir/ritonavir | 15-Sep-00     | In combination with other antiretroviral agents for the treatment of HIV-1 infection in adults and pediatric patients age six months and older. |
| 21205   | Trizivir      | Abacavir sulfate, lamivudine, & zidovudine | 14-Nov-00 | Either alone or in combination with other antiretroviral agents for the treatment of HIV-1 infection. |
| 21335   | Gleevec       | Imatinib mesylate   | 10-May-01     | Use of 50 and 100 mg capsules for the treatment of patients with chronic myeloid leukemia (CML) in blast crisis, accelerated phase, or in chronic phase after failure of interferon-alpha therapy. |
| 21356   | Viread        | Tenofovir disoproxil fumarate | 26-Oct-01 | In combination with other antiretroviral agents for the treatment of HIV-1 infection in adults. |
| 21272   | Remodulin     | Treprostinil sodium | 21-May-02     | Use of 1.0, 2.5, 5.0, and 10.0 mg/ml injection for the treatment of pulmonary arterial hypertension (PAH). |

*Updated through April 30, 2004.*
### TABLE 3.7. NDAs approved under Subpart H based on surrogate endpoints. Part III.

| NDA   | Trade name | Generic name                  | Approval date | Indication for treatment                                                                                                                                                                                                 |
|-------|------------|-------------------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 21196 | Xyrem      | Sodium oxybate                | 17-Jul-02     | Provides for the use of Xyrem Oral Solution for the treatment of cataplexy associated with narcolepsy (restricted use, not on surrogate endpoint.)                                                                     |
| 21492 | Eloxatin   | Oxaliplatin injection         | 9-Aug-02      | Provides for the use of Eloxatin in combination with infusional fluorouracil/leukovorin (5-FU/LV) for the treatment of patients with metastatic carcinoma of the colon or rectum whose disease has recurred or progressed during or within 6 months of completion of first line therapy with the combination of bolus 5-FU/LV and irinotecan. |
| 21481 | Fuzeon     | Enfuvirtide injection         | 13-Mar-03     | Provides for the use of Fuzeon in combination with other antiretroviral agents, for the treatment of HIV-1 infection in treatment experienced patients with evidence of HIV-1 replication despite ongoing antiretroviral therapy. |
| 21588 | Gleevec    | Imatinib mesylate tablets     | 18-Apr-03     | Provides for the use of Gleevec for the treatment of patients with chronic myeloid leukemia (CML) in blast crisis, accelerated phase, or in chronic phase after failure of interferon-alpha therapy.                          |
| 21399 | Iressa     | Gefitinib tablets             | 5-May-03      | Provides for the use of IRESSA as monotherapy for the treatment of patients with locally advanced or metastatic non-small cell lung cancer after failure of both platinum-based and docetaxel chemotherapies.                         |
| 21602 | Velcade    | Bortezomib injection          | 13-May-03     | Provides for the use of Velcade for the treatment of multiple myeloma patients who have received at least two prior therapies and have demonstrated disease progression on the last therapy.                             |
| 21320 | Plenaxis   | Abarelix injectable suspension| 25-Nov-03     | Provides for the use of Plenaxis for the palliative treatment of men with advanced symptomatic prostate cancer and specific symptoms (restricted use, approval not based on surrogate).                      |

*Updated through April 30, 2004.*
TABLE 3.8. *NDA Supplements approved under Subpart H based on surrogate endpoints. Part I.*

| NDA  | Supp | Trade name | Generic name | Approval date | Indication for treatment |
|------|------|------------|--------------|---------------|--------------------------|
| 50697 | N    | Biaxin     | Clarithromycin (tablets) | 23-Dec-93 | Disseminated mycobacterial infections due to Mycobacterium avium and Mycobacterium intracellular. |
| 20636 | SE1 009 | Viramune | Nevirapine | 11-Sep-98 | Provides for the inclusion of pediatric information into the labeling. |
| 50718 | SE1 006 | Doxil | Doxorubicin hydrochloride (liposomal formulation) | 28-Jun-99 | Treatment of metastatic carcinoma of the ovary in patients with disease that is refractory to both paclitaxel- and platinum-based chemotherapy regimens. |
| 21156 | N    | Celebrex  | Celecoxib | 23-Dec-99 | To reduce the number of adenomatous colorectal polyps in Familial Adenomatous Polyposis (FAP), as an adjunct to usual care. |
| 19537 | SE1 038 | CIPRO | Ciprofloxacin hydrochloride | 30-Aug-00 | Inhalational anthrax (post-exposure). |
| 19847 | SE1 024 | CIPRO | Ciprofloxacin hydrochloride | 30-Aug-00 | Inhalational anthrax (post-exposure). |
| 19857 | SE1 027 | CIPRO | Ciprofloxacin hydrochloride | 30-Aug-00 | Inhalational anthrax (post-exposure). |
| 19858 | SE1 021 | CIPRO | Ciprofloxacin hydrochloride | 30-Aug-00 | Inhalational anthrax (post-exposure). |
| 20780 | SE1 008 | CIPRO | Ciprofloxacin hydrochloride | 30-Aug-00 | Inhalational anthrax (post-exposure). |
| 21335 | SE1 001 | Gleevec | Imatinib mesylate | 1-Feb-02 | Patients with Kit (CD117) positive unresectable and/or metastatic malignant gastrointestinal stromal tumors (GIST). |

*Updated through April 30, 2004.*
TABLE 3.9. NDA Supplements approved under Subpart H based on surrogate endpoints. Part II.

| NDA  | Supp  | Trade name | Generic name          | Approval date | Indication for treatment                                                                                     |
|------|-------|------------|-----------------------|---------------|-------------------------------------------------------------------------------------------------------------|
| 21107| SE8 005| Lotronex   | Alosetron hydrochloride| 7-Jun-02      | Restricted use of Lotronex for women only with severe diarrhea-predominant irritable refractory to conventional therapy (not based on surrogate). |
| 20541| SE1 010| Arimidex   | Anastrozole tablets   | 5-Sep-02      | Provides for the use of ARIMIDEX for adjuvant treatment of postmenopausal women with hormone receptor positive early breast cancer. |
| 21335| SE1 004| Gleevec    | Imatinib mesylate 100 mg capsules | 20-Dec-02 | Provides for the use of Gleevec for the treatment of newly diagnosed adult patients with Philadelphia chromosome positive chronic myeloid leukemia (CML). Follow-up is limited. |
| 21335| SE5 003| Gleevec    | Imatinib mesylate tablets | 20-May-03 | Provides for the use of Gleevec for the treatment of pediatric patients with Ph+ chronic phase CML whose disease has recurred after stem cell transplant or who are resistant to interferon alpha therapy. |

*Updated through April 30, 2004.*
3.3 Use of Surrogate Markers in Anti-viral Drug Products

Surrogate markers have been widely used in anti-viral drug therapies. It is one of the first areas that surrogates were used, as a response to the AIDS epidemic and the thrust to bring potential therapies to the market within the earliest time frame.

Various biological markers have been considered during early drug development processes in anti-HIV therapies. They included CD4 count, p24 and ICD p24 antigen level, $\beta_2$-microglobulin, neopterin, HIV-1 RNA, HIV-1 DNA among a few. The cumulative evidence base suggests that both CD4 count and HIV-1 RNA provide important prognostic factor for AIDS. Some surrogate markers such as $\beta_2$-microglobulin and neopterin have proved to be of limited use in a clinical trial.

It has been indicated that the natural history of HIV-1 infection can be characterized by increased HIV-1 RNA level leading to CD4 count depletion which in turn leads to AIDS and eventually death. Following the initial HIV-1 infection, there is a latency period of up to 7 years where little virus is detected in the blood but there is still virus particles being produced on a daily basis. It was thought that if virus replication can be completely blocked by potent anti-retroviral drug combinations, it would take between two and three years of treatment to completely eradicate the virus from the infected host.

HIV-1 RNA as a surrogate endpoint has several unique properties. First, HIV-1 RNA is a marker of the severity of the disease – the higher it is, the more severe the infection. Second, it has been shown repeatedly that AIDS-defining illness is much less frequent when HIV-1 RNA is below a certain threshold, e.g., 5000 copies/ml. Third, HIV-1 RNA is usually high at the time of initial HIV-1 infection, and often increases near the time of an AIDS-defining illness such as an opportunistic infection. Following the 1997 NIH workshop and the subsequent publication of two guidance documents by the Department of Health and Human Services, the consensus was to monitor HIV-1 viral load and CD4 count of HIV-infected patients on a routine basis to make treatment decisions.

In August 1999, FDA issued a draft Guidance for Industry discussing Clinical Considerations for Accelerated and Traditional Approval of Anti-Retroviral Drugs Using Plasma HIV RNA (http://www.fda.gov/cder/guidance/index.htm). Although accelerated approvals are routinely based on changes in endpoints such as CD4 cell counts and plasma HIV RNA levels, clinical endpoint trials assessing effects on mortality and/or disease progression had been a requirement for traditional approvals prior to July
1997. With the availability of potent anti-retroviral drug regimens and sensitive assays for assessing plasma HIV RNA, the standards of clinical practice evolved to a paradigm emphasizing maximal and durable HIV RNA suppression.

To evaluate feasibility of using HIV-1 RNA as a study endpoint, a collaborative group of pharmaceutical, academic and government scientists investigated relationships between treatment-induced changes in HIV-1 RNA and clinical endpoints from ongoing and completed anti-retroviral trials. In several analyses of multiple trials involving more than 5000 patients, a clear association was seen between initial decreases in plasma HIV-1 RNA within first 24 weeks, and a reduction in the risk of clinical progression and death. This relationship was observed across a range of patient characteristics including pretreatment CD4 counts and HIV-1 RNA levels, prior drug experience, and treatment regimen. Based on these data, it was proposed that the accelerated approvals could be based on studies that show a drug’s contribution toward shorter-term reductions in HIV-1 RNA (e.g., 24 weeks), whereas traditional approvals could be based on trials that show a drug’s contribution toward durability of HIV-1 RNA suppression (e.g., at least 48 weeks). In addition, the changes in CD4 cell counts need to be consistent with observed HIV-1 RNA changes (Hughes et al. 2000).

According to the 1999 Guidance, studies in a broad range of patient populations (gender, age, and race) and a range of pre-treatment characteristics (e.g., advanced and early disease, heavily pre-treated and treatment naïve) are recommended to characterize the activity of the drug in at least two adequate and well-controlled trials with a minimum of 24 weeks duration to support accelerated approval. In combination therapies, analyses at some earlier time points (e.g., 16 weeks) have proven to be less discriminatory. Every attempt is to be made to design randomized, blinded, controlled trials that provide all study patients with treatment regimens according to a standard clinical practice. If the studies are designed as superiority trials, add-on or substitution comparisons can be included, where the regimen with the experimental drug should show superiority to the control regimen. If equivalence trials using substitution comparisons are to be designed, it is important that the contribution of the substituted drug to the regimen’s overall activity be previously characterized in the population of interest.

Historically, zidovudine (ZDV) was approved in 1987 based on 17 weeks survival. The next product, didanosine (ddI) was approved in 1991 based on surrogate endpoint of CD4 counts with a limited indication in patients refractory to AZT failures. It was not until 1992 that the accelerated approval mechanism was used in the approval of dideoxycytidine (ddC). Since then many other HIV drugs have been approved under this regulation. For approvals prior to 1995, the accelerated approval was based on either change
in CD4 count or time-averaged change in CD4 count (DAVG). Between 1995–1998, HIV-1 RNA load was gradually being more frequently used. The metric used for HIV-1 RNA included change from baseline, DAVG or the percentage of patients below a certain threshold. After 1998, most of the accelerated approvals have been based on the criteria of having HIV-1 RNA <400 and/or 50 copies/ml. The endpoints used in traditional approvals of anti-HIV agents were primarily based on disease progression (DP) prior to 1997. From 1997 onwards, the traditional approvals are mostly based on HIV-1 RNA, either as percentage of patients having less than 400 copies/ml or the time to virologic failure.

According to Gilbert et al. (2001), the selection of primary endpoints for AIDS trials is complicated by the long clinical course of the disease, the frequent onset of anti-viral drug resistance, and the limitations in data for validating surrogate endpoints. However, increasing the objectivity of the selection process in the future requires expansion of available information for the elucidation of the complex relationship between various surrogate endpoints and clinical endpoints. Only through vigilant collection of clinical outcomes data (e.g., through routine collection of death event data from national death records) and data from long-term studies that monitor virologic, immunologic, and clinical information throughout sequences of regimens can this goal be achieved.

Figure 3.3 summarizes the endpoints traditionally used in accelerated and traditional approval of anti-HIV drugs. The endpoints on the left axis refer to the surrogate endpoints used for accelerated approval; endpoints on the right axis refer to the clinical endpoints used for traditional approval. The horizontal axis gives the approval timelines.
TABLE 3.10. Crixivan. Basis for accelerated approval (CD4 count: comparison of MK-containing arms to ZDV).

| Statistic | Crixivan (MK) vs. zidovudine (ZDV) | MK+ZDV vs. ZDV |
|-----------|-----------------------------------|---------------|
| Study 028 | Difference                        | 66            | 69            |
|           | \( p \)-value                     | \(<0.0001\)   | \(<0.0001\)   |
|           | 95\% CI                           | 42-89         | 45-93         |
| Study 033 | Difference                        | 62            | 47            |
|           | \( p \)-value                     | \(<0.0001\)   | \(<0.0001\)   |
|           | 95\% CI                           | 40-84         | 25-69         |

Next, we discuss two examples of therapeutic agents that have undergone the accelerated approval and eventually went through the traditional approval.

3.3.1 Crixivan: A Case Study

Crixivan (indinavir sulfate), also referred to as MK-639 or simply MK or IDV, was submitted in 1996 for accelerated approval based upon change from baseline in CD4 cell counts. Change from baseline of HIV-1 RNA was also considered as a secondary endpoint.

Two Phase III studies (Study 028 and 033) were examined for this review, (see Table 3.10), and the regulatory decision was based on the interim analyses of the surrogate markers. Study 028 was a double-blind study in 224 patients with no prior nucleoside analogue experience. The patients were randomized to receive one of the three treatment regimens — the test drug (MK)+Zidovudine (ZDV)+ddI, MK monotherapy or ZDV+ddI. The comparisons of each arm containing MK versus the control arm were conducted using ANOVA adjusting for center and CD4 strata at baseline.

Study 033 was performed in 266 subjects with prior ZDV experience and was randomized to one of the three regimens — MK+ZDV+lamivudine (also known as 3TC), MK monotherapy and ZDV+3TC. Analysis plans were similar to Study 028.

Two short term (24-week) Phase II studies were also examined in order to provide preliminary efficacy information regarding triple combination therapy. It was seen that the results were convincing enough to warrant
accelerated approval.

Crixivan was submitted for traditional approval following completion of the pivotal trials (see Table 3.11). The clinical endpoint was defined as the first occurrence of death from any cause or the diagnosis of AIDS as predefined in the protocol. The comparisons were to be based on time-to-first failure methods, including Kaplan-Meier, log-rank test and Cox proportional hazards regression models. Study 033, later conducted as AIDS Cooperative Trial Group (ACTG) 320, used the area under the response-time curve for each patient divided by the time from randomization to the last available evaluation of the patient minus the baseline value (AUCMB). The ACTG Data and Safety Monitoring Board (DSMB) monitored the course and conduct of this study. One interim look was planned after 250 events or one year, and the Peto and Pike stopping boundary was used. The trial was to be considered for early stopping if the nominal $p$-value $<0.001$. The trial was indeed stopped early by DSMB after 1156 patients were enrolled.

When considered in the light of the results of Trial 028 and the patterns seen over time in ACTG 320, it appeared that the failure to reach the traditional 0.05 level is the result of the premature discontinuation of ACTG 320. The achieved significance level was still felt to be sufficient to support the results of study 028 that Crixivan is associated with a reduction in rate of progression or death due to HIV.

### 3.3.2 Viramune: A Case Study

Accelerated Approval

Viramune (nevirapine, or NVP) belongs to a new class of anti-retroviral agents called non-nucleoside reverse transcriptase inhibitor (NNRTI). The accelerated approval of this drug was sought in patients with advanced HIV-1 infection whose current anti-retroviral therapy is no longer deemed adequate. Three studies, two in nucleoside experienced population and one in nucleoside naive population, were submitted under accelerated approval to support the claim that the addition of nevirapine to one or more nucleoside drugs provides an improvement in surrogate markers for HIV disease (see Table 3.12). For each study, the surrogate endpoints were CD4 cell count and HIV-1 RNA level in an eight-week window of time at the end of the studies.

Study 1037 was a randomized, double blind, placebo-controlled study comparing ZDV/NVP to NVP monotherapy in 60 patients with prior ZDV experience for 3–24 months and CD4 cell counts between 200 and 500. Subjects were followed for 28 weeks with scheduled visit every 2 weeks in
TABLE 3.11. *Crixivan. Basis for traditional approval (time to first clinical event analysis – treatment comparison).*

| Study       | Treatment comparison | Stratified log-rank test (two-sided $p$-value) for time to first clinical event factor |
|-------------|----------------------|--------------------------------------------------------------------------------------|
| Study 028   | MK+ZDV vs ZDV        | 0.0001                                                                               |
|             | MK vs ZDV            | 0.0001                                                                               |
|             | MK+ZDV vs MK         | 0.22                                                                                 |
| ACTG 320    | MK+3TC+ZDV vs 3TC+ZDV| 0.0021*                                                                               |

* From randomization-based test, required 0.001 to achieve 5% level.

the beginning and every 4 weeks after the fourth week.

Study 1031 (ACTG 241) was a randomized, double blind, placebo-controlled study in 400 patients comparing ZDV/ddI/NVP to ZDV/NVP with similar schedule as Study 1037. Eight of the 16 participating centers, with 200 patients, were to be included in a virology substudy, in which HIV-1 RNA were to be collected in addition. The subjects were to be followed for 48 weeks on CD4 count.

Study 1046 was an international randomized, double blind placebo-controlled trial in 120 patients comparing ZDV/ddI/NVP to ZDV/NVP and ZDV/ddI with same dosing regimen for 52 weeks after the start of therapy.

The studies were analyzed using ANOVA models with baseline CD4 strata and center as covariates.

It is seen that addition of nevirapine to one or nucleosides has been shown to produce an increase in CD4 cell counts and a small decrease in HIV-1 RNA levels. The lack of significance in Study 1046 may be attributed to the much smaller sample size (50/arm versus 200/arm).

Traditional Approval

For traditional approval, the sponsor submitted five randomized, controlled clinical trials. Study 1090, the Atlantic trial, and another trial with the acronym INCAS, were planned pivotal trials: trials ACTG 193a and ACTG 241 were provided as supportive evidence.

Trial 1090 was a placebo-controlled study designed to compare efficacy of
TABLE 3.12. *Viramune.* Basis for accelerated approval.

| Endpoint Metric | N | Treatment | Control | \( p \)-value |
|-----------------|---|-----------|---------|-------------|
| Study 1031      |   | Z/D/N*    | Z/D     |             |
| Mean change week 20-28 | 328 | 26        | -5      | .001        |
| Mean change week 40-48 | 328 | 6         | -16     | .002        |
| AUCMB week 28  | 392 | 23        | 6       | .001        |
| AUCMB week 48  | 392 | 20        | 0       | .001        |
| RNA            |   | Mean change week 20-28 | 155 | -0.27 | -0.08 | .137 |
| RNA            |   | Mean change week 40-48 | 149 | -1.14 | 1.11  | .024 |
| RNA            |   | AUCMB week 28 | 188 | -0.57 | -0.27 | .001 |
| RNA            |   | AUCMB week 48 | 188 | -0.43 | -0.17 | .003 |
| Study 1037      |   | Z/N       | Z       |             |
| Mean change week 12-16 | 55 | 53        | -31     | .001        |
| Mean change week 20-28 | 55 | 14        | -31     | .009        |
| AUCMB week 16  | 60  | 44        | -11     | .001        |
| AUCMB week 28  | 60  | 22        | -24     | .001        |
| RNA            |   | Mean change week 12-16 | 55 | 0.03 | 0.01 | 0.525 |
| RNA            |   | Mean change week 20-28 | 55 | 0.16 | 0.12 | 0.590 |
| RNA            |   | AUCMB week 16 | 60 | -0.38 | -0.01 | .001 |
| RNA            |   | AUCMB week 28 | 60 | -1.16 | -0.04 | .001 |
| Study 1046      |   | Z/D/N     | Z/D     | Z/N        | \( p \)-values |
| CD4            |   | Mean change week 12-16 | 117 | 95 | 44 | .44 | .08 | .01 |
| CD4            |   | Mean change week 20-28 | 113 | 78 | 22 | .18 | .05 | .001 |
| CD4            |   | AUCMB week 16 | 72 | 62 | 57 | .58 | .77 | .39 |
| CD4            |   | AUCMB week 28 | 87 | 67 | 47 | .23 | .28 | .02 |
| RNA            |   | Mean change week 12-16 | -1.76 | -1.55 | -5.56 | .35 | .001 | .001 |
| RNA            |   | Mean change week 20-28 | -1.72 | -1.43 | -5.55 | .14 | .001 | .001 |
| RNA            |   | AUCMB week 16 | -1.61 | -1.44 | -9.99 | .24 | .002 | .001 |
| RNA            |   | AUCMB week 28 | -1.63 | -1.41 | -8.5 | .15 | .001 | .001 |

\*Z = ZDV; D = ddI; N = NVP.

NVP when used in combination with 3TC and other anti-retroviral therapies in NNRTI naïve patients with CD4 counts \( \leq 200 \) cells/mm\(^3\). The primary efficacy endpoint was time to clinical disease progression, subsequently changed to time to virologic failure as defined as increase in HIV-1 RNA above limit of quantitation (BLQ). The planned primary analysis, a stratified Fisher’s exact test on percentage of subjects without virologic failure at Week 48, stratified by prior anti-retroviral therapy, HIV disease status, baseline CD4 count, and baseline HIV-1 RNA found nevirapine to be superior to placebo with a \( p \)-value <0.001.

The INCAS trial was designed to compare one triple-drug regimen, indicated by NVP+ddI+ZDV, to two dual-drug regimens (ddI+ZDV, NVP+ZDV). The primary endpoint was percent BLQ by 48 weeks in HIV-1 infected anti-retroviral naïve patients with CD4 cell counts of 200–600 cells/mm\(^3\) without AIDS-defining illness or active invasive infection or malignancy. The primary analysis found adding nevirapine to ddI+ZDV background gave a significant increase in sustained viral suppression from 19% to 45% (log-rank \( p \)-value <0.001). It was also found that ddI+ZDV was statistically significantly superior to NVP+ZDV (log-rank \( p \)-value <0.001), indicating that nevirapine should not be used with only one NNTI.
The Atlantic trial was designed to compare efficacy of three different triple-drug regimens comparing NVP with indinavir (IDV) and NNTI 3TC when used in conjunction with ddI and stavudine (also known as d4T). The primary efficacy endpoint is percent BLQ at 48 weeks. This trial was conducted in asymptomatic NNTI naive patients with CD4 counts >200 cells/mm$^3$ and HIV-1 RNA ≥500 copies/ml. The primary analysis used 95% two-sided confidence intervals for the difference in success rates, using normal approximation to the binomial. The trial was felt to have too small sample size and the confidence intervals were too wide to support a firm conclusion that nevirapine is no less than 10% worse than indinavir or lamivudine (3TC).

Unlike the previous case study, this example highlights a regulatory decision that was less straightforward. However, it was flexible enough to keep the totality of the drug experience in order to meet the demand for newer treatment regimens faster.

### 3.4 Use of Surrogate Markers in Anti-cancer Drug Products

Traditionally, therapies for cancer patients have been approved on the basis of objective response to the agent (tumor shrinkage) together with direct evidence that the therapy produces measurable clinical benefit. Typical approval endpoints have been included, such as response rate together with increased patient survival, decreased recurrence rate, increased disease-free interval, and/or improved quality of life. It has been assumed that durable, complete clinical response (complete disappearance of detectable tumor) is a valid surrogate for such clinical benefit, but it is only infrequently achieved. Much more commonly, partial tumor shrinkages are induced, and evidence has accumulated that such responses are often directly linked to longer or better patient survival. In fact, for some new agents, the FDA began to rely on a reasonable high rate of verifiable objective partial response to the therapy as a basis for approval of agents to treat refractory malignancies without requiring evidence of improved survival or quality of life even prior to 1996. Subsequently, additional trials have been conducted to confirm or expand the product’s indication. Although the predictive value of partial responses may still be a matter of discussion and study for all types of cancer patients, the FDA had concluded that for patients with refractory malignant diseases or for those who have no adequate alternative, clear evidence of anti-tumor activity is a reasonable basis for approving the drug. In these cases, studies confirming a clinical benefit may be appropriately completed after approval.
In March 1996, U.S. President Bill Clinton and Vice President Al Gore issued a National Performance Review as a part of reinventing the government initiative. This document discussed accelerating approval as well as expanding access to anti-cancer agents. In the introduction, it stated “The Food and Drug Administration has demonstrated a longstanding commitment to the prompt consideration and, when appropriate, early approval of new therapies for cancer patients.” To speed up the entire process further, the FDA is adopting a uniform policy that will permit accelerated approval of a significant number of new cancer therapeutics. In the past, the FDA has approved cancer therapies on the basis of an agent’s ability to produce an effect on the well-established and long-recognized criteria such as survival, improved quality of life, and relief of symptoms, as well as objective disease regression. When partial response of disease (measurable but incomplete tumor shrinkage) has been noted in patients who have extensive or metastatic cancer, it is often correlated with other approval criteria. Because of this experience, it is believed that for many cancer therapies it is appropriate to utilize objective evidence of tumor shrinkage as a basis for approval, allowing additional evidence of increased survival and/or improved quality of life associated with that therapy to be demonstrated later. By utilizing objective response as a surrogate endpoint in clinical trials, the FDA will decrease the total time needed for marketing approval in many situations.

Although the accelerated approval provisions have been applicable to promising treatments for cancer patients who do not benefit from or cannot tolerate available therapy, this approval mechanism had not been frequently utilized prior to 1996, largely because general agreement on reasonable surrogate endpoints had been lacking.

Under the 1996 initiative, the FDA substantially expanded the use of accelerated approval process based upon verified and recognized demonstration of objective tumor shrinkage. For approval, potential effectiveness of the treatment should outweigh its toxicities and post-approval studies will usually be required to further define the utility of the new agent for the approved and/or other indications, either alone or in combination with other agents. The FDA can also apply accelerated approval provisions to certain products intended to remove a serious or life-threatening toxicity of cancer treatment based on post-approval studies that demonstrate that surrogate measures correspond to clinical benefit and/or effect of therapy on survival.

The greater utilization of the accelerated approval provisions for cancer treatment not only has an important impact on the original applications but also on supplemental application for secondary indications. The actual use of cancer agents may be far broader than the approved indications. Because of the nature of cancer therapy, the approved label does not nec-
essarily convey all the medical conditions for which the agent is used or may be useful. Nonetheless, the FDA-approved label should accurately convey as many as agent’s uses as are properly supported by data.

The greater utilization of the accelerated approval provisions for cancer treatment not only has an important impact on the original applications but also on supplemental application for secondary indications. The actual use of cancer agents may be far broader than the approved indications, and because of the nature of cancer therapy, the approved label does not necessarily convey all the medical conditions for which the agent is used and may be useful. Nonetheless, the FDA-approved label should accurately convey as many as agent’s uses as are properly supported by data.

The type and quantity of clinical data that is required will vary depending on the cancer indication under study, the availability and acceptability of other therapies, and the specific observations in the studies. According to the Guidance to the Industry document dated December 1998 (http://www.fda.gov/cder/guidance/index.htm), there is flexibility regarding the data requirements. In the refractory cancer setting, for example, where therapies with meaningful benefit are unavailable, non-randomized studies showing that a new treatment provides a significant objective response rate with tolerable treatment toxicity may be adequate to support approval under the accelerated approval regulations. In this setting, objective response rates are considered a surrogate endpoint reasonably likely to predict a clinical benefit. Evidence to confirm that clinical benefit can be obtained after approval. In those cases where durable complete responses can be attained, non-randomized studies showing a significant rate of durable complete responses can be persuasive evidence of effectiveness.

During 1992–2002, 15 NDAs involving 13 drugs have been submitted to the Division of Oncologic Drug Products in Center for Drug Evaluation and Research (CDER). Of them, 10 were based on single-arm phase II studies and used objective response as a surrogate endpoint. Only 5 were based on randomized trials. The details are given in Tables 3.13 and 3.14.

This brings up several important trial design issues about optimal accrual of patients in the trials and the extent to which the changing circumstances can impede the conduct of planned studies. Consider the following case study where this scenario has been brought to bear.

3.4.1 Doxil: A Case Study

Doxil (doxorubicin HCl liposome injection) was approved under the accelerated approval mechanism for “the treatment of metastatic carcinoma
TABLE 3.13. Single-arm trials with no concurrent comparator in the Division of Oncologic Drug Products.

| Drug                   | Year | Indication                          | Sample size | Trial details                  |
|------------------------|------|-------------------------------------|-------------|--------------------------------|
| Liposomal doxorubicin  | 1995 | Kaposi’s sarcoma second line         | 383         | 77 of 383 identified refractory |
| Amifostine (Ethyol)    | 1996 | To decrease cisplatin toxicity in NSCLC | 100         | Two trials, 50 patients each    |
| Docetaxel (Taxotere)   | 1996 | Breast cancer second line            | 483         | 6 US trials total 309; 3 Japanese trials 174 |
| Irinotecan (Capeptosar) | 1996 | Colon cancer                        | 132         | Single-arm trial                |
| Capecitabine (Xeloda)  | 1998 | Breast cancer refractory            | 162         | Single trial in patients in stage IV disease |
| Liposomal doxorubicin  | 1999 | Ovarian cancer refractory           | 145         | 3 studies                      |
| Temozolomide (Temodar) | 1999 | Anaplastic astrocytoma refractory   | 162         |                                |
| Gemtuzumab ozogomycin  | 2000 | AML                                 | 142         | 3 studies                      |
| Imatinib (Gleevec)     | 2001 | CML in BC, AC, or CP after interferon failure | 1027     | 3 studies                      |
| Imatinib (Gleevec)     | 2001 | GIST                                | 147         | Single study 2-arm              |

of the ovary in patients with disease that is refractory to both paclitaxel and platinum-based chemotherapy regimens. Refractory disease is defined as disease that has progressed while on treatment, or within 6 months of completing treatment.” In November 1998, the drug was assigned an orphan drug designation, given that no drug has been approved for the treatment of ovarian cancer refractory to platinum compounds and paclitaxel. In December 1998, a supplemental NDA was submitted for the above indication containing data from three Phase II non-comparative studies in relapse or refractory ovarian cancer. The primary analysis was based on the surrogate endpoint of response rate on 176 patients. This application also submitted data from an interim analysis of an ongoing Phase III study (Study 30–49) comparing Doxil with Topotecan. The accelerated approval was granted in June 1999. The traditional approval was to be based on the timely completion and final results of Study 30–49.

Study 30–49 (performed May 1997–March 1999) was designed to show safety and efficacy in patients with relapsed ovarian cancer following failure with platinum based chemotherapy in 474 patients. The study was stratified by platinum sensitivity and bulky disease and designed to show superiority of Doxil to Topotecan in either time to progression (TTP) or survival, with a supporting trend demonstrated for the other endpoint. The secondary
TABLE 3.14. Approvals based on randomized trials in the Division of Oncologic Drug Products.

| Drug               | Indication                             | Year | Endpoint                          |
|--------------------|----------------------------------------|------|-----------------------------------|
| Dexrazoxane (Zinecard) | Reduction of doxorubicin cardiomyopathy | 1995 | LVEF, Cardiac heart failure       |
| Liposomal cytarabine (Depocyte) | Lymphomatous meningitis                | 1999 | Cytologic response                |
| Celecoxib (Celebrex)   | Reduction of adenomatous polyps        | 1999 | Number of polyps                  |
| Oxaliplatin (Eloxatin) | Second-line colorectal cancer          | 2002 | Objective response, Time to progression |
| Anastrozole (Arimidex) | Adjuvant post-menopausal ER+           | 2002 | Disease-free survival             |

Outcomes were objective response rate (ORR), response duration, survival and safety. If Study 30–49 did not demonstrate the clinical benefit of Doxil, the sponsor would have to perform another study to show clinical benefit of the drug in ovarian cancer.

In June 2000, the sponsor informed the FDA that the planned treatment analysis for Study 30–49 did not demonstrate superiority in TTP, but showed significant survival advantage of Doxil over Topotecan in the platinum-sensitive group, with approximately 50% of the patients still alive. However, in the platinum-refractory subset, the patient population for which it is to be indicated, the results were marginally in favor of the control (hazard ratio = 1.01 [0.78, 1.31]) (Gordon 2003). This made regulatory decision significantly harder, and accelerated approval was granted after recommendation from a panel of external experts in an Oncology Advisory Committee (ODAC) meeting. Results presented at the ODAC are summarized in Table 3.15 (Hamburger 2003). Based on this scenario, it was agreed that a final survival analysis is to be performed on Study 30–49 when a 90% of the patients (planned size is 474) died or were lost to follow up. The final survival result of Study 30–49 is currently undergoing regulatory review.

A second protocol to prove clinical benefit was required. This Phase IV protocol (SO200), initiated in 2000 and currently enrolling, was an open-label inter-group study between Doxil and carboplatin versus carboplatin in 900 platinum-sensitive patients with recurrent epithelial ovarian carcinoma after failure of initial, platinum-based chemotherapy, to be performed jointly with an oncology group, SWOG. The primary endpoint is overall survival and the secondary endpoints are progression-free survival (PFS), confirmed complete response (CR), time to failure (TTF), and toxicity. It is currently enrolling patients.

This experience brings forth some of the challenges surrounding Phase IV
TABLE 3.15. Doxil. Median time to progression (TTP) and overall survival (OS) time in weeks at the end of planned treatment.

| Population          | Doxil (N) | Topotecan (N) | p-value |
|---------------------|-----------|---------------|---------|
|                     | TTP       |               |         |
| All patients        | 18.4 (239)| 18.3 (235)    | 0.632   |
| Platinum-sensitive  | 29.9 (109)| 26.7 (111)    | 0.387   |
| Platinum-refractory | 9.1 (130) | 14.3 (124)    | 0.941   |
|                     | OS        |               |         |
| All patients        | 58.7 (239)| 56.7 (235)    | 0.964   |
| Platinum-sensitive  | 110.7 (109)| 84.7 (111)  | 0.027   |
| Platinum-refractory | 34.6 (130)| 41.4 (124)    | 0.126   |

commitment trials in oncology, highlighted by the case study.

- The times to complete the Phase IV commitments are often longer than anticipated. In the Doxil case study, after the end of the planned treatment analysis, the primary endpoint was modified to become overall survival. Time to reach 90% event endpoint in Study 30–49 took more than 3.5 years.

- Multiple parties are often involved in finalization and implementation of the Phase IV trials. In the Doxil experience, the transfer and clinical responsibilities had to be coordinated between the sponsor, SWOG, other cooperative groups, National Cancer Institute (NCI), and with the FDA.

- The competition for accrual among other ongoing trials is often so fierce that it impedes the progress of the trial.

- After the accelerated approval, a drug can be prescribed to patients with that indication outside of a clinical study, making it harder to accrue patients needed for completion of Phase IV commitment.

3.5 Use of Surrogate Markers in Cardiovascular Drug Products

The use of surrogate markers in cardiovascular drug products has received mixed response – a rising enthusiasm for providing efficacious drugs at the
earliest possible time along with experiences tempered with some unexpected results in some products.

Surrogates can be early or late in the causal chain—cholesterol (a biochemical variable), blood pressure (a pathophysiologic variable), coronary vessel diameter (a morphological variable), or left ventricular hypertrophy (a morphological variable). However, some are closer to certain clinical events such as myocardial infarction and heart failure. Some surrogates are not etiologic but are thought to reflect activity of an underlying process that leads to an adverse event.

The risk of reliance on a surrogate is that the pathway connecting surrogate endpoint to the clinical endpoint may not be clear. It is widely accepted that elevated blood pressure is a direct cause of stroke, heart failure, and renal failure and accelerated coronary disease and that reducing blood pressure reduces morbidity and mortality. However, before the controlled outcome studies of hypertensive drugs were performed in 1960s, there was an active debate that blood pressure was an “adaptive” response to the vascular disease and that lowering it would be harmful (Freis 1990). Recent data on the benefits of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors may have partially settled the role of surrogate markers in cholesterol-lowering drugs, but the real value of other surrogate markers in cholesterol-lowering drugs may not be as clear. The safety database needed to characterize adequately the risk-benefit ratio is often not extensive enough in accelerated approval submissions, leading to the common phrase “there is no surrogate for safety” (Temple 1999).

Distinguishing concern about the validity of the surrogate from the more general question of safety is important because it affects the kind of data that can be used to assess the benefits and risks of treatment. If there is doubt about the surrogate itself, only an outcome study in the specific disease can determine the value of the drug. But if the validity of the surrogate is accepted, studies in a variety of settings may be pertinent to assessment of safety. For example, a drug lowering blood pressure may be about as certain to provide clinical benefit, as would be an anti-anginal drug. However, the safety profile of the drug has to be established, probably from a moderate study in hypertension or angina or from another population more vulnerable to cardiovascular toxicity or from pharmacologically related agents in either population. The absence of outcome studies in certain anti-hypertensives (calcium channel blockers and angiotensin-converting enzyme (ACE) inhibitors) has been cited by critics of the use of surrogates. Table 3.16 summarizes the surrogate endpoints used in cardiovascular drugs.
TABLE 3.16. Surrogate endpoints used in the Division of Cardiovascular Drug Products.

| Condition                          | Approval endpoint                                      | Postmarketing outcome studies |
|------------------------------------|--------------------------------------------------------|-------------------------------|
| Hypertension                       | Change in blood pressure                              | No                            |
| Hyperlipidemia: initial approval    | Change in blood lipid level                           | Yes\(^1\)                     |
| Hyperlipidemia: clinical benefit    | Survival, rate of myocardial infarction               | No\(^2\)                      |
| Hyperglycemia                      | Change in blood sugar levels, glycosylated hemoglobin  | No                            |
| Heart failure: symptoms            | Exercise, symptoms, together with evidence (except for ACE inhibitors) that there is no adverse effect on survival | No                            |
| Heart failure: long-term benefit   | Survival, hospitalization                             | No\(^2\)                      |
| Angina, effort                     | Exercise, symptoms                                    | No                            |
| Angina, vasospastic                | Angina rate                                            | No                            |
| Silent ischemia                    | Outcome (acute myocardial infarction, survival)       | No\(^2\)                      |
| Ventricular arrhythmia: symptoms   | Symptoms, with evidence of no harm                     | No\(^2\)                      |
| Ventricular arrhythmia: life-threatening | Symptoms, with evidence of no harm; survival           | No\(^2\)                      |
| Atrial arrhythmia                  | Symptoms, delayed recurrence, evidence of no adverse effect on survival | No\(^2\)                      |
| Acute coronary syndrome, postangioplasty/coronary artery bypass graft | Outcome (death, acute myocardial infarction, urgent intervention) | No\(^2\)                      |
| Acute myocardial infarction        | Survival                                               | No\(^2\)                      |
| (thrombolysis)                     |                                                        |                               |
| Orthostatic hypotension            | Decreased orthostatic blood pressure                   | Yes\(^3\)                     |

\(^1\) By Agreement, the sponsors voluntarily agreed to conduct post-marketing studies.

\(^2\) Studied pre-marketing.

\(^3\) Required under the FDA Accelerated Approval Rule.

### 3.5.1 Anti-hypertensive Drugs

Effect of blood pressure is the basis for approval of new hypertensive drugs. The most persuasive support for the surrogate endpoint of blood pressure is experience from numerous long-term outcome studies showing a clear effect on stroke and at least favorable trends on cardiovascular events and survival rates. In addition, substantial epidemiological evidence indicates that blood pressure is continuously related to the risk of stroke and coronary heart disease. Few active drugs have shown any other factor to modulate directly hypertensive benefit, but direct comparisons of high-dose diuretics and beta-blockers showed no real difference (Collins et al. 1990), even for cardiovascular events for which they are expected to be superior due to
post-infarction benefit and lack of hypokalemic effects. Although the FDA emphasized the importance of such comparisons in the past, comparative studies have not been required of individual sponsors. If ongoing large trials demonstrate differences in outcome with drugs approved using same surrogate procedures, that policy will change (Temple 1999).

3.5.2 Anti-platelet Drugs

Currently, platelet aggregation inhibitors or anticoagulants in various settings (post-infarction or stroke, peripheral vascular disease, acute coronary syndrome, post-angioplasty or post-bypass) are studied using clinical endpoints (death, new infarction, and urgent procedural intervention). As yet, although various anti-platelet treatments have a long and growing record of success in preventing adverse outcomes, there is no effect on a platelet aggregation or coagulation surrogate endpoint that has been convincingly shown to correspond to a clinical benefit and to define the risk of bleeding.

3.5.3 Drugs for Heart Failure

Increased mortality with two classes of inotropic agents and an inotropic vasodilator drug clearly indicates that hemodynamic or symptomatic benefit in heart failure does not predict improved survival. Therefore, for a drug to be approved for heart failure symptom improvement, evidence of a symptomatic benefit needs to be supported by showing that there is no adverse mortality effect. Long before the adverse outcome effects studies of inotropes were observed (Packer et al. 1993), the FDA concluded that “there should be reasonable assurance that survival in high-risk patients is not impaired; the controlled trials thus need to be of sufficient size to detect a substantial increase in mortality” (Temple 1987). This conclusion was based in part on early suggestions of rapid deterioration in open studies of inotropes and in part on the known adverse effects of digoxin.

3.5.4 Drugs for Angina and Silent Ischemia

Anti-anginal drugs are approved based on improvement in exercise tolerance or reduction in symptoms of angina; no current treatments have been shown to improve outcome. Safety of anti-anginal drugs is well supported by studies of calcium channel blockers and beta-blockers in post-infarction settings. Silent ischemia, like symptomatic ischemia, predicts an increased rate of death and myocardial infarction, and it has been proposed that a
reduced rate of silent episodes should be a basis for approval. As of now, the FDA has not accepted this suggestion (Temple 1990) concluding instead that the drugs for this indication need to show an effect on a clinical end-point, such as survival or rate of new infarction. It did not seem reasonable that the drugs known only to affect ischemia would provide benefit, when the same drugs used to treat symptomatic angina has not been able to show improved outcome. It also seemed at least possible that ischemia stimulated growth of collateral vessel, which could improve outcome (Temple 1988).

3.5.5 Ventricular Arrhythmias

The most controversial example of an erroneous surrogate is the stunning results of the Cardiac Arrhythmia Suppression Trial (CAST). Details of this trial will be discussed in Section 3.5.6. It definitely established that effective suppression of ventricular premature beats (VPB) does not decrease mortality, despite the well-established association between elevated VPB rates and early arrhythmic death. But although the markedly adverse outcome was certainly unexpected, labeling for encainide and flecainide before the CAST study specifically pointed out the absence of known survival benefit from VPB suppression, the lack of information on safety and effectiveness in the post-infarction state, and the drug’s ability to cause worsened arrhythmias. The indicated uses for both drugs were limited to patients with documented life-threatening arrhythmias and symptomatic patients with non-sustained ventricular and frequent VPBs. Since the CAST results were reported, approval of drugs for ventricular arrhythmias that are not immediately life-threatening has required showing improved survival benefits and no adverse effect on survival in case of symptomatic claim. At the present time, no drugs have been able to meet this standard.

3.5.6 The CAST Experience: A Case Study in Ventricular Arrhythmia

The occurrence of ventricular premature depolarizations in survivors of myocardial infarction is a risk factor for subsequent sudden death, but whether anti-arrhythmic therapy reduces risk is not clear. CAST was undertaken to evaluate the effect of anti-arrhythmic therapy, such as encainide, flecainide or moricizine, in patients with asymptomatic or mildly symptomatic ventricular arrhythmia after myocardial infarctions.

The purpose of the study was to test the hypothesis that suppression of ventricular ectopy after a myocardial infarction reduces the incidence of sudden death. The design of the study was multicenter, randomized, placebo con-
controlled with a preliminary, open-label titration to ensure that all patients would respond to at least one of the drugs. The primary endpoint was death or cardiac arrest with resuscitation, either of which due to arrhythmia.

The results of this study were unexpected. The flecainide and encainide arms of this trial were stopped early, after a mean follow up of 10 months. Of 89 deaths or cardiac arrests total, 63 patients were on active drugs versus 26 on placebo ($p = 0.0001$) and of death or cardiac arrest due to arrhythmia, 43 patients were on active drugs versus 16 on placebo ($p = 0.0004$). The conclusions were that the results indicate that encainide or flecainide, when used to prevent ventricular arrhythmias post-myocardial infarction, are detrimental to survival. The study continued limited to the arms of moricizine versus placebo.

After the flecainide and encainide arms of the CAST I were discontinued, a continuation of the CAST I, the CAST II used moricizine to determine if suppressing asymptomatic or mildly symptomatic ventricular ectopy post-myocardial infarction (MI) reduces the incidence of sudden death from ventricular arrhythmias. There were 1325 patients with EF greater than or equal to 40%, who were within 4 to 90 days of having an MI, and who had greater than or equal to 6 repetitive ventricular complexes. The CAST II was a multicenter, randomized, placebo-controlled study in which patients received placebo or up to 900 mg/day of moricizine as necessary to suppress arrhythmias. The primary endpoint was sudden death. There was a 14-day exposure phase and a 2-year long-term evaluation phase.

This study was terminated early because in the 14-day exposure period, there was excess mortality in the moricizine arm (17 deaths in 665 patients) as opposed to the no therapy or placebo group (3 deaths in 660 patients). A less than 8% chance of finding a survival benefit was found if the study was completed. It was concluded that the use of moricizine to suppress asymptomatic or mildly symptomatic ventricular premature depolarizations post-MI is ineffective and increases mortality.

These experiences point to an interesting fact that surrogate markers may not always be useful and have to be validated extensively before being used as a regulatory tool.

### 3.6 Statistical Issues Related to Accelerated Approval

The accelerated approval process is an important tool for therapeutic agents in serious or life-threatening diseases where no existing therapies are ex-
pected to be of help. For example, in oncology, survival and improvement in patient-reported symptoms are considered clinical endpoints while objective response rate and time to progression are considered meaningful surrogate markers. There are several unique situations for drugs submitted under accelerated approval process. It is to be noted that none of them have been used in a regulatory setting so far, so implications of these methods on the drug development procedure are not ascertained.

- Because placebo is considered unethical in most clinical settings appropriate for accelerated approval, the trials are usually conducted in single-arm Phase II studies, sometimes not even in randomized trials.
- The accelerated approval is based on a very small sample size, making regulatory decisions about benefit-risk ratio very hard to characterize.
- The safety databases for such approvals are minuscule, posing serious consideration about the safe use of the drug in a widespread clinical setting.
- After accelerated approval, it is often not possible to perform the Phase IV commitments in the original population, and clinical experiences bridging the original population and the enriched population can be hard.

Some specific statistical issues have arisen in the accelerated approval process. Because the accelerated approval is based on a limited database of patients in which the drug has shown a positive finding, there are two stages in which the regulatory decision has to be taken and hence two points where a Type I error ($\alpha$) has to be controlled. There has been some discussion (Sridhara et al. 2001) that instead of considering the two approval processes as independent events, the more appropriate paradigm would be to distribute the $\alpha$ over both events.

If we consider these series of events in clinical trial terms, the accelerated approval can be treated as a decision based on an interim analysis on a pre-specified surrogate endpoint. The full or traditional approval can then be granted based on the clinical endpoint at the conclusion of Phase IV commitments, which can be treated to be the final analysis. Under this situation, there are several possible scenarios after accelerated approval is granted based on a surrogate endpoint:

1. The drug development continues as planned and full approval is granted based on the Phase IV commitments.

2. Based on the Phase IV commitments, the new drug does not demonstrate significant effect with respect to the desired final clinical outcome. This can happen if:
• the study conducted for the accelerated approval was a false positive study;
• the surrogate marker used in the accelerated approval study is not predictive of the clinical endpoint;
• the assumptions and/or the design of the accelerated and final approval studies were not appropriate;
• the final clinical risk/benefit ratio is markedly different than what was originally anticipated, e.g., unexpected negative mortality effect.

Under any of the scenarios discussed, the studies have to be designed with sufficient power to detect overall significant difference with respect to the clinical endpoint at the end of the studies. The studies also have to have sufficient assay sensitivity to detect significant difference with respect to the pre-specified surrogate endpoint at the interim analysis stage for the accelerated endpoint.

There have been some methodologies proposed recently for the two-stage design (see Figure 3.4). Two of them will be discussed briefly in the discussion, primarily from a theoretical standpoint.

The first method by Shih et al. (2003) proposes a two-stage design with clinical endpoint $T$ and surrogate endpoint $S$. At the end of the first stage, both $S$ and $T$ are evaluated according to the flowchart in Figure 3.5. At the end of the first stage, the data may support early termination of the trial for clinical benefit based on $T$ or may support accelerated approval based
FIGURE 3.5. Two-stage approval method proposed by Shi et al. (2003).

on the surrogate $S$. If neither is true, then the trial continues to the second stage. At the end of the second stage, $T$ is evaluated for clinical benefit. This is summarized in Figure 3.5. The “final approval” Type I error rate ($\alpha_F$) for the clinical endpoint is given by

$$\alpha_F = \alpha_{F1} + \alpha_{F2},$$

(3.5)

where

$$\alpha_{F1} = P(|Z_{T1}| > c_{T1} \mid H_{0T}),$$

$$\alpha_{F2} = P(|Z_{T1}| < c_{T1}, |Z_{T2}| > c_{T2} \mid H_{0T}).$$

$H_{0T}$ is the null hypothesis for $T$, and $Z_{T1}$ and $Z_{T2}$ are normally distributed test-statistics based on the data available for $T$ at the first and second stages, respectively. If $\alpha_{F1} = 0.001$, then $c_{T1} = 3.29$. If the correlation $\rho(Z_{T1}, Z_{T2}) = \sqrt{0.5}$, then to maintain $\alpha_F = 0.05$ or 0.04, $c_{T2}$ needs to be equal to 1.962 or 2.06, respectively.

The authors raise a question as to what false positive rate, or Type I error, for the surrogate endpoint should we control for. The “accelerated approval” Type I error rate ($\alpha_A$) is given by

$$\alpha_A = P(|Z_{T1}| < c_{T1} \text{ and } |Z_S| > c_S \mid H_{0S}, H_{0T}),$$

(3.6)

where $H_{0S}$ is the null hypothesis for $S$ and $Z_S$ is a normally distributed test-statistic based on the data available for $S$ at the first stage. Some of the choices can be:

- control $\alpha_F$ at the 0.05 level;
• control $\alpha_F$ at the 0.05 level and $\alpha_A$ at the 0.01 level;

• control $\alpha_F + \alpha_A$ at the 0.05 level;

• any other appropriate choice.

The second proposed method, due to Yang et al. (2002) and Shridhara et al. (2002), analogously considers a two-stage design with clinical endpoint $T$ and surrogate endpoint $S$. At the end of the first stage, both $S$ and $T$ are evaluated according to the flowchart in Figure 3.6. However, unlike Shih’s method, there is an additional condition to be satisfied before concluding that the interim data supports accelerated approval. The Type I error rate for the overall clinical endpoint is given again by (3.5) with

$$
\alpha_{F1} = P(Z_{T1} > c_{T1}|H_0T),
\alpha_{F2} = P(Z_{T1} < c_{T1}, Z_{T2} > c_{T2}|H_0T).
$$

But, in addition, it is proposed that the following probability,

$$
P(c_{T1, \alpha^*} < Z_{T1} < c_{T1} \text{ and } Z_S > c_S \text{ and } Z_{T2} < c_{T2} | H_0T), \quad (3.7)
$$

should be controlled at some level $\gamma$. The last equation is the joint probability of a positive surrogate outcome and a nominally positive ($\alpha^*$) clinical benefit at interim, and a non-significant final clinical benefit outcome. If this probability is less than a certain level, say, $\gamma = 0.30$, then the interim positive results on the surrogate endpoint and a nominally positive ($\alpha^*$) clinical outcome provide a reasonable level of evidence to support the
TABLE 3.17. Probability of a study to support accelerated approval under an alternative of 80% power to detect a clinical benefit and 90% power for surrogate endpoints. The probabilities are calculated based on the O’Brien-Fleming version of the Lan-DeMets spending function.

| Information fraction $(t)$ | Corr$(Z_S, Z_T)$ $(\rho)$ | Level of $\gamma$ $0.10$ | $0.20$ | $0.30$ |
|---------------------------|--------------------------|----------------|-------|-------|
| $\frac{1}{4}$            | $0.1$                    | $0.4627$       | $0.5885$ | $0.6618$ |
|                           | $0.5$                    | $0.4850$       | $0.6082$ | $0.6778$ |
|                           | $0.9$                    | $0.5087$       | $0.6307$ | $0.6967$ |
| $\frac{1}{3}$            | $0.1$                    | $0.5190$       | $0.6293$ | $0.6887$ |
|                           | $0.5$                    | $0.5428$       | $0.6478$ | $0.7023$ |
|                           | $0.9$                    | $0.5696$       | $0.6713$ | $0.7205$ |

drug’s accelerated approval (in the absence of other supportive information). With $c_{T1}$ and $c_{T2}$ specified by some spending function, and if we also specify $\gamma$, then we can solve for $\alpha^*$, and thus, $c_{T1,\alpha^*}$. An example is shown in Table 3.17.

The criterion provides some level of assurance of a clinical benefit in the event that the confirmatory trial may not materialize due to patient crossover, changing standards of care, other available new treatments, etc. The method does require that at the end of the first stage, there should be at least a certain fraction of the expected total events to have occurred.

As a consequence, the submission of evidence based on the surrogate endpoint for an accelerated approval will be slightly delayed until the desired fraction of the expected total events has been achieved. This represents a compromise between the real possibility that the post-marketing trials already underway, or yet to be conducted, to confirm the positive finding on the surrogate endpoint may never materialize.

### 3.7 Surrogate Markers at Other Phases of Drug Development

At earlier stages of drug development, both positive (efficacy) and negative (safety) effects of a drug can be characterized using a variety of measurements or response endpoints. These effects include clearly clinically pertinent effects (clinical benefit or toxicity), effects on a well-established surrogate (such as blood pressure or QT interval in cardiovascular dis-
ease), or effects on a more remote biomarker (such as ACE inhibition or bradykinin levels in cardiovascular disease). All of these measurements can be expected to show exposure-response relationships that can guide therapy, suggest dose/dose intervals, or suggest further study. In many cases, multiple response endpoints are more informative than single endpoints for establishing exposure-response relationships. Methods to combine attributable proportions or relative effects of two or more surrogate markers for the same true clinical endpoint have to be developed. Specifically, less clinically persuasive endpoints (biomarkers, surrogates) can help in choosing doses for the larger and more difficult clinical endpoint trials and can suggest areas of special concern.

In addition, surrogate endpoints can be used to link with external sources of information on the disease or on other treatments. They can be used to integrate the data across all phases to build an evidence base, including validation. This database can be analyzed and mined for the relationship of surrogate endpoints to the disease states, other markers, and patient covariates.

Surrogate marker methodology can play a significant role in drug safety and risk assessment area. The databases can be mined for the signs of potential toxicities. For example, certain liver enzyme tests (ALT, AST, bilirubin) can be considered as surrogate markers for hepatic toxicity. This will enable early detection of potential problems later in the drug development process.

Finally, surrogate markers play an important role in drug development in identifying faster and more focused pathways to bringing a promising drug to patients. However, care needs to be taken in ensuring that the markers are pre-specified, equivocally validated, and predictive of the final clinical benefit to the patients.