

2.

THE CONCEPTS OF EFFICACY AND SAFETY

2.

THE CONCEPTS OF EFFICACY AND SAFETY

The concepts of efficacy and safety have not been suddenly discovered or created. They have always existed in medical thought. In an intuitive sense, an efficacious and safe medical technology is one that “works” and causes no undue harm. That statement may sound naive to individuals working in the field of health today. However, for a major portion of the history of medicine, efficacy and safety were measured by that intuitive standard. Furthermore, that intuitive standard still lies at the heart of medical practice, but the meaning and measurement of those concepts have evolved with increased sophistication of scientific methods in medicine.

This chapter introduces the concepts of efficacy and safety. It begins with a brief discussion of the nature of efficacy and safety knowledge, presents the characteristics and concept of efficacy, then of safety, and finally, discusses efficacy and safety in relation to each another.

THE NATURE OF EFFICACY AND SAFETY KNOWLEDGE

Measurement of efficacy and safety is in essence an examination of interventions in the processes by which various phenomena affect health and disease. Neither these phenomena (whether they be biological, psychological, or social) nor the interventions (often, technologies) need be thought of as having a fully predictable mechanistic effect. A probabilistic view of effects—that is, when an event occurs, there is a range of possibilities that other events will occur—is more useful. The concept of probability is used to summarize the effects of causal variables which are unknown or not taken into account. Thus, we can speak of estimating or evaluating efficacy and safety, but not exactly determining them. Specific technologies have certain probabilities of effects; therefore, efficacy and safety information is normally expressed in terms of probabilities.

EFFICACY

There is no shortage of definitions for efficacy; nor is there a lack of confusion relating to distinctions between terms such as efficacy, effectiveness, benefit, and efficiency. Table 1 on the following page lists several definitions of efficacy.

Despite the sometimes substantial differences among the various interpretations of efficacy, one can isolate four critical factors that, taken together, form a comprehensive view of the concept.

Table I.—Selected Definitions of “Efficacy”

Source	Term defined	Definition	Relation to four factors (See below)
Federal Food, Drug, and Cosmetic Act (363)	Effectiveness, Efficacy (interchangeable)	A drug is effective if it has “the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof”	Benefit: Explicit Population affected: Implied Medical problem: Explicit Condition of use: Not included
A. Cochrane (72)	Efficacy (interchangeable with effectiveness)	“The effect of a particular medical action in altering the natural history of a particular disease for the better”	Benefit: Explicit Population affected: Not included Medical problem: Explicit Conditions of use: Not included
World Health Organization (435)	Efficacy	Benefit or utility to the individual of the service, treatment regimen, drug, preventive or control measure advocated or applied	Benefit: Explicit Population affected: Explicit Medical problem: Explicit Conditions of use: Not included
Discursive Dictionary of Health Care (347)	Efficacy (as a variant of effectiveness)	“The degree to which diagnostic, preventive, therapeutic, or other action or actions (undertaken under ideal circumstances) achieves the desired result”	Benefit: Explicit Population affected: Not included Medical problem: Not included Conditions of use: Explicit
Office of Technology Assessment, in this report	Efficacy	The probability of benefit to individuals in a defined population from a medical technology applied for a given medical problem under ideal conditions of use	Benefit: Explicit Population affected: Explicit Medical problem: Explicit Conditions of use: Explicit

The factors are:

1. Benefit to be achieved,
2. Medical problem giving rise to use of the technology,
3. Population affected, and
4. Conditions of use under which the technology is applied.

1. Benefit: The fact that a technology’s efficacy depends heavily on its benefit to the recipient seems a simple concept. Yet the question of what outcomes represent benefits is not so simply answered. Outcome criteria have usually been restricted to measurement of mortality and morbidity; less consideration has been given to life expectancy (longevity) or psychosocial and functional factors (40,41). The definition of benefit to be used will vary depending on the goals of the investigator and the type of technology being assessed.

A range of relevant outcomes can be considered in regard to a particular technology (227). A curative technology, for example, is efficacious only if it has a direct causal relationship to a positive patient outcome. In other cases, however, the consideration of intermediate criteria may be appropriate. For example, the benefit resulting from use of diagnostic technologies can be examined at five levels (116):

- 1) Technical capability—Does the device perform reliably and deliver accurate information?
- 2) Diagnostic accuracy—Does use of the device permit accurate diagnoses?
- 3) Diagnostic impact—Does use of the device replace other diagnostic procedures, including surgical exploration and biopsy?
- 4) Therapeutic impact—Do results obtained from the device affect planning and delivery of therapy?
- 5) *Patient* outcome—Does use of the device contribute to improved health of the patient?

If it is assumed that the function of a diagnostic technology, such as skull X-ray, is to perform accurate diagnoses of individuals' illnesses, the evaluation of benefit concentrates on the second level. If the diagnostic technology is expected to affect therapy or eventual patient outcome, then the fourth and fifth levels would be examined. Studies at the fourth and fifth levels may be difficult to conduct because long-term followup is required. As a result of this difficulty and the emphasis on diagnostic accuracy, evaluations in terms of therapeutic planning and patient outcome are infrequently performed.

The specification of benefit is often difficult for other classes of technologies as well. For example, is the efficacy of coronary bypass surgery to be evaluated in terms of its ability to give relief from symptoms (e. g., pain) or in terms of increased longevity for the patient? Thus, two different measures of benefit may possibly yield two different statements of efficacy for the same technology. This concept is illustrated by case study 8 on coronary bypass surgery in chapter 3.

2. Medical Problem: A technology's efficacy can be evaluated only in relation to the diseases or medical conditions for which it is applied. Obviously, one would not spend much time evaluating the efficacy of plaster cast applications for controlling hypertension. In general, however, the specification of medical problems is complex and can lead to controversy regarding the evaluation of the efficacy of a particular technology. For example, hysterectomies have been performed for a variety of medical conditions: pre-malignant states and localized cancers, descent or prolapse of the uterus, and obstetric catastrophes such as septic abortion (see chapter 3, case 11). They may also be performed as prophylaxis to avoid possible later cancer or pregnancy. If the efficacy of hysterectomy has been estimated for one of these diseases or medical conditions, it cannot be assumed automatically that the procedure will have similar efficacy for the others.

3. Population Affected: The effect of a medical technology varies depending on the individual treated. Sometimes, however, enough uniformity of effect exists to permit careful generalizations (163). These generalizations, or extrapolations, apply to the specific population type within which the original observations were made and should be supported by valid and reliable statistical techniques. For example, in the late 1960's the Veterans Administration (VA) conducted a multi-institutional controlled clinical trial of treatment for hypertension using the drugs hydrochlorothiazide, reserpine, and hydralazine (399) (see chapter 3, case 12). The treatment was shown to be efficacious for patients with diastolic blood pressure above 105 mm mercury. But, all the patients in the trial were males. Thus, the treatment could be considered to be efficacious (based on that trial and other evidence) for the population studied, males, but no automatic assumptions can be made concerning its efficacy for females.

Assumptions cannot be made because there are physiological and other differences among various population types. Children under certain ages, for example, may be affected by the same drug quite differently than adults. Therefore, the population undergoing treatment needs to be specified when the efficacy of a medical technology is discussed.

4. Conditions of Use: The outcome of the application of a medical technology is partially determined by the skills, knowledge, and abilities of physicians, nurses, and other health personnel, and by the quality of the drugs, equipment, institutional settings, and by support systems used by those personnel during the application. Cardiac surgery, as a commonly cited example, may result in a better outcome when conducted by skillful, well-trained surgeons who frequently perform such operations than when conducted by surgeons who rarely use that technology. Similarly, a drug's benefit may be greater if correct dosages are administered at the correct times. Also, the interaction of a drug with other drugs may affect the benefit. A situation where the physician is skillful and experienced, medication is administered carefully, and the patient receives the best care possible must be described as ideal. By definition, not all physicians are the most skillful, and not all conditions of use are of the highest possible quality. Average conditions of use inherently contain a great many variables, such as physician skill, that may differ from one hospital to another, and from one application of a technology to another. Thus, it is valuable to have an outcome measure that is not dependent on the differing variables inherent in average conditions of use. Efficacy is this measure. By defining efficacy as benefit under *ideal* conditions of use, a reasonably consistent measure for that factor is introduced. No conditions of use are absolutely ideal, but, for most purposes, carefully controlled research settings can serve as a substitute for ideal circumstances. These carefully controlled situations are frequently found in research hospital settings. For example, the efficacy of ambulatory maternal care can be studied in clinics, home situations, or hospitals. The essential criterion is "best possible control of conditions."

When the four factors described above are specified for the application of a specific medical technology, a relatively comprehensive statement has been made as to that technology's efficacy. Because a definition is merely a description of the properties of an entity, these four variables or factors can serve to define the concept of efficacy. This report uses the following definition of efficacy, not because it is necessarily more "correct" than others, but because it can be useful for discussion. It explicitly declares several key variables that, together, describe the potential usefulness of a medical technology.

Efficacy: The probability of benefit to individuals in a defined population from a medical technology applied for a given medical problem under ideal conditions of use.

This report differentiates efficacy from effectiveness. Effectiveness is concerned with the benefit of a technology under average conditions of use. An effective technology has positive benefits for those people who are treated with the technology in a typical medical setting. Although the efficacy of a drug, for example, may be evaluated for individuals in a research setting, its effectiveness in an average setting may be influenced by variables such as those mentioned above. These variables, such as proper administration of a drug, are more rigorously controlled in a research setting. Thus, the efficacy and the effectiveness of a drug may differ.

Though they can be viewed as distinct, efficacy and effectiveness are closely related concepts. The effectiveness of a technology is estimated by methods similar to those used to estimate its efficacy; however, estimating effectiveness is often more difficult because

of the absence of rigorously controlled settings. To be desirable, of course, a medical technology should be both efficacious and effective (as is, for example, polio vaccine). A medical technology can be efficacious and of limited effectiveness (e. g., a technology that benefits individuals but can be applied by only a few highly trained physicians). But a technology that is not efficacious cannot be effective.

SAFETY

Safety, like efficacy, is a relative concept: no technology is ever completely safe, or completely efficacious. In the beginning of this chapter, a safe technology was described intuitively as one that “causes no undue harm.” Despite the apparent simplicity of that informal definition, it reflects a critical property of the concept of safety: that safety represents a value judgment of the acceptability of risk. Risk can be thought of as “a measure of the probability and severity of harm to human health” (218). This definition of risk implies that investigators and policymakers should be concerned with both the nature of the risk and the probability of its occurrence. For example, a low but measurable probability of death can be more significant than a high probability of experiencing pain, discomfort, or other minor impairments.

Thus, if the risks of using a medical technology are acceptable (to the patient, physician, society, or other appropriate decisionmaker), the technology may be considered “safe” in that instance. Safety can then be defined as a judgment of the acceptability of the risk associated with a medical technology (90). That definition is useful to organizations, such as the Food and Drug Administration (FDA), which need both to consider the risk of a technology, such as a drug, and to decide whether and under what circumstances that risk may be considered acceptable. If FDA decides that a technology has certain risks which are likely to be acceptable to a sufficient number of decisionmakers, and if it is efficacious, the agency will approve that technology for marketing.

As with efficacy, several factors must be specified when risk and safety are discussed. The *medical problem* for which the technology being evaluated is applied must be specified, not only because the medical problem or condition of the patient will often affect the action of the technology and thus the associated risks, but also because the judgment of acceptable risk depends on the type and severity of the medical problem. For example, technologies used to treat Hodgkin’s disease* have types of risks that can at times be severe, although the probability of their occurring may be relatively low. A second malignancy may develop as a result of using radiotherapy and chemotherapy. Also, treatment may cause bone marrow suppression, pneumonitis, or several other deleterious effects (273). These risks, however, must be compared to the benefits of a normal life span, which is very often the direct result of treatment. Given these alternatives, the patients may regard the treatment as acceptably safe; that is, the risks are acceptable under the specific circumstances.

The *population affected* is also an important factor to be specified for reasons similar to those given regarding efficacy. For example, persons above a certain age or below a certain age may be especially susceptible to undesirable side effects of a drug, or they may be less able than most adults to withstand the rigors of a prolonged surgical procedure. Thus, the risks to those persons would be greater and more severe.

● Hodgkin’s disease is a form of cancer that affects the lymphatic system.

The risk associated with a particular medical technology also depends on the *conditions of use* under which the technology is applied. The reasoning for the inclusion of this factor parallels that presented in the previous section, *Efficacy*.

For the purposes of this report, then, risk may be defined as follows:

Risk: A measure of the probability of an adverse or untoward outcome occurring and the severity of the resultant harm to health of individuals in a defined population associated with use of a medical technology applied for a given medical problem under specified conditions of use.

This definition covers risk under ideal (research) settings, under average or typical settings, and under conditions where quality is below average. This coverage is afforded by the specification of “conditions of use.” Normally, when “efficacy and safety” judgments are being discussed, risk is assumed to be measured under ideal conditions of use.

Given this definition of risk, safety can be specified.

Safety: A judgment of the acceptability of risk in a specified situation.

EFFICACY AND SAFETY

Efficacy and safety are separate concepts; they can be measured and discussed as distinct properties of a medical technology. Efficacy is defined in terms of a benefit; safety, in terms of a risk. There are, though, many similarities between the two concepts. Neither efficacy nor safety is absolute. Both are discussed in terms of *probability* and *magnitude* of benefit or harm. Also, both are specified by several common factors: medical problem, population affected, and conditions of use. Most importantly, however, each can be fully evaluated only in terms of the other. A technology may provide benefits, but the *value* of those benefits depends on the risks involved in using the technology.

The controversy surrounding the use of mammography illustrates the interdependency of these concepts (see chapter 3, case 4). The benefits of reduced or delayed mortality due to using mammography for detection of breast cancer must be balanced against the risk of developing cancer from radiation emitted by the mammography device. The benefits and the risks are estimated separately, but the value of the technology depends on a comparison of the two estimates. In the case of mammography, for example, the Breast Cancer Screening Consensus Development Panel, assembled by the National Institutes of Health, “found no convincing justification for routine mammographic screening for women under the age of 50” (385). Efficacy and safety evaluation, then, is one specialized form of benefit-risk analysis.

Although efficacy assessments and safety assessments are for the most part symmetrical, at least four factors differ:

1. Ranges of effects,
2. Number of people affected,
3. Whether effects are known or expected, and
4. Time period of effects.

1. **Range of Effects:** In assessing efficacy, a limited number of specific benefits are usually sought. A certain drug, for example, may be tested for its ability to reduce blood pressure to a safe level in hypertensive individuals. The researcher often does not expect that drug to cure or ameliorate other disease conditions. Assessing safety, however, involves consideration of the broadest range of risks that can be assessed within practical limitations.

2. **Number of People Affected:** An efficacious medical technology results in benefits for patients with a given medical condition. Preferably, the technology will benefit a high proportion of people having the condition. Measurement and assessment of risk, however, consider the negative health effects of a technology for even a small proportion of patients. For example, although a technology may be beneficial for many patients, FDA may judge it unsafe if only a small proportion of those benefiting suffer significant, unacceptable, negative effects. The trade-off between benefits and risks will depend on the perceived magnitude and value of both benefits and risks for the proportions of people affected in each case.

3. **Known or Expected Benefits vs. Unknown or Unexpected Risks:** When the efficacy of a new technology is tested, a specific type of benefit is, in general, expected. Other benefits are usually ancillary to the outcome sought. In assessing risk, however, the negative outcomes are often unknown or unexpected. And, unlike the ancillary benefits, the significance of these effects must be considered to the extent practicable before a technology is deemed of acceptable risk. When thalidomide was tested as a sleeping pill, no major negative effects were discovered. Its effects upon the fetus were not tested, and thalidomide was marketed as a safe drug. The birth defects that resulted vividly demonstrate the need to consider risks from many perspectives. However, even with extensive examination of possible risks, we cannot expect absolute safety.

4. **Time Period of Effects:** The benefits derived from the use of a technology often may be observed sooner than the adverse effects. The time difference in the manifestation of adverse and beneficial effects is particularly characteristic of therapeutic technologies. Some deleterious effects, such as surgical complications or certain adverse drug reactions, can be observed almost immediately; others may not occur for years after the treatment. In some cases, the offspring may suffer more harm than the patient.