Reforming FDA Policy

Lessons from the AIDS Experience

Joanna E. Siegel and Marc J. Roberts

Recent changes in drug regulatory policy introduced as a result of the AIDS epidemic have been widely interpreted by both detractors and supporters as portending a major change in the drug approval system. The policies proposed or implemented so far, however, do not constitute basic reforms. They target AIDS and other life-threatening illnesses, rather than address the standards for drug approval generally. For the most part, recent reforms formalize policies and procedures used by the Food and Drug Administration (FDA) for many years to allow access to unproven drugs in exceptional circumstances.

We shall argue that although AIDS activism to date has not motivated fundamental changes in the drug approval system, such changes are in fact warranted. The AIDS activists' critique—that a long and inflexible drug approval process actually endangers sick patients—is not exceptional, but is a particularly dramatic example of the problems of an excessively restrictive drug approval process. A better process would more adequately consider the consequences of delays in accessing new drugs as well as the risks associated with their early and widespread availability. We outline our interpretation of the problems related to the current safety and efficacy standard and propose changes in its implementation. We believe that regulation must become more flexible, differentiated, and sophisticated than our current nominal approach.

Joanna E. Siegel is an assistant professor and Marc J. Roberts is a professor of political economy and health policy in the Department of Policy and Health Management at the Harvard School of Public Health.

Traditional Regulation of the Pharmaceuticals Market

The recent history of pharmaceuticals regulation began with the Food, Drug, and Cosmetic Act of 1938, which required that drugs be proven safe before they could be marketed. This law was amended in 1962 (the Kefauver-Harris amendments) in the wake of the thalidomide scandal and following extensive hearings on deceptive drug marketing practices. The amendments added the requirement that drug companies provide evidence of efficacy and simultaneously broadened FDA discretion in approving new drugs.

The post-1962 drug research process requires seven to ten years to complete, including an average of two years of preclinical testing (laboratory and animal studies) followed by several years of human trials. Human testing is divided into three stages. Phase I studies, usually conducted with twenty to eighty healthy volunteers, establish the drug's pharmacological action and safe dosage levels. Phase II studies, involving 100 to 3,000 patients, are pilot controlled studies to assess drug effectiveness and to identify side effects. Phase III studies are extensive clinical trials, conducted on 1,000 to 3,000 patients, to confirm efficacy and to detect infrequent adverse effects. Following Phase III, the drug company submits the new drug application to the FDA for review, typically a two- to three-year process.

Extensive regulation in the pharmaceuticals market is intended to protect consumers. A special need for protection in this market is often explained by the complexity of drug products and the particular vulnerability of buyers. The sick are seen as

uniquely susceptible to exploitation by ruthless sellers because of emotional vulnerability or diminished capacity. In *United States* v. *Rutherford* (1979) the Supreme Court noted: "Since the turn of the century, resourceful entrepreneurs have advertised a wide variety of purportedly simple and painless cures for cancer, including liniments of turpentine, mustard, oil, eggs, and ammonia; peatmoss; arrangements of colored floodlamps; pastes made from glycerin and limburger cheese .... Congress could reasonably have determined to protect the terminally ill, no less than other patients, from the vast range of self-styled panaceas that inventive minds can devise." In addition, buyers themselves (doctors or patients) have too few observations of a drug's effects to judge whether a claim of safety or efficacy is true. Because experience will generally be a poor guide, consumers must have other means of determining the desirability of a product. Finally, the effects of "mistakes" in using drugs may be large and irreversible.

Economists argue for regulation on the basis of the public good character of drug research. In an unregulated environment knowledge of new drugs would be underprovided if introducing or selling products were possible without it. Advertising and other marketing strategies would often be cheaper methods of capturing market share, although such alternatives might not improve consumers' choices.

Recent changes in drug regulatory policy introduced as a result of the AIDS epidemic do not constitute basic reforms; they mainly formalize policies and procedures used by the FDA to allow access to unproved drugs. A better process would more adequately consider the consequences of delays in accessing new drugs and the risks associated with their early and widespread availability.

Knowledge of new drugs would also be underdiffused in an unregulated market. Since the cost for the marginal use of such knowledge is zero, its price should be zero. Charging for drug information would lead to its underconsumption, but not charging would lead to its underproduction.

Both protectionist and public goods arguments support intervention in the drug market, but not necessarily the exclusion of unapproved drugs. Economists generally consider policies that limit consumers' options to make them worse off. A policy of limiting product variety can be justified in the pharmaceuticals market, however, because of the search and information costs in an unrestricted market. To illustrate, assume that there are products that are difficult to distinguish, that decrease well-being if consumed mistakenly, and that few customers would willingly choose over available alternatives. If these products are allowed on the market, buyers must invest time and resources to avoid costly mistakes. If they are banned, few are injured, and most consumers gain in reduced search costs.

In practice few disagree that regulation of the drug market is desirable. The argument focuses instead on whether the current regulatory policies—the limits to product variety—impose costs that are excessive. The problems associated with the strict requirements for drug approval became evident soon after the Kefauver reforms were enacted. Numerous studies described increased research costs in the pharmaceuticals industry, reduced innovation, increased time to drug approval, and a drug lag—a delay before new drugs available in Europe became available in the United States. These costs of a strict approval process led Milton Friedman to protest that the existing rules are "doing vastly more harm than good." Less strident critics assert that the restrictions are simply doing more harm than necessary. While regulation of the pharmaceuticals market is preferable to free-market distribution of drugs, the challenge is to reform the system of regulation in such a way as to lower its costs.

The AIDS Reforms. The AIDS epidemic brought a shift in the actors and alliances concerned with drug regulation. AIDS activists joined the forces voicing frustration over the FDA's restrictive stance. As consumers—the constituency the FDA was supposed to protect—AIDS activists were particularly difficult for the agency to ignore. Well-organized, highly visible, and knowledgeable about the drug approval process, they made effective use of the media to communicate their concerns. Extensive congressional hearings explored charges of a lack of responsiveness on the part of the FDA and the appropriateness of its restrictions on access to drugs.

The FDA responded to this pressure through several measures designed to expedite access to AIDS treatments: treatment use of investigational new drugs (IND), liberalized interpretation of import regulations, subpart E regulations to speed the approval process, and the parallel track proposal

to allow drug access early in the testing process.

FDA regulations for the treatment use of investigational new drugs, formalized in 1987, are intended to expedite access to promising new drugs before approval, usually during Phase III testing. Once a drug is granted treatment IND status, physicians may obtain it directly from the pharmaceutical company sponsor. To quality for treatment IND status, a drug must treat a serious or immediately life-threatening illness for which there is no satisfactory alternative treatment. The drug must be undergoing clinical trials, and its sponsor must be actively pursuing approval. Available evidence must indicate that the drug is safe and that it may be effective.

The treatment IND procedures formalize access for all patients and physicians to experimental drugs formerly available only in individual cases or in isolated circumstances. These procedures allow manufacturers to recover part of the costs of research, production, and distribution earlier in the process so that broader distribution is practical. The restricted scope of the treatment IND measure is also notable, however, particularly the continuing requirement for efficacy. Because of the efficacy requirement, treatment IND status has been used mainly for the distribution of drugs already in late stages of clinical testing or as a bridge between testing and approval.

Although the FDA has the authority to exclude unproven drugs from the United States, a policy adopted in 1988 allows individuals to import drugs for their own use. The individual must identify a supervising physician and may import only a short-term (three-month) supply. Before this regulation, AIDS patients had illegally imported drugs approved in other countries through "buying clubs." The new policy formalized a tacit FDA practice of noninterference with this activity and required only that the clubs send drugs directly to individual users to avoid the prohibition against the import of commercial quantities.

The FDA amended the investigational new drugs regulations with Subpart E provisions in 1988. Subpart E provisions encourage sponsors of drugs for life-threatening and severely debilitating diseases to meet with FDA officials early in the testing process to design the later stages of research. Although the requirements for approval are not changed, these provisions are intended to speed the approval process by condensing the later stages of testing to obtain the necessary evidence of efficacy.

The parallel track procedures, proposed in 1990, were advanced in response to the FDA's conservative interpretation of the treatment IND requirement for effectiveness. The parallel track would allow the many patients ineligible for or without access to clinical trials to obtain experimental drugs directly from pharmaceutical companies. The qualifying drugs would thus become available to patients concurrently with testing.

The major difference between treatment IND status and the proposed parallel track would be the stage at which drugs would be eligible. Most treatment IND statuses have been granted to drugs during or after Phase III testing, although drugs may be considered as early as Phase II. Expanded availability through the parallel track could be initiated following Phase I trials, the small-scale trials to establish safe dosage. Only a "promise" of efficacy would be expected on the basis of animal or initial human studies.

To expedite access to AIDS treatments the FDA authorized the treatment use of investigational new drugs, liberalized its interpretation of drug import regulations, issued regulations to speed the approval process, and proposed allowing drug access early in the testing process.

The parallel track is proposed only for AIDS or HIV-related illness, only for cases with no available satisfactory alternative, and only for patients unable to participate in clinical trials. A prototype of the parallel track was carried out with the drug DDI, which was subsequently approved.

The recent changes, although clearly needed, do not challenge the fundamental assumptions of past FDA policies. They represent departures targeted only to exceptional circumstances. Anthony Fauci of the FDA has explained: "[We are trying] ultimately to get drugs proven to be safe and effective or unsafe and ineffective .... On the way to that, we have to take into consideration the needs of people who don't have any other options." But the safety and efficacy standards impose costs on a broader range of consumers than those with terminal illness.

Problems with the Safety and Efficacy Standards. The notions of safety and efficacy have guided our drug regulatory policy for thirty years. This paradigm suggests that a drug is first found to be safe

or unsafe. If it is safe, its efficacy is then tested, and it is approved for use if found effective. Actually, for several reasons, the dilemmas posed by drug regulation are much more complex.

First, the concepts of safety and efficacy have meaning only in the context of an ill individual. A drug that poses unacceptable risks for a patient with a minor illness might provide an important improvement in length or quality of life for a person with a debilitating illness. Similarly, many drugs that improve outcome but do not cure an illness are "effective" relative to the patient's initial prognosis.

Second, drug approval decisions considering safety and efficacy must depend on the available alternatives. If a new drug poses greater risks to a patient than an existing drug accomplishing a similar therapeutic outcome, it is not "safe." If no other treatment exists, it may well be "safe" as compared with the risks associated with the illness.

Pharmaceutical regulation should be based on a more complex test than safety and efficacy. The appropriate test for drug approval should be whether a drug represents a reasonable option in specified patient circumstances.

Third, many effects of new drugs are unknown. An important aspect of the FDA's dilemma is that regulators typically do not know exactly how a drug will affect patients. The agency must decide whether to allow various uses of the drug while this uncertainty is being diminished by further study.

Finally, the complexity of assessing drug effects is a function of the variability and uncertainties characterizing potential consumers. For example, how long will a given AIDS patient live without AZT? Such information is described by one probability distribution while the action of the drug is described by a second probability distribution. The individual's situation is multidimensional. Beyond surviving or dying, an individual will experience many consequences from illness and from drugs taken to combat illness that range from hair loss and pain to life-threatening side effects. Individuals will also place different values on various outcomes. Some care more about avoiding pain, others about a longer life.

While the language of current drug regulation implies that approved drugs should be safe and effective, almost every drug raises the possibility of an outcome that potential users would not welcome. The standard reflects an ideal rather than the reality of regulatory choices and therefore provides little guidance in actual decisions. Because of this lack of fit, risk avoidance and the historical biases of the agency have disproportionately influenced regulatory decisions.

Pharmaceutical regulation should be based on a more complex test than safety and efficacy. An appropriate test should reflect the costs of limiting patients' access to drugs as well as the costs of allowing access. It should provide a consistent principle for use in regulatory decisions that conforms more closely with the complexities of these decisions and the tradeoffs required. We would argue that the appropriate test for drug approval should be whether a drug represents a reasonable option in specified patient circumstances.

The Reasonable Option Standard

On an operational level a reasonable option principle would move the FDA further in the direction of perfecting consumer choice by eliminating those options that are clearly undesirable for most patients or that have satisfactory alternatives for most consumers. Standards of proof should be stringent where risks of disease are low, where treatments are currently available, or where the possible side effects are more severe. In these situations, consumers can afford to wait to ensure that a new compound is in fact a reasonable option. The strict requirement would seek to avoid situations such as those that occurred in the early marketing of the antibiotic chloraphenicol. A number of patients who could have been effectively treated with other antibiotics died from the side effects of this drug.

If a disease has high costs—if it is life-threatening, severely disabling, or painful—a drug may be a reasonable option even if toxic or of uncertain safety or efficacy. A patient in these circumstances will often prefer a "promise" of efficacy to his prognosis without intervention. In these cases delay can impose extremely high costs on patients, and disincentives to drug development occur in areas where innovation is badly needed. The relative cost of the current efficacy requirement in this situation is excessive. Our alternative would imply a lighter burden of proof.

We thus propose that the FDA reconsider both the nature of its standards and the amount of evidence it should require to show that its chosen

standard has been met. In Samuel Broder's words, "[t]he more risk that a patient faces from the natural consequences of the disease, the more one needs to be inclined to act."

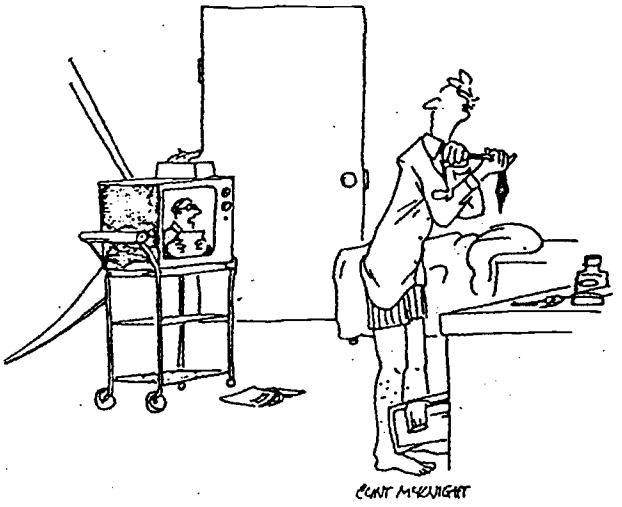
Facilitating Implementation of the Reasonable Option Standard. Broder's comment highlights the FDA's recognition of risk-benefit tradeoffs, despite the absence of a consistent application of this recognition in FDA decisions. An additional obstacle to implementation has been the lack of explicit intermediate approval mechanisms.

The existing approval process interferes with access to drugs because it has been too inflexible to recognize a range of intermediate situations where approval decisions should reflect risk tradeoffs. Currently, regulators have two basic options: approving or denying approval. (The latter is frequently equivalent to a "gather more evidence" option.) A more useful approval process would improve the targeting of drugs that are a reasonable option for a small number of patients and mitigate the risk and finality of approving drugs when important uncertainties remain.

Two types of changes are needed. The first is for drugs that are quite clearly both effective and toxic. Approval in these cases should be designed to facilitate appropriate use of the drug and decrease the FDA's risk in releasing it. Options should include conditions for prescription and provisions for review. In principle, these drugs should be approved if they confer substantial benefits otherwise unavailable to patients. The amount of acceptable risk would still be determined relative to the severity of the disease, but for debilitating illnesses the balance of incentives would shift in favor of releasing the drug.

The existing approval process interferes with access to drugs because it has been too inflexible to recognize a range of intermediate situations where approval decisions should reflect risk tradeoffs.

The second situation occurs where significant uncertainty remains as to the drug's effectiveness, but early results indicate that it could be the best available option. In this case the FDA should approve the drug with extensive provisions for postapproval research and review. These options do not apply



"The Food and Drug Administration today banned all forms of physical activity when five laboratory rats became, quote, 'super tired' after running on their wheel for a few hours."

only to terminal illness, but incorporate the relative risks of drug and disease as well as the available alternatives at any level of disease severity.

The mechanisms we suggest are not new, and they do not require modification of the current statute. The question is one of interpreting the safety and efficacy requirements—the level of proof and certainty the FDA actually demands.

To make drug approvals conform to the principle of a "sliding scale," the FDA must develop and publicize policies describing its interpretive practices and the intermediate options. The process would be controversial. Organized medicine might well protest the increased control over the practice of medicine implied by approval's restricting usage, and enforcement systems (for example, monitoring or liability) would generate debate. Carefully specifying new policies would, however, provide an explicit guide within the FDA that clarified objectives and improved consistency. Equally important, this specification would provide a basis for review by external observers such as the permanent oversight committee recently proposed by the National Committee to Review Current Procedures for Approval of New Drugs for Cancer and AIDS.

Implications of Changes in Drug Approval. Perhaps the most worrisome aspect of changes in the regulatory requirements relates to potential effects on drug research. It is argued that proposed changes would leave drug companies less interested in conducting clinical trials and patients less willing to participate in them. The proposed parallel track

has already generated considerable controversy as the first measure that would permit broad access to essentially experimental drugs, albeit only a few of them. Relaxation of the burden of proof could be expected to raise similar concerns.

The current requirement to prove efficacy before a drug can be marketed provides a strong financial incentive for drug companies' conducting clinical trials. With lower standards of proof some incentives would remain. The prospect of expedited marketing should encourage development of methods for quickly establishing initial evidence of effectiveness even if substantial proof can be ascertained only in longer-term studies. Other incentives could also be structured to encourage research, including liability burdens and levels of third-party reimbursement. The ultimate success of a drug should depend on evidence from trials to assure the appropriate investment of resources.

Reducing the evidence required to meet the FDA standard could undermine the incentives to conduct drug evaluation research. This danger can be offset to a significant extent by requiring postmarketing research and imposing other conditions on approval.

The other dilemma is whether patients would participate when drug companies wish to hold trials. Currently, patients regard trials as the primary source of treatment with experimental drugs. The practice of excluding more vulnerable population subgroups (such as children) has recently been criticized for this reason, despite the risk that originally justified such exclusion. If drugs are available before trials are completed or even begun, obtaining the drug would no longer be an incentive to participate. Trial participation may be less attractive than purchasing a drug on the market because of study requirements for blinding and randomization or treatment guidelines.

A partial solution would be to improve clinical trial design so that patients would be willing to participate even when they have other sources of a drug. Recent design changes motivated by criticism of AIDS trials include comparisons of multiple drugs, shorter courses of experimental treatment, midstream adjustment of treatment strategies, and increased use of surrogate markers. When trial participation requires less sacrifice, patients would be more willing to participate to advance future knowledge. Minimizing patient sacrifice is also consistent with a basic principle governing trial participation: a trial should be open only to patients for whom the choice of a recommended treatment remains substantially uncertain.

Because some sacrifice on the part of trial subjects is inevitable, other incentives to participation should be strengthened. These currently include closer medical follow-up and free care. For early drug studies access to otherwise unavailable drugs would remain an incentive for some patients.

Effects on research would not be the only impact of altered standards. Pharmaceuticals account for about 8 percent of national health care expenditures; increased availability of unproven drugs could result in higher health expenditures with less certain health returns. Because consumers are relatively insensitive to medical prices, looser standards might increase the burden on third-party payers to avoid inappropriate drug use. For drugs approved with usage restrictions, reimbursement could be contingent upon appropriate use, just as some medical procedures require specific indications. Payers might also demand certain levels of efficacy or rely on restrictive formularies as a condition for coverage. These measures would be consistent with a broader trend toward establishing effectiveness of medical procedures.

In sum, reducing the evidence required to meet the FDA standard could undermine the incentives to conduct drug evaluation research. This danger can be offset to a significant extent by requiring postmarketing research and imposing other conditions on approval. Because approval would always be based on some indications of a compound's effects, the earliest stages of research would not be jeopardized. Incentives to generate relevant information early in trials would motivate development of new research strategies.

Conclusion

AIDS activists have forced us to think about the regulation of drugs in a new way. For many years we have been content to live with a high level of errors of omission to minimize errors of commission. We have paid the price of a higher standard of proof in exchange for a higher level of protection from drugs. AIDS has emphasized the cost of this form of caution.

The FDA has responded to AIDS patients' pressure for expedited approval and access to unproven drugs by broadening and formalizing preexisting mechanisms. These procedures have systematized and improved patients' access, but they have not fundamentally changed the way the FDA conceptualizes the relationship among drug access, the approval process, and drug research. Whether increased access should occur by means of the approval mechanisms remains a question.

We have argued that a change in the FDA's interpretation of the efficacy requirement is necessary. The author of the Kefauver amendments chose to require "substantial" evidence of efficacy precisely to allow latitude in drug approval. Indeed, FDA policy regarding unproven drugs and devices has evolved gradually during the past decade. To facilitate a more contextual and flexible interpretation of efficacy, the FDA must have incentives to approve drugs on the basis of lower standards of proof and intermediate options that lessen the risks of approval. Explicit policies allowing conditional and probationary forms of drug approval should be developed and oversight mechanisms established.

Changes in access would have offsetting effects. With a lower burden of proof, instances in which patients use drugs that turn out to be inadvisable would increase in frequency. These mistakes would have both dollar costs and health costs. But benefits would accrue because other drugs, which fulfill their initial promise, would be made available sooner and to more patients than under the old procedures. The FDA should protect current and future patients' interests in research by mandating postapproval studies and utilizing as well as possible the experience generated in wider early usage of drugs. Financial and other incentives affecting research and distribution of approved drugs should be examined carefully.

The recent reforms introduced by the FDA have improved access in the case of AIDS and other life-threatening illnesses. While these illnesses have been treated as exceptions, they are actually just one end of a continuum. A more flexible drug approval process would improve equity and efficiency for a broader spectrum of patients. The "reasonable option" interpretation of safety and efficacy should allow consumers greater freedom in weighing the risks of drugs against the risks of illness in situations in which risks of illness are great. In many circumstances access to toxic drugs and drugs of uncertain efficacy may be warranted. Past regulatory developments have moved the locus of control excessively far from patients in these situations.

Protection from errors that benefit no one must remain a goal of drug regulation. But while free access to ineffective or harmful drugs is a dubious social policy, access to less-proven or risky drugs that does not unreasonably raise everyone's "search and avoidance" costs is consistent with the objectives of drug regulation.

A more flexible drug approval process would improve equity and efficiency for a broader spectrum of patients. Access to less-proven or risky drugs that does not raise "search and avoidance" costs is consistent with the objectives of drug regulation.

Selected Readings

Annas, G.J. "Faith (Healing), Hope and Charity at the FDA: The Politics of AIDS Drug Trials." *Villanova Law Review*, Vol. 34 (1989).

Byar, D.P., Schoenfeld, D.A., et al. "Design Considerations for AIDS Trials." *New England Journal of Medicine*, Vol. 323 (1990).

Merigan, T.C. "You Can Teach an Old Dog New Tricks." *New England Journal of Medicine*, Vol. 323 (1990).