



FDA Standards — Good Enough for Government Work?

Jerry Avorn, M.D.

The cliché “good enough for government work” implies that lower standards are acceptable for a job sponsored by a public agency. But in biomedical research, the opposite is usually true. The National

Institutes of Health has always had tough standards; its newly constrained funding is leading to an even more stringent review process, so that near-perfect evaluation scores are now required to win support. Similarly stringent criteria prevail at the National Science Foundation. Yet there is one area of biomedicine in which the government allows — even defends — a minimal standard that would be unacceptable anywhere else in research. It is the set of evidentiary requirements maintained by the Food and Drug Administration (FDA) for the approval of new drugs.

This is not to suggest that the

FDA condones sloppiness — quite the opposite. Like a patient with obsessive-compulsive disorder, the agency is single-mindedly preoccupied with demanding the meticulous performance of a series of relatively simple acts — proving that a new medication is superior to a usually irrelevant comparison treatment (such as placebo) in achieving a potentially irrelevant outcome (such as a surrogate measure). The sloppiness resides not in the quality of execution the FDA requires, which is high, but in the questions it asks.

Several drug-approval decisions illustrate the problem. Some concern the most lucrative kind of

medications: those taken for extended periods by huge numbers of basically healthy people. Such “lifestyle” drugs may initially be evaluated for the treatment of a real clinical problem, such as severe obesity, but there may be no clear consensus defining the “mild” end of the disease-nondisease continuum. As a result, the market can be cranked up by aggressive promotion to both patients and prescribers. Here the comparison of a drug’s benefits and risks is vitally important, but the government generally does not require such assessment. Consider the latest two “epidemics” facing Americans: overweight and insomnia. For both conditions, the application of current regulatory standards can result in important clinical and economic problems.

The most notorious example of an appetite-control drug that

the FDA deemed good enough for approval was dexfenfluramine (Redux), the *d*-isomer of the decades-old, minimally effective fenfluramine that became part of the fen-phen diet-pill craze of the late 1990s. Fenfluramine was known to cause pulmonary hypertension that could be fatal, and its *d*-isomer was expected to do so as well. But despite this risk, the drug was approved in light of the supposedly more worrisome epidemic of obesity that it might help to thwart. What were the medication's credentials? In its pivotal preapproval trials, patients randomly assigned to receive dexfenfluramine lost an average of about six pounds more than those assigned to placebo. No meaningful improvement was demonstrated in blood pressure, lipid levels, or glyce-mic control. But the costly product worked better than nothing at the $P < 0.05$ level, and it was therefore approved. Practitioners and "pill mills" throughout the country then used it to treat millions of women who wanted to shed a few pounds for cosmetic reasons. The expected pulmonary hypertension complications occurred, as did an unanticipated side effect, cardiac valvulopathy. The blockbuster was withdrawn from the market in its first year of use.

We now await the FDA's review of rimonabant (Acomplia), the first endocannabinoid-receptor inhibitor to be brought before it. The drug is said to reduce appetite, lipid levels, and the desire to smoke. If its manufacturer seeks an antiobesity indication,

the agency may well use its favored criterion: modestly greater weight loss than that achieved with placebo, even if it is only temporary. Patients who took the drug in controlled trials had higher rates of withdrawal because of neuropsychiatric and gastrointestinal

***A broader vision
is possible even under
the FDA's
current mandate.***

disorders than did control subjects.¹ But if the FDA applies its usual standards, the drug could be on the market despite these problems.

The second "epidemic" the public is being warned about is insomnia, and analogous concerns apply to the increasingly widely used hypnotics. The FDA has tended to approve sleep aids if they are superior to placebo in terms of polysomnography-laboratory measures such as sleep latency (the interval between the time a person attempts to fall asleep and the onset of sleep measured on an electroencephalogram). The studies typically last a few weeks at most, even though long-term use of the drug is often anticipated. The recently evaluated melatonin-receptor agonist ramelteon (Rozerem) was approved on the basis of brief sleep-laboratory studies that used one- or two-night assessments to demonstrate a reduction in sleep latency.² (A more

relevant study of patient-reported sleep onset did not demonstrate a difference between the drug and placebo in patients younger than 65 years of age.³)

Daytime somnolence caused by severe chronic insomnia can be a real problem for some patients, but commonly taken hypnotics can also cause next-day drowsiness, cognitive impairment, and an increased risk of falling, especially among older patients who are their most frequent users. How does one weigh these very real risks against benefits defined in the sleep laboratory, especially in the case of long-term use? How much better will the new agents be than their predecessors in terms of these clinically relevant matters? The FDA's usual trial standards ignore these questions, and comparative studies of different agents are not part of the evaluation process.

The problem of minimalist government requirements extends beyond overly simple surrogate measures and the failure to demand relevant comparison trials. The FDA recently approved BiDil, a branded combination of hydralazine and isosorbide dinitrate, on the grounds that it was uniquely effective in treating congestive heart failure in black patients. The problem here was not the end points studied, which included death — the cleanest outcome measure known to medicine. Rather, it was the premise on which the drug was approved for use specifically in that racial group. The contention that the hydralazine-nitrate combination

works better in blacks than in whites was based on a post hoc analysis of racial subgroups enrolled in a larger trial in which the combination did not perform particularly well.³

This interesting observation could have been enormously important in helping us understand the pathophysiology of this common and often devastating condition in a particularly vulnerable population. A plausible next step would have been to test the racial-difference hypothesis in a controlled trial enrolling both blacks and whites in order to look for differences in outcomes as well as predictors of those differences, including genetic markers, self-identified race, diet, and other risk factors. But no such study was required for approval. Instead, the racial-difference assumption was embraced as fact, and the new pivotal study enrolled only blacks; it found that adding BiDil to their regimens worked better than adding placebo. The drug was then approved as a new and expensive treatment for congestive heart failure specifically for blacks.

Defenders of the FDA point out that its enabling legislation requires the agency to approve a drug if it is found to be effective in well-conducted clinical trials. The pharmaceutical industry has argued that this criterion is met if a product is shown to be better than placebo in achieving a surrogate outcome in a short study. But a broader vision is possible, even under the current legal mandate. The agency has been willing to flex its regulatory muscles

in the interest of better science in the past. For example, it will not allow a company to claim that its osteoporosis drug prevents fractures if the trial data demonstrate only an increase in bone mineral density. Thus, it should be well within the purview of the FDA to decide what “effective” really means.

The agency is also required to determine whether a drug is safe enough for use, a decision that can be reasonably made only in relation to the drug’s actual clinical usefulness. Since all drugs have side effects, it is not a stretch to expect that the approvability of a drug should take into account whether its risks are acceptable in light of its real-world effectiveness. This would require evaluating clinical benefit by means of a more relevant measure than short trials with surrogate outcomes. It would also require consideration of a drug’s efficacy and safety as compared with alternative therapies. If such studies are not required as part of the approval process, it seems that we don’t have any way to ensure that they are ever conducted; as a result, they usually are not.⁴

Manufacturers have claimed that such evaluation requirements would make preapproval testing too lengthy and expensive, but that is not a compelling argument. The sums spent by the large pharmaceutical companies on meaningful research and development are less than a third of what they spend on marketing, promotion, and administration.⁵ A rebalancing of this relationship would be

quite feasible and could generate more clinically useful content for all those promotional activities. The better prescribing that such improved data would make possible would surely save the country more than the new approach would cost, since it would allow doctors, patients, and payers to understand the true value of a costly new product. Important new drugs that meet urgent and serious health needs could still be provisionally approved on the basis of less demanding studies but with a required reassessment a few years later to evaluate more relevant outcomes. Such a reassessment could be required of the manufacturer as a condition of keeping the drug on the market, or it could be undertaken by independent drug-evaluation units, with the results disseminated broadly to inform decisions on prescribing and purchasing drugs.

Some in the industry would argue that the lowest possible standard of efficacy is good enough and that an act of Congress would be required to change the current rules. But such an act is not inconceivable. Increasing public concern about efficacy–risk–cost trade-offs may move this agenda forward in Washington, especially if Medicare becomes the nation’s largest drug purchaser in 2006. The ballooning cost of that program may bring together clinical scientists, advocates of prudent federal spending, and even free-market aficionados, all demanding more useful standards. The idea that government approval should be based on what a new drug really

does for patients may soon come into its own.

Dr. Avorn is a professor of medicine at Harvard Medical School and chief of the Division of Pharmacoepidemiology and Pharmacoeconomics at Brigham and Women's Hospital, Boston.

An interview with Dr. Avorn can be heard at www.nejm.org.

1. Van Gaal L, Rissanen AM, Scheen AJ, Ziegler O, Rossner S. Effects of the cannabinoid-1 receptor blocker rimonabant on weight reduction and cardiovascular risk factors in overweight patients: 1-year experience from the RIO-Europe study. *Lancet* 2005;365:1389-97.
2. Ramelteon (Rozerem) (package insert). Lincolnshire, Ill.: Takeda Pharmaceuticals North America, July 2005. (Accessed August 8, 2005, at <http://www.fda.gov/cder/foi/label/2005/021782lbl.pdf>.)
3. Carson P, Ziesche S, Johnson G, Cohn JN.

- Racial differences in response to therapy for heart failure: analysis of the vasodilator-heart failure trials. *J Card Fail* 1999;5:178-87.
4. Avorn J. Powerful medicines: the benefits, risks, and costs of prescription drugs. New York: Alfred A. Knopf, 2004.
 5. Centers for Medicare & Medicaid Services. Health care industry market update: pharmaceuticals. Baltimore: Centers for Medicare & Medicaid Services, January 2003. (Accessed August 8, 2005, at http://www.cms.hhs.gov/reports/hcimu/hcimu_01102003.pdf.)

The Justice Department's Case against the Tobacco Companies

Michael C. Fiore, M.D., M.P.H., Paula A. Keller, M.P.H., and Timothy B. Baker, Ph.D.

The case brought by the Department of Justice against the tobacco industry — the largest civil litigation in U.S. history — has become mired in controversy. Projections suggest that about 23 million Americans, or one of every two current smokers, will die prematurely because of a disease caused by tobacco use. The lawsuit called for actions to remedy harms that have resulted directly from misconduct on the part of the tobacco industry. Unfortunately, Justice Department attorneys shifted gears at the 11th hour, drastically reducing the remedies proposed by their own expert witness, and reports surfaced that government witnesses were pressured to water down their testimony. The reasons for this behavior remain murky, but the likely long-term effects seem clear: more latitude for the tobacco companies, more new smokers, and more smoking-related illness and death.

The origins of the case can be traced back to 1999, when in

the wake of the Master Settlement Agreement of 1998 (which did not earmark monies for tobacco control), President Bill Clinton announced in his State of the Union address that the Justice Department would begin litigation against the tobacco companies. The initial filing in late 1999 was based on efforts to recover Medicare funds expended as a result of tobacco-caused illness and on the civil federal Racketeer Influenced and Corrupt Organizations (RICO) Act, which provides a mechanism for preventing and restraining unlawful racketeering activity.

After five years of preparation, the trial began in September 2004 and was divided into two parts — a liability phase and a remedies phase. Rulings regarding both phases are expected in late 2005. In the liability phase, the Justice Department focused on what it called the “seven pillars of fraud” in portraying tobacco-industry misconduct (see list).

In response, the tobacco companies argued that, even if they had engaged in improper behavior in the past, they had ceased to do so after the Master Settlement Agreement and had become law-abiding corporate citizens.

The proposed penalties presented by the Justice Department during the remedies phase of the trial were constrained by two prior court rulings. In 2000, presiding U.S. District Court Judge Gladys Kessler ruled that penalties in the case could not be used to offset Medicare costs. Then, in February 2005, with the trial in its fifth month, an appellate court decided that a \$280 billion disgorgement remedy was not “forward looking” and would not fulfill its ruling that RICO remedies must “prevent and restrain” future wrongful conduct (a decision that the Justice Department appealed to the Supreme Court on July 18). The Justice Department responded by developing a new series of penalties, designed