THE EFFECT OF GOVERNMENT REGULATION ON THE EMERGENCE OF A NEW MEDICAL TECHNOLOGY

Jack B. Homer
System Dynamics Group
Alfred P. Sloan School of Management
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

This paper explores the possible paths of emergence of a new medical technology and how those paths might be altered by government regulations of the sort now promulgated by the Food and Drug Administration (FDA). The purpose of the paper is to help clarify the role of FDA regulation in a dynamic context. The analysis focuses on the idea that an emerging technology's effectiveness may change over time and that the benefits and losses due to regulation may themselves have a dynamic character. An increasingly complex story of the emergence (or dissemination and development) process is told with the help of causal-loop diagrams. Results from a preliminary system dynamics model based on this story are illustrated and discussed. They suggest that the FDA's actions may have unintended effects, such as slower development of a technique, which may or may not be harmful. They also suggest that, in certain cases, post-marketing surveillance and communication of results may be at least as important an activity for the FDA as pre-marketing evaluation.

Background

The proliferation of new medical technologies in recent years has produced excitement about the opportunities to improve the length and quality of life but has also produced much concern over the sometimes high costs and risks involved. The stakes have steadily been raised in the medical profession, and for every long string of life-saving successes there seems to have been at least one disaster or near-disaster, such as the liquid sulfanilamide tragedy in 1938 which cost nearly 100 lives. [1] The effect of these dramatic failures has been to sensitize the public to the potential benefits of broad government regulations that would effectively protect the unsuspecting consumer from dangerous, inferior, or inappropriately-applied medical technologies. Both patients and their physicians now demand official assurance that widely available drugs and devices have been thoroughly tested and accurately labeled.

In 1962, following the thalidomide disaster in Europe which revealed shoddy drug evaluation practices in the U.S., the Food, Drug, and Cosmetic Act of 1938 was amended to include requirements that manufacturers demonstrate both safety and efficacy of new drugs prior to marketing, by providing the FDA with "substantial evidence" from well-controlled investigations. [2] In 1976, the Act was again amended to provide requirements for certain new medical devices similar to the requirements that already existed for new drugs. [3] Marketing approval for a drug or a device may be withdrawn or distribution

restricted to a limited number of investigators at any time, if new evidence casts a reasonable doubt on the technology's safety or efficacy. [4] The purpose of the proof-of-efficacy requirements codified in the 1962 and 1976 Amendments was to reduce the consumer losses attributable to ineffective new drugs and devices. Careful testing would ideally identify all potential problems with a technology before it is widely distributed and enable the FDA to effectively prevent a "run-away" of catastrophic magnitude.

Primarily since 1962, the FDA has become one of the largest and most complex regulatory institutions in this country. Not surprisingly, it has also become one of the most controversial of the Federal agencies and has been criticized at various times for acting too slowly or too quickly, for being too lax or too strict, for being overly consumer-oriented or overly industry-oriented, and for attempting to control too much or too little of the research, development, and dissemination process. In fact, some critics claim that the FDA's rigid interpretation of the 1962 Amendments has actually been responsible for a net loss to society, citing the reduced flow of all medical technologies (not only ineffective ones) to the marketplace since the late 1950s. [5] [6] [7]

The Dynamics of Usage and Effectiveness

Prologue

This section of the paper sets the stage for a discussion of the possible effects of regulation on various patterns of usage and effectiveness for a new medical technology. It begins with the description of a model which produces two well-known patterns of usage for medical technologies, namely an S-shaped pattern for a "successful" technology and a rise-and-fall pattern for an "unsuccessful" technology. [8] [9] [10] [11] The differences between these two patterns is attributable to different input assumptions for the technology's effectiveness, a concept which is re-examined and seen to be potentially more flexible than the initial model allowed. The boundary of the model is therefore expanded to permit a dynamic view of effectiveness, and a number of important new feedback loops appear. With a richer structure, the model exhibits an additional usage pattern which has not, to the best of the author's knowledge, appeared in the empirical research literature to date. Nonetheless, the relative sparseness of this literature and the reasonableness of the simulation results have encouraged the author to accept the possible existence of the "new" behavior and to investigate the effects of regulation on it.

The full model was developed under the auspices of the National Heart, Lung, and Blood Institute for the purpose of evaluating the likely impacts of alternative government policies on

the emergence of percutaneous transluminal coronary angioplasty (PTCA), a new technique for the treatment of coronary artery disease, one of the major killers in the United States today. This procedure uses a specially designed, balloon-tipped catheter to reduce obstruction in coronary arteries. PTCA is considered an alternative for some patients to the popular but considerably more costly and invasive coronary artery bypass graft (CABG) surgery and has attracted much attention recently among practitioners and policy-makers. Because of the author's familiarity with PTCA and its inherent richness as a case study, PTCA will be used as the primary source of examples in the following model description.

A Model with Fixed Effectiveness

The first model to be considered interprets the dynamics of usage as primarily a technology dissemination issue, focusing on the acceptance and possible rejection of a new drug or device over time. A somewhat simplified diagram of this model is presented in Figure 1.

The supply of procedures (usage) is measured on a flow basis (for example, 1000 PTCAs per year) and will be equal to demand, unless practitioners of the technique are overloaded with referrals. Overloading will not become a problem, however, if the opportunities for practice are readily recognized and there are no impediments to obtaining the resources (including specific materials and training) that are necessary to become a practitioner. Demand for the procedure is generated by the recommendation of physicians who are not necessarily

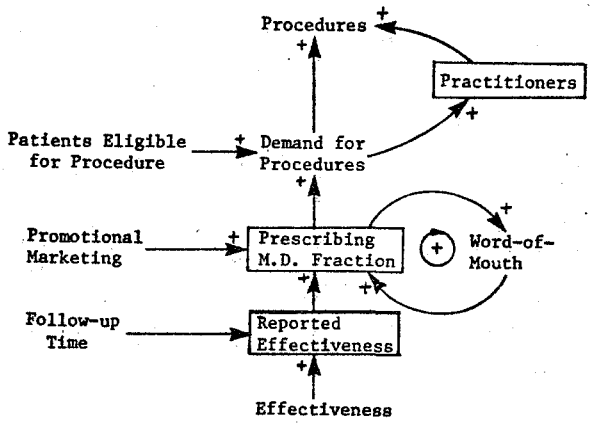


Figure 1. A Model with Fixed Effectiveness

practitioners but are qualified to determine the patient's likely eligibility for the procedure. In the case of PTCA, these prescribing physicians are primarily cardiologists who perform various diagnostic procedures to determine the precise nature and severity of their patients' heart troubles and decide on the best course of therapy, which may consist of medication, CABG surgery, or PTCA. The physician's assessment of patient eligibility is largely determined by existing protocols or indications established by experts in the field and communicated to him or her by medical journals, by manufacturers via promotional material and package inserts, or by fellow colleagues.

These same three basic sources of information are also responsible, to varying degrees, for making the physician aware of the new technology and persuading him to recommend it to the indicated patients. [8] [12] Figure 1 indicates that promotional marketing, journal reports on the technique's effectiveness, and word-of-mouth among colleagues, may all affect the fraction of physicians who have accepted and are prescribing the technology. Although all three of these influences on adoption are represented dynamically in the model as elements of various feedback loops [13], the most important of these loops involves the word-of-mouth effect. This says that as the fraction of prescribing physicians increases, there will be more physicians informing their colleagues about the technology and persuading them to try it.

Rejection or disadoption of a new medical technology generally will not take place unless new evidence indicates a level of effectiveness which is both lower than was originally reported and lower than at least certain prescribing physicians can tolerate. The reported level of effectiveness may differ from the actual if the full range of effects has not yet been observed or appreciated. Extremely rare or delayed effects may require years to be detected, especially if the technology causes cancer or genetic damage. This inherent problem in evaluating medical technologies has been brought to the fore by the delayed discovery, in a number of instances, of severe adverse reactions, such as thromboembolism from oral contraceptives and thyroid cancer from head-and-neck irradiation. [1] [14] As the period of

follow-up evaluation increases and more of the technology's after-effects are observed, reports of effectiveness will tend to become more accurate.

Effectiveness refers to the extent to which the procedure directly or indirectly improves the "health status" of the average recipient, taking into account all of the short-term and long-term benefits and risks associated with the procedure. [15] In the case of PTCA, the results of the procedure can range anywhere from the immediate disappearance of all chest pain and a complete return to normal activities, to death from complications such as arterial dissection; and the long-term effects on length and quality of life are still not known for certain. [16] Clearly, the existence of a range of possible outcomes for any technology, especially a risky one, makes it difficult to assign a single number called "effectiveness" which can summarize the technology's medical value. However, the following definition works nicely, given the model's implicit assumption that the patient plays a purely passive role in the decision to prescribe: Effectiveness is the degree to which the average physician would judge that the established protocols for patient eligibility are justified, if he or she were aware of the complete distribution of patient outcomes. [17]

The model discussed above can generate two distinct paths of emergence, which are illustrated in Figures 2 and 3 over a ten year time horizon. For both simulations, it was assumed that the procedure has adverse after-effects which require three years for complete

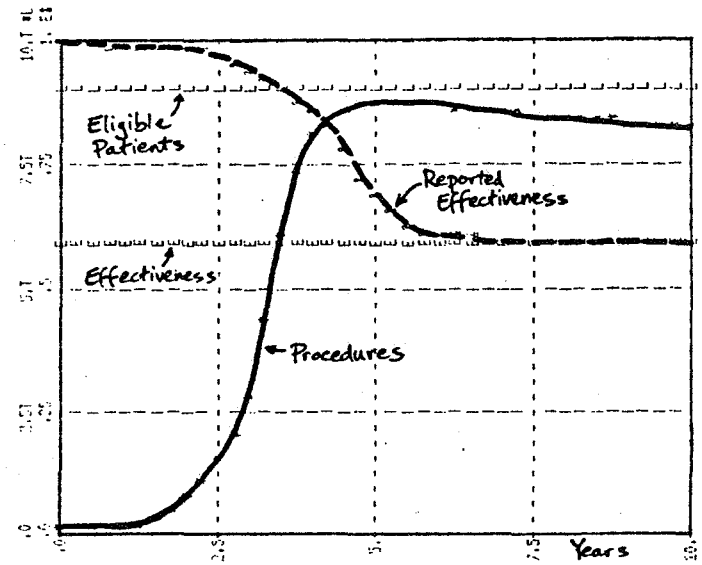


Figure 2. S-Shaped Usage Pattern for an Effective Technology

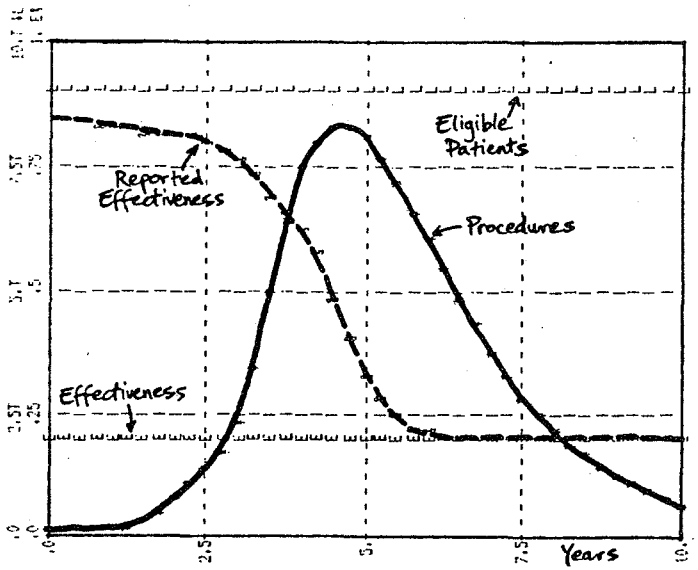


Figure 3. Rise-and-Fall Usage Pattern for an Ineffective Technology

observation. [18] The only difference between the two scenarios is that the first assumes a level of effectiveness which is acceptable to most physicians, while the second assumes a largely unacceptable level of effectiveness.

In Figure 2, the initial level of reported effectiveness is quite high and the word-of-mouth effect has no trouble in driving usage upward until it is close to the total number of eligible patients by year 5. As the adverse after-effects are discovered and reported between years 3 and 6, however, enthusiasm for the technology wanes slightly and the number of procedures declines accordingly. The rejection which occurs as a result of mixed effectiveness remains small in comparison to the forces encouraging acceptance, and usage stabilizes at about 90% of its maximum value. The technology may be considered successful.

In Figure 3, the initial level of reported effectiveness is lower than it was in Figure 2, but word-of-mouth is again successful in producing rapid growth in procedures for the first four years of the simulation. However, by year 5, reported effectiveness has declined to a point which is unacceptable to most physicians. [19] Rejection of the technology begins at this time and continues steadily through the end of the run. The main reason for the relatively slow decline in usage is that loyal prescribing physicians are able to continue persuading some of their colleagues to accept the technology, as long as the discouraging evidence is not conclusive; that is, as long as recent reports have not completely displaced older notions of the

technology's value. This result supports empirical evidence which suggests that the system may become more inflexible after a large fraction of physicians have adopted the technology. [20] Nevertheless, the technology represented in Figure 3 is ultimately unsuccessful due to its low effectiveness.

A Revised View of Effectiveness

The preceding discussion indicated that evaluations of a technology's effectiveness may change substantially over time because of delays in perceiving the full range and distribution of outcomes, that is, the "true" effectiveness of the technology. The task of evaluation is further complicated, however, by the fact that the technology's effectiveness may itself change over time, so that evaluation results may be obsolete by the time they are published. [15] [20] Effectiveness may change either because the technology itself is evolving or because the circumstances under which it is applied are variable. Many complex new technologies are subject to continuous modification and improvement by both manufacturers and experts in the field. The materials and equipment used in PTCA, for example, have already undergone a number of important modifications and additions since the technique's introduction in 1977, including new catheter shapes and sizes.

The circumstances of usage which may have an impact on effectiveness include practitioner expertise and the protocols or criteria physicians use for selecting patients. Complex techniques

such as CABG surgery or PTCA require a high level of skill gained through repeated experience, and even the proper administration of drugs often requires a detailed knowledge of dosage and interaction effects that is at least partially acquired through trial and error. [21] [22] Aside from practitioner incompetence, effectiveness may be low because the technique is being used on too broad a class of patients relative to its inherent capability. If various subsets of the patient population can be identified and their differential outcomes pinpointed, then effectiveness may be increased by narrowing the patient selection criteria to those patients most certain to benefit from the technology. [23] [24] For example, the PTCA procedure was originally attempted with patients suffering from either single- or multiple-vessel coronary artery disease. Observation of higher risk for multiple-vessel patients, however, led to an awareness of the technique's limitations (at that time) and suggested greater selectivity in the use of PTCA. [16] [27]

A Model with Variable Effectiveness

The model to be discussed below permits effectiveness to change along with its three determinants: patient selection criteria, practitioner skill, and technical capability. These factors will be introduced in order and their behavioral implications discussed.

The structural consequence of allowing the patient selection criteria to vary in response to evaluations is shown in Figure 4. The idea here is that practitioners will attempt to adjust their protocols

to a point that appears to be justified by recent evaluations. In equilibrium, the reported effectiveness will equal 1, indicating that the selection criteria are apparently fully justified by the results. In order for this goal to be achieved, however, evaluations must consider the differences in outcomes associated with various subsets of the eligible patient population, the results must be efficiently communicated to practitioners, and practitioners must be willing to change their patient selection criteria. In other words, the selection criteria must be flexible in order for the negative loop in Figure 4 to be influential in controlling effectiveness. Unfortunately, this is not always the case. [26]

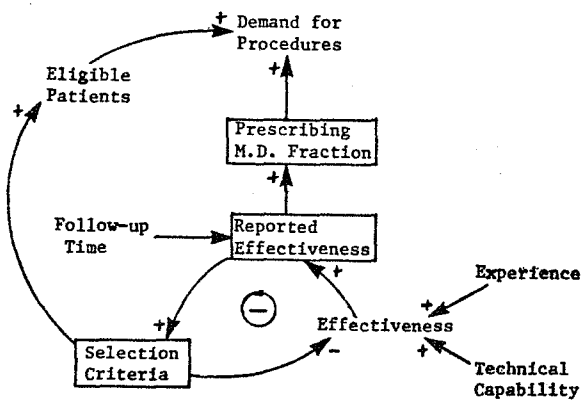


Figure 4. Variable Selection Criteria

As a result of introducing the assumption that the patient selection criteria are responsive to evaluations, a new mode of behavior becomes possible, one in which usage overshoots and then undershoots its final equilibrium value. The additional assumptions needed to produce this third path of emergence are the same ones responsible for the rise-and-fall behavior seen in Figure 3; namely, that the technology is initially quite low in effectiveness and that this fact is not detected for several years. In response to reports of low effectiveness, the selection criteria are narrowed, which accelerates the decline in demand already occurring due to rejection. But as the selection criteria are narrowed, effectiveness rises, and physicians are eventually persuaded to readopt a technology which is less broadly applicable but much more effective than it originally was.

Figure 5 shows the structural consequence of assuming that practitioner skill is variable and a function of experience. Although the process is more complex than the diagram indicates, the average level of experience per practitioner will drop as a result of growth in practitioners. In essence, rapid growth implies a larger fraction of novice practitioners than usual. Lower experience leads to lower effectiveness, and that, when perceived, leads to slower growth in demand. Finally, slower growth in demand leads to slower growth in practitioners. The major effect of the negative loop just described is to decrease effectiveness and the growth in usage somewhat during the technology's adoption phase.

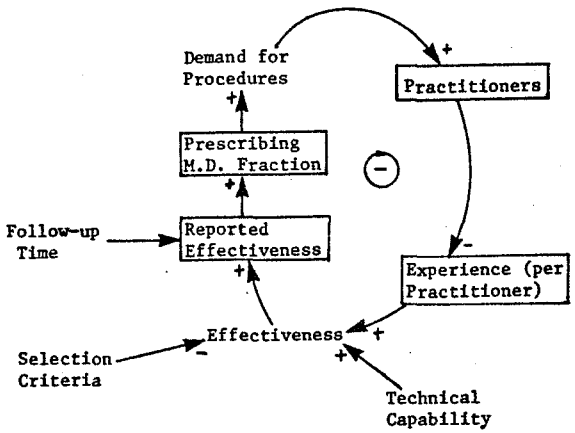


Figure 5. Variable Practitioner Skill

Figure 6 shows the structural consequence of assuming that the technique's inherent (but only indirectly measurable) capability can improve to some degree. Improvements may occur along various dimensions, including safety, accuracy, scope, and specificity. [27] Modifications of a medical technology are generally engineered, produced, and distributed by manufacturers, whose ideas may be largely based on the suggestions of innovative practitioners. [28] [29] As the technology's capability advances, these opportunities will gradually decrease, until it is no longer perceived as worthwhile to invest more time and money into modifying the technology. [30] The consequences of improving the technology are two-fold. First, perceived technical

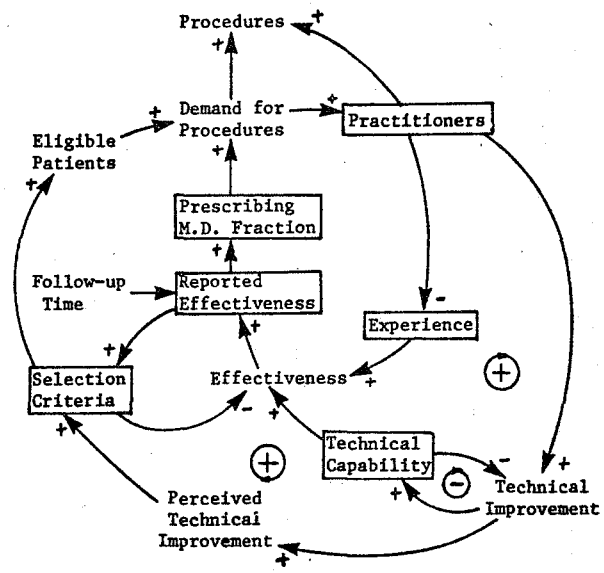


Figure 6. A Model with Variable Effectiveness, Including Technical Improvement

improvements are the major driving force behind the expansion of patient selection criteria. For example, the development of smaller PTCA catheters that can enter previously inaccessible coronary arteries has permitted a greater variety of applications. Second, as long as the selection criteria do not expand too quickly relative to the rate of technical improvement, increasing technical capability will result in greater effectiveness for a technology previously low in effectiveness.

Figure 6 illustrates three potentially important feedback loops created by the inclusion of technical improvement. The identified negative loop represents the fact that there is a virtual limit to such improvement. The positive loops indicate that as the technology improves and the selection criteria and effectiveness increase, demand for the technique will also increase, which may entice more innovative practitioners into the field; this may result in an even faster or at least a sustained, rate of technical improvement for a number of years.

Figure 7 displays a simulation run of the complete model with variable effectiveness, featuring the overshoot-and-undershoot usage pattern discussed previously. The assumptions for this simulation are: (1) effectiveness is initially quite low, because the selection criteria are too broad relative to the true technical capability and because practitioner skill is moderate [31]; (2) reported effectiveness is fairly high for several years until the full range of after-effects is observed; (3) the selection criteria are flexible; (4) technical capability can be increased to approximately twice its initial value. The appropriate time horizon for this scenario is about twenty years, in comparison to the ten year simulations shown in Figures 2 and 3.

The story that can be told for Figure 7 is really only an elaboration upon what has already been said about the overshoot-and-undershoot pattern. During the initial period of adoption, skill remains moderate on average, as a result of the inflow of

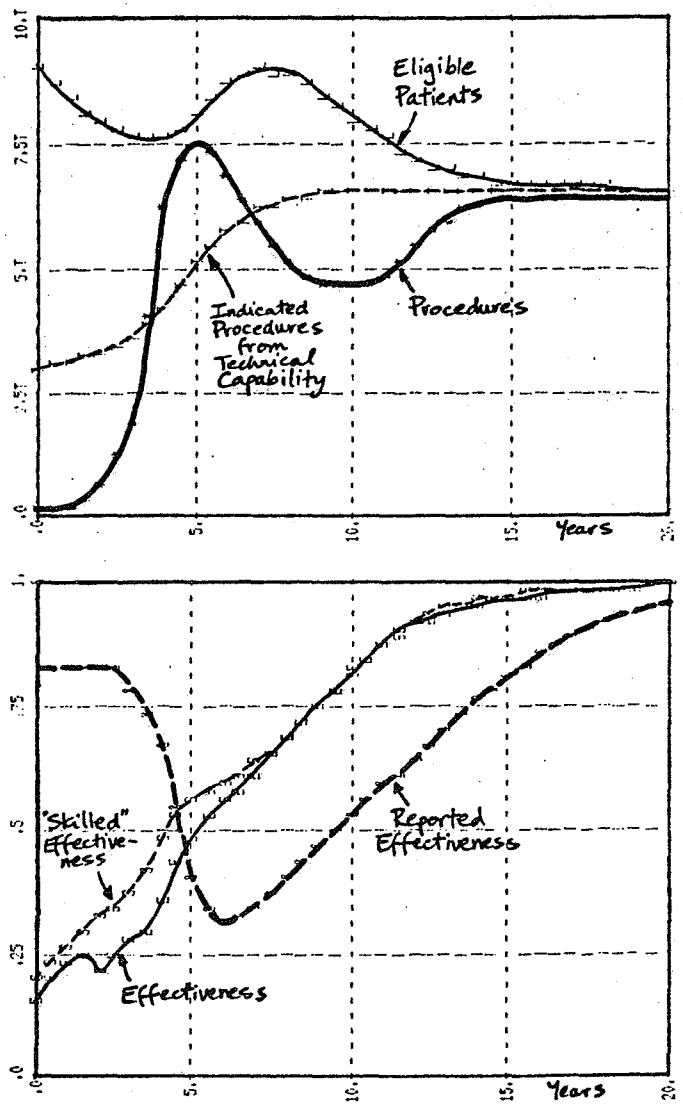


Figure 7. Simulated Usage Pattern for a Technology with Increasing Effectiveness—Unregulated Scenario

inexperienced practitioners. This accounts for the gap between "skilled" and actual effectiveness through year 7 or so. Technical improvements increase the potential number of effective procedures for about eight years, until decreasing returns finally shut off attempts to modify the technique. These improvements are responsible for expansion of the selection criteria starting in year 4 and continuing through year 7, even though the reported effectiveness is rather low during much of this period. The first reports of low effectiveness are essentially ignored by protocol-setting practitioners, who perceive that improvements in the technology have made this information obsolete. To some extent, this perception is accurate, as a comparison of actual and reported effectiveness from years 5 to 7 indicates. However, as evaluations continue to report mixed outcomes, the selection criteria are narrowed and effectiveness climbs steadily. As reported effectiveness rises, most of those who rejected the technology between years 5 and 9 come to believe that the narrower selection criteria have been vindicated by the results, and they become prescribers again. By the end of the simulation, usage has risen to nearly its full potential, and effectiveness is near its equilibrium value of 1.

The Dynamics of FDA Regulation

Structure

In the present context, the FDA regulations affecting new medical technologies may be considered collectively as an endogenous policy of the health care system which reduces both demand and supply until a convincing body of evaluative data suggesting adequate effectiveness has been accumulated. On the demand side, the regulations include pre-marketing restrictions on advertising and promotion. On the supply side, the regulations will make it difficult for physicians to obtain the materials and the institutional permission required to practice the technique, before it has been approved for widespread use. Figure 8 indicates, in a simplified way, the structural consequences of adding regulation to the model described in the last section. From the diagram, it would appear that regulation does little more than reinforce the loops involving reported effectiveness already present in Figure 6. Indeed, the main intent of regulation is to strengthen the existing role of evaluations in controlling usage, so as to oppose those forces that might encourage rapid and early growth in spite of inadequate or discouraging data.

Although the FDA's actions may have their intended effects, there may be unintended effects as well. If skill is a factor in effectiveness, then outcomes observed during the slow-growth period of regulatory restrictions may be significantly different than the outcomes observed after those restrictions are lifted. [32] The

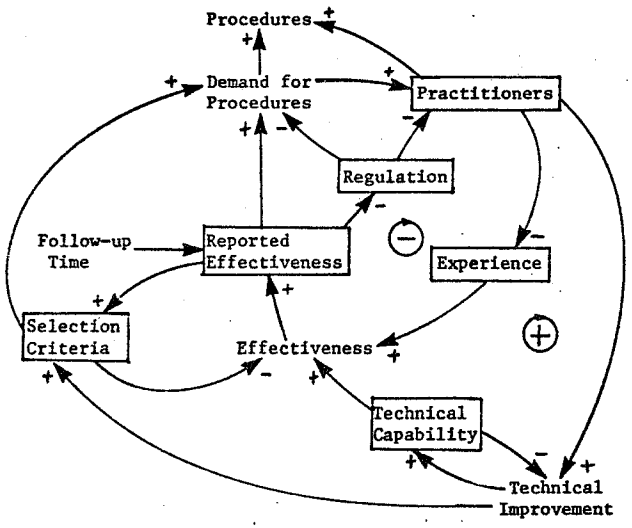


Figure 8. A Model with Variable Effectiveness and Endogenous Government Regulation

implication is that the FDA may be persuaded by early encouraging results to approve marketing of a technology that should be placed only in the hands of experts. Another unintended effect is that by limiting the inflow of practitioners, the FDA may delay the technology's improvement. This may be of critical importance in the case of an initially ineffective technology which can, through improvement, become effective. The indicated positive loop in Figure 8 may correspond to a self-fulfilling prophecy of ineffectiveness, if the FDA's restrictions prevent or severely delay this improvement.

Behavior

If the technology's effectiveness is assumed to be fixed, conclusions from the model can contribute little to the current debate surrounding the FDA. The model's behavior under this assumption (not shown here) confirms the popular idea that government regulation serves mainly to delay the introduction of a new medical technology. [33] In the case of a permanently effective technology of the sort represented in Figure 2, this delay results in lost opportunities to help patients. In the case of a permanently ineffective technology of the sort represented in Figure 3, this delay results in monetary and possibly health-related savings associated with not using the technology. The total amount lost or saved in each of these polar cases is a function of the length of the regulatory delay. For an effective technology, the shorter the delay the better. For an ineffective technology, the longer the delay the better. This much should be obvious.

If effectiveness is allowed to vary, the effect of regulation on a new technology's emergence may be rather complex, and the policy's benefit or cost to society may change dramatically over time. Figure 9 displays a simulation run of the complete model under the same basic assumptions used for generating Figure 7, but with an endogenous FDA regulation policy in place. As it turns out, this policy delays the initial growth in procedures by about two years and entirely eliminates the overshoot seen in Figure 7. Although improvement of the technique is also delayed (by about 1 1/2 years), there is

little difference in effectiveness between the two scenarios through year 5, because, in part, of greater practitioner skill under the slow-growth conditions of regulation. Starting in year 5 and continuing until about year 11, effectiveness is actually higher with regulation than without it. This occurs not only despite the delay in technical improvement, but also, ironically, as a consequence of it. In particular, the first reports of low effectiveness are not treated by physicians as obsolete to nearly the same extent that they were in Figure 7; as a result, the selection criteria drop rapidly through year 6 and are never far above their final equilibrium value thereafter. The relatively narrow criteria lead to greater effectiveness, but they also depress demand. The lower usage seen with regulation than without it through year 12 is, however, only partially attributable to narrower selection criteria. More importantly, the regulatory restrictions enforce slow growth during most of this period, and they are not lifted until the early discouraging reports are displaced (in the regulators' minds) by an accumulation of the newly supportive evidence.

Perhaps the best way to summarize the differences between Figures 7 and 9 is by noting that usage is greater in the unregulated case than it is in the regulated case during both the initial period of low effectiveness and the later period of high effectiveness. If one accepts the generalization that regulation tends to be beneficial in the case of a permanently ineffective technology and detrimental in

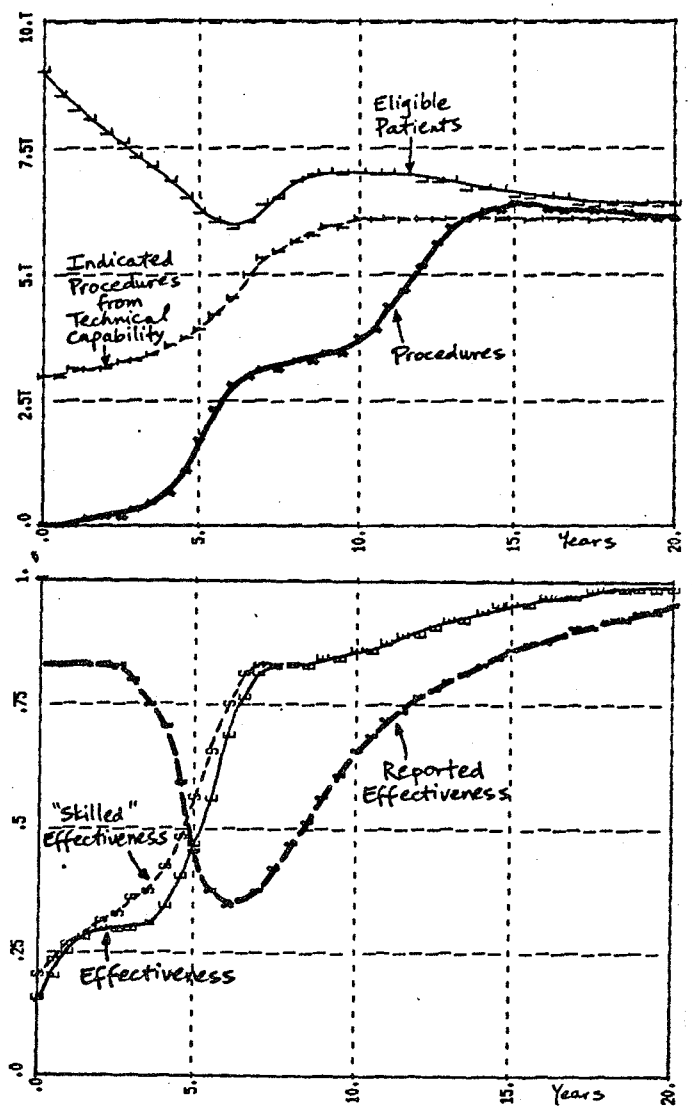


Figure 9. Simulated Usage Pattern for a Technology with Increasing Effectiveness—Regulated Scenario

the case of a permanently effective technology, then it seems natural that the benefits of regulation should decrease as an initially ineffective technology becomes effective. Indeed, a rudimentary cost-benefit comparison of these two simulations (not shown here) indicates that the net benefit that initially accrues from regulation can quickly reverse and become a net loss after a "borderline" value of effectiveness (from society's viewpoint) is exceeded. [34]

Conclusion

Discussions of government regulation often note that any attempt to protect the public from potential hazards will invariably be at the expense of some of the fruits of the regulated industry. [1] Clearly, regulation involves tradeoffs. However, it should be recognized that these tradeoffs are not necessarily a fixed entity, unamenable to government policy. The fact that regulators make decisions under conditions of great uncertainty does not excuse them from attempting to increase the odds of their success. This paper has demonstrated, for example, that the flexibility of patient selection criteria for a new medical technology may have an important influence over its ultimate chances for being effective. This suggests that an important activity of the FDA might be to increase the medical community's responsiveness to information that indicates a need to adjust

the selection criteria for a new technology. In fact, it has been argued that if the FDA stepped up its post-marketing surveillance and information dissemination functions, it could afford to relax some of its strict pre-marketing requirements. [32]

While it is appropriate to expect that the FDA's actions can improve the chances of a given technology's success, the public should not be led to believe that the FDA can actually guarantee the effectiveness of every new technology. Unless the American people are willing to tolerate extremely long delays in the marketing approval process, there will always be some technologies whose after-effects will only be seen after they are widely disseminated. The FDA should probably put a greater emphasis than it has on creating mechanisms or augmenting existing ones that tend to improve the resiliency of the health care system, instead of focusing entirely on an attempt to perfect the evaluation process. The public would be best served by flexible government policies which provide helpful feedback where it is needed rather than by rigid policies in search of an ideal.

References and Notes

[1] Crout, J.R. The nature of regulatory choices. University of Rochester: Center for the Study of Drug Development, PS 7812, January 1978.

[2] The FDA requires that these investigations be carried out under three carefully monitored phases: (I) laboratory pharmacology testing for safety, (II) limited clinical efficacy testing, and (III) larger clinical trials.

[3] Specifically, all devices whose purpose is to support or improve life or prevent suffering, or which are potentially risky--collectively designated as "Class III" devices--require "substantial evidence" of safety and efficacy before marketing approval may be granted.

[4] United States Congress. Federal food, drug, and cosmetic act, as amended. Government Printing Office, Washington, D.C., October 1976.

[5] Peltzman, S. The benefits and costs of new drug regulation. In: Regulating new drugs. Edited by R.L. Landau. University of Chicago Center for Policy Study. 1973: 113-211.

[6] Havighurst, C.C. Federal regulation of the health care delivery system. University of Toledo Law Review. 1975; 6: 577-590.

[7] Wardell, W.M. Rx: More regulation or better therapies? Regulation. 1979; 3: 25-33.

[8] Coleman, J.S., E. Katz, H. Menzel. Medical innovation: A diffusion study. New York: Bobbs-Merrill, 1966.

[9] Russell, L.B. The diffusion of hospital technologies: Some econometric evidence. Journal of Human Resources. 1977; XII: 482-501.

[10] Baruch, J. The diffusion of medical technology. Medical instrumentation. 1979; 13: 11-13.

[11] Fineberg, H.V. Gastric freezing--a study of diffusion of a medical innovation. In: Medical technology and the health care system: A study in the diffusion of equipment-embodied technology. Washington, D.C.: National Academy of Sciences, 1979; 173-200.

[12] Bauer, R.A., L.H. Wortzel. Doctor's choice: The physician and his sources of information about drugs. Journal of Marketing Research. 1966; III: 40-47.

[13] For example, the influence of journal reports on usage is magnified by an increasing quantity of data, which is linked to the number of procedures that have been evaluated and therefore to the number of procedures that have been performed.

[14] Favus, J.F., A.B. Schneider, M.E. Stachura, et al. Thyroid cancer occurring as a late consequence of head-and-neck irradiation: Evaluation of 1056 patients. New England Journal of Medicine. 1976; 294: 1019-1025.

[15] Committee on Technology and Health Care, Assembly of Engineering, National Research Council, and Institute of Medicine. The evaluation of equipment-embodied technology. Washington, D.C.: National Academy of Sciences, 1979.

[16] Proceedings of the workshop on percutaneous transluminal coronary angioplasty: June 15-16, 1979. Sponsored by Cardiac Diseases Branch, Division of Heart and Vascular Diseases, National Heart, Lung, and Blood Institute. Bethesda, MD: National Institute of Health, 1980.

[17] This definition is similar in flavor to the FDA's, which states that an effective technology is one that meets the claims made for it by the manufacturer. See [5].

[18] In the case of PTCA, some believe that three years might be an appropriate follow-up time because of the possibility of renewed or accelerated deterioration of the treated coronary arteries.

[19] The borderline level of reported effectiveness at which physicians find evaluative data neither encouraging nor discouraging is assumed to be 0.4. By year 5, reported effectiveness has declined to about 0.3, so that new data is discouraging on the whole.

[20] Fineberg, H.V., H.H. Hiatt. Evaluation of medical practices: The case for technology assessment. New England Journal of Medicine. 1979; 301: 1086-1091.

[21] Luft, H.S., J.P. Bunker, A.C. Enthoven. Should operations be regionalized? An empirical study of the relation between surgical volume and mortality. New England Journal of Medicine. 1979; 301: 1364-67.

[22] Block, H. Toward better systems of drug regulation. In: Regulating new drugs. Edited by R.L. Landau. University of Chicago Center for Policy Study. 1973: 243-265.

[23] Kraus, W.A., D.O. Davis. Utilization and cost-effectiveness of crenial computed tomography at a university hospital. Journal of Computer Assisted Tomography. 1978; 2: 209-214.

[24] Bendixen, H.H. The cost of intensive care. In: Costs, risks, and benefits of surgery. Edited by J.P. Bunker, B.A. Barnes, F. Mosteller. New York: Oxford University Press. 1977: 372-384.

[25] Gruntzig, A.R., A. Sennig, W.E. Siegenthaler. Nonoperative dilatation of coronary-artery stenosis: Percutaneous transluminal coronary angioplasty. New England Journal of Medicine. 1979; 301: 61-67.

[26] Some empirical evidence suggests that even when large controlled studies cast serious doubt on the effectiveness of an established practice, changes in use may be slow in coming. For example, see: Chalmers, T.C. The impact of controlled trials on the practice of medicine. Mt. Sinai Journal of Medicine. 1974; 41: 753-759.

[27] Blume, S. Aspects of the dynamics of medical technology. In: Research on health research. Edited by E. Heikinnen, H. Vuori, T. Laaksovirta, P. Rosenquist. Helsinki: Publications of the Academy of Finland. 1980: 195-213.

[28] von Hippel, E. The dominant role of users in the scientific instrument innovation process. Research Policy. 1976; 5: 212-239.

[29] In the case of a technology for which state-of-the-art improvements involve matters more related to the physician's technique than to the manufactured product, such as CABG surgery, modifications may be almost exclusively developed by practitioners.

[30] McLoughlin, W.G. Fundamentals of research management. New York: American Management Association, Inc. 1970.

[31] The initial positive level of experience is attributable to preclinical uses of the technology, such as animal studies.

[32] Lasagna, L. Current status of international drug regulation. University of Rochester: Center for the Study of Drug Development, PS 7704, June 1977.

[33] Since the passage of the 1962 Amendments, the time between a manufacturer's submission of an application to market a new drug and the application's approval has increased by over two years (from seven months to 34 months) on average. See [5] and [7].

[34] For this analysis, the following assumptions were made:
(1) Net benefit per procedure = ($14,000)((Effectiveness-0.4)/0.6), where 0.4 is the assumed borderline effectiveness, and $14,000 is the approximate difference in cost between PTCA and CABG surgery;
(2) A discount rate of 10% per year is applied to future benefits.
Under these assumptions, the cumulative discounted net benefit under regulation exceeds that under no regulation by $5.7 million in year 5, but is exceeded by $4.1 million in year 10 and by $14.8 million in year 20.
For a discussion of cost-benefit and cost-effectiveness analysis in medical policy-making, see: Weinstein, M.C., W.B. Stason. Foundation of cost-effectiveness analysis for health and medical practices. New England Journal of Medicine. 1977; 296: 716-721.

Appendix: Model Genesis and Documentation

The modeling study on which this paper was based began as a student project at MIT in the spring of 1980. This project (on which the author was a consultant) attempted to model the constraints on diffusion of a new medical technology (PTCA) in competition with an established regimen (CABG surgery). The author was employed by the National Heart, Lung, and Blood Institute (NHLBI) from June 1980 to June 1981, to expand the scope and purpose of the model so that the Institute could examine, in a broad sense, the effect that government policies (such as sponsoring large-scale clinical trials) might have on the emergence of a technology like PTCA. No direct changes of policy were expected to come as a result of this analysis, which was seen more as a speculative or first-step exploratory effort. Although the NHLBI was the official client, policies of the FDA and the Health Care Financing Administration (HFCA, in charge of Medicare reimbursement policy) were also examined. The author was free to develop the model as he saw fit, with the only requirements being that the policy levers be clearly specified, that the model variables in general be recognizable to the relevant policy-makers, and that the model be parameterized to represent PTCA. No NHLBI-funded extension of the PTCA study is planned at the present time.

CLASS II DOCUMENTATION STANDARDS
FOR SIMULATION MODELS

1. ACCESS TO MODEL:

Name of Model: EMEDT (Emerging Medical Technology)

Name and current address of the senior technical person responsible for the model's construction: Jack B. Homer, System Dynamics Group, MIT, Cambridge, Mass.

Who funded the model development? National Heart, Lung, and Blood Institute (NHLBI)

In what language is the program written? DYNAMO II

On what computer system is the model currently implemented? MIT's IBM 370/168

What is the maximum memory required to store and execute the program? 400K bytes (linked to IBM's CMS editor)

What is the length of time required for one typical run of the model? 3 minutes

Is there a detailed user's manual for the model? no

2. PURPOSE OF THE MODEL:

For what individual or institution was the model designed? Office of Program Planning and Evaluation of NHLBI

What were the basic variables included in the model?

Procedures and their Effectiveness, Practitioners and their Experience, Patient selection criteria, Prescribing M.D. fraction, Evaluation reports and data on effectiveness, Technical modification/improvement/capability

Over what time period is the model supposed to provide useful information on real world behavior?

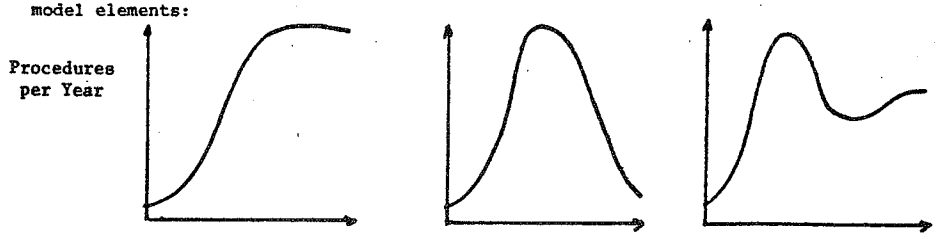
Up to 30 years, typically

Was the model intended to serve as the basis of:

an academic exercise designed to test the implications of a set of assumptions or to see if a specific theory would explain historical behavior	no
communication with others about the nature and implications of an important set of interactions	yes
projecting the general behavioral tendencies of the real system	yes
predicting the value of some system element(s) at some future point in time	no

3. MODEL SPECIFICATION AND THEORETICAL JUSTIFICATION:

Provide three diagrams illustrating the extreme behavior modes exhibited by the major model elements:



I. S-Shaped II. Rise-and-Fall III. Overshoot/ Undershoot

If they are not included in the body of the paper indicate where the reader may find:

a model boundary diagram that indicates the important endogenous, exogenous and excluded variables — MIT SD Group Memo D-3270-5

a causal influence diagram, a flow diagram, the computer program and definitions of the program elements — MIT SDG Memo D-3318 for Equation Description

Is the model composed of:

simultaneous equations _____

difference or differential equations x

procedural instructions _____

Is the model deterministic x or stochastic _____

continuous x or discrete _____

4. DATA ACQUISITION

What were the primary sources for the data and theories incorporated in the model?

Data NHLBI usage data, journal articles

Theory Physicians and evaluators at NHLBI, Manufacturer representatives, Medical diffusion literature, Technical innovation and development literature

What percent of the coefficients of the model were obtained from:

measurements of physical systems 0

inference from social survey data 20

econometric analyses 0

expert judgment 10

the analyst's intuition 70

What was the general quality of the data? Poor or non-existent

5. PARAMETER ESTIMATION

If they are not given in the publication, where may the reader obtain detailed information on the data transformations, statistical techniques, data acquisition procedures, and results of the tests of fit and significance used in building and analyzing the model? No formal techniques used; no write-up on this topic

6. MODEL PERFORMANCE AND TESTING

Over what period was the model's behavior compared with historical data? _____

3 years of historical data on PTCA

What other tests were employed to gauge the confidence deserved by the model? _____

Feedback from experts, comparison of output with previously reported behavior of other technologies

Where may the reader obtain a detailed discussion of the prediction errors and the dynamic properties of the model? For dynamic properties, see D-3270-5

7. APPLICATIONS

What other reports are based upon the model? Final report to NHLBI (D-3318)

Name any analysts outside the parent group that have implemented the model on another computer system. none

List any reports or publications that may have resulted from an evaluation of the model by an outside source. none

Has any decision maker responded to the recommendations derived from the model? no

Will there be any further modifications or documentation of the model? yes

Where may information on these be obtained? Jack Homer's doctoral dissertation will provide full model documentation. Expected completion date: May 1982