Users’ Guides to the Medical Literature

XIV. How to Decide on the Applicability of Clinical Trial Results to Your Patient

Antonio L. Dans, MD; Leonila F. Dans, MD; Gordon H. Guyatt, MD, MSc; Scott Richardson, MD; for the Evidence-Based Medicine Working Group

**CLINICAL SCENARIO**

You are the attending physician on duty when a poor, 45-year-old man presents to the emergency department of a general hospital in the Philippines. He has severe chest pain for 2 hours, associated with clammy perspiration. Physical examination reveals a blood pressure of 110/70 mm Hg, a pulse rate of 92 beats per minute, a normal first heart sound, and clear lungs. An electrocardiogram discloses 3-mm ST-segment elevation in the inferior leads. As intravenous lines are placed, and the patient is prepared for admission to the coronary care department, you consider whether you should offer this patient a thrombolytic agent. Though your response is that the impoecious patient cannot afford the treatment, you ponder the right course of action in a richer patient. As your duty ends that night, you resolve to prepare for the next patient admitted for an acute myocardial infarction (MI) by retrieving the best evidence on the use of thrombolytics.

**THE SEARCH**

Streptokinase is the only thrombolytic agent that your patients might afford. You, therefore, confine your search to this drug, trying to locate the best trial or, if possible, a meta-analysis. Using Grateful Med software (National Library of Medicine, Bethesda, Md), you select myocardial infarction from the list of medical subject headings used to index articles. On the second subject line, you use the term streptokinase. You limit your search to English-language articles, and to find quantitative reviews or original studies, you use the term meta-analysis or randomized controlled trial as the publication type.

You retrieve a systematic meta-analysis of randomized trials that deal only with effectiveness and not toxicity. You, therefore, also review a single trial from ISIS-2 Collaborative Group that you choose on the basis of its size (17 000 patients), strong design (including double-blinding), and the wide variety of settings in which the study was undertaken. You refer to earlier Users’ Guides to evaluate the validity of the studies, as well as the magnitude and precision of the treatment effects and toxicity.

The articles pass the validity criteria, and the treatment reduced the event rate from 17.4% to 12.8%. This outweighs the potential harm of “bleeds requiring transfusion,” which occurred in 0.5% of patients treated with streptokinase compared with 0.2% in the placebo group.

An answer does not come easily to the last question: “How can you apply the results to your patients?” Asians constituted a small minority of the patients in the trials, and you are uncertain about your hospital staff’s ability to cope with technical requirements for administering the drug or dealing with any complications.

As clinicians look more often to randomized controlled trials (RCTs) to guide their clinical care, they must decide how to apply RCT results to individual patients in their practice setting. This Users’ Guide addresses the issue of applicability, which involves the implications of the trial results for patient care. Applicability is closely related to concepts of generalizability and external validity, but is broader in its scope, including issues related to the overall impact of treatment in individual patients. In considering applicability, clinicians first must decide whether the biology of the treatment effect will be similar in patients they are facing; second, their patients’ risk of a target event, which the treatment is designed to prevent; third, the adverse effects that may accompany treatment; and fourth, their own ability to deliver the intervention in a safe and effective manner.

Clinicians managing patients who differ economically, racially, and culturally from those recruited in typical clinical trials face particular challenges in addressing applicability. Such patients include those from the inner cities of North America, the Native American reservations, or less industrialized countries. Clinicians seeing these patients cannot afford to repeat every trial simply because of doubts regarding applicability. The end result is that applicability becomes a fait accompli—an issue that may often be ignored rather than confronted.

Earlier in this series, we addressed the applicability problem in the Users’ Guide for articles about therapy or prevention: “A better approach than rigidly applying the study’s inclusion and exclusion criteria is to ask whether there are compelling reasons why the results should not be applied to the patient. A compelling reason usually won’t be found, and most often you can generalize the results to your patient with confidence.”

---

From the Departments of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario (Dr Guyatt); the Departments of Internal Medicine (Dr A. L. Dans) and Pediatrics (Dr L. F. Dans), University of the Philippines College of Medicine, Manila; and the Department of Medicine, University of Texas Health Sciences Center, San Antonio (Dr Richardson).

The original list of members (with affiliations) appears in the first article of this series (JAMA. 1993;270:2093-2095). A list of new members appears in the 10th article of this series (JAMA. 1993;270:2093-2095). The following members contributed to this article: Deborah J. Cook, MD, MSc; Hertz Gerstein, MD, MSc; Ann Holbrook, MD, PharmD, MSc; Lex Irwig MBBS, PhD; Virginia Moyer, MD, MPH; and Thomas B. Newman, MD.

Reprints: Gordon H. Guyatt, MD, MSc, McMaster University Health Sciences Centre, 1200 Main St W, Room 2C12, Hamilton, Ontario, Canada L8N 3Z5.

JAMA, February 18, 1998—Vol 279, No. 7

©1998 American Medical Association. All rights reserved.

Downloaded from www.jama.com at FAY TOWELL MEDICAL LIB, on December 10, 2007
Physicians may encounter problems following this advice. We didn’t give a good definition of a “compelling reason” or provide guidelines on how to systematically address the question. In this article, we correct these deficiencies by presenting a set of guidelines for evaluating the applicability of the results of RCTs to populations other than the participants. We present the guides as questions that probe for situations when clinicians may be forced to reject applicability. We phrase the questions so that a “yes” answer will lead clinicians to suspect a problem of applicability. Table 1 summarizes the guides, categorizing them into biologic issues (which help us decide if the treatment can work), socioeconomic issues (which help us decide if the treatment will work), and epidemiologic issues (which help us decide if the treatment will be). As we discuss each issue, we will offer sources of information that will help physicians answer their questions.

**THE GUIDES—BIOLOGIC ISSUES**

### Are There Pathophysiologic Differences in the Illness Under Study That May Lead to a Diminished Treatment Response?

Diseases with a single name may represent conditions with important pathophysiologic differences. These differences can sometimes lead to diminished treatment responses due to divergence in pathogenetic mechanisms or biological differences in the causative agent. Hypertension in blacks, which has been observed to be relatively responsive to diuretics and unresponsive to β-blockers, provides an example of the former. This selective response reflects a state of relative volume excess that investi- gated now theorize may have served protective functions in their hot and arid ancestral environments. Malaria provides an example of a condition that may vary because of biological differences in the causative agent. Malaria treatment protocols vary depending on drug resistance patterns. In these examples, clinicians should anticipate variation in response to treatment and should temper hasty conclusions regarding the applicability of trial results.

**Sources of Evidence**

Sources of information regarding disease pathophysiology in populations include basic and laboratory studies, animal studies, genetic studies, and observational studies documenting pathologic changes in affected individuals and evaluating the biology of causative agents (eg, surveys on drug resistance patterns of infectious diseases). In some cases, variation in response to treatment may be the first clue to a difference in pathophysiology. This was the case in the example of hypertension in blacks.

To address our scenario of applicability of streptokinase to the treatment of MI in the Philippines, we reviewed a case series of autopsies performed on Filipino patients who had MI. Pathologic changes in the coronary arteries and myocardium were similar to those noted among North Americans, while nonatherosclerotic causes of coronary disease were rare. Clinical surveys have demonstrated that Filipinos share the same risk factors for coronary disease as North Americans. Thus, we can be confident that disease pathogenesis is similar.

### Are There Patient Differences That May Diminish the Treatment Response?

Between-population differences in response to treatment may arise from differences in drug metabolism, immune response, or environmental factors that affect drug toxicity. Differences in drug metabolism may directly influence the efficacy of a treatment regimen. If they are not identified, slow metabolizers of a drug could face the risk of greater toxic effects, while a significant decrease in efficacy might occur in rapid metabolizers. Such differences are usually based on genetic polymorphism in the activity of metabolizing enzymes. A well-known example is hepatic N-acetyltransferase, an enzyme with increased activity among Asians. For this reason, clinicians offer higher drug dosages for agents such as isoniazid, hydralazine, and procainamide hydrochloride. Other examples of genetic polymorphism include pseudocholinesterase activity in the metabolism of suxamethonium and glucose-6-phosphate dehydrogenase activity in the metabolism of sulfonamides and other drugs.

Differences in patients’ immune response may also modulate treatment effect. Haemophilus influenzae vaccine, for example, has a lower efficacy in Alaskan natives than in nonnative populations. Finally, environmental factors may affect response to therapy. For instance, the incidence of thyroid dysfunction from amiodarone differs in low vs high iodine environments. Pharmacokinetic and bioavailability studies are important sources of evidence regarding differences in treatment response. Such studies generally require small sample sizes and commonly available equipment. Unfortunately, for a wide variety of drugs, technology for assays remains unavailable. Reasonable alternatives include dose-ranging and descriptive studies of patients receiving treatment, which can also provide information on immune response to vaccines and environmental factors that may increase or decrease the toxic effects of drugs. Postmarketing surveillance studies and large RCTs require large sample sizes and long-term follow-up, but (as in the example of the decreased effect of H influenzae vaccine in Alaskan natives) may provide definitive information about differential response to therapy.

Although we found no studies evaluating the pharmacokinetic profile of streptokinase when given to Filipinos, postmarketing studies show that Filipinos experience the same reperfusion arrhythmias and bleeding complications when given streptokinase at the same dose as North Americans. These studies provide some assurance of similarities in the response to adverse effects of treatment.

**SOCIAL AND ECONOMIC ISSUES**

When satisfied that biologic differences do not compromise treatment applicability, clinicians must examine constraints related to the social environment that may diminish treatment effectiveness.

### Are There Important Differences in Patient Compliance That May Diminish the Treatment Response?

To the extent that groups of people exhibit different compliance with treatment, clinicians may expect variation in treatment effectiveness. Variability in compliance between populations may stem from resource limitations in a particular setting or less obvious attitudinal or behavioral idiosyncrasies. Both types of problems may, for example, affect the safety of outpatient administration of anticoagulants. Neither indigent pa...
tients or their society may be able to afford repeated clinic visits and tests for treatment monitoring. Alcoholic patients, whatever their financial situation, may be less likely to comply with monitoring. Inadequate monitoring, whatever the reason, increases bleeding risk from overanticoagulation, shifting the balance between benefit and harm (even to the point where harm outweighs benefit).

Sources of Evidence

While clinicians perform poorly at untutored guessing of patient compliance, a systematic examination of compliance in individual patients, or groups of patients, is likely to aid in identifying varying compliance patterns. Clinicians may also refer to more general sources of evidence, such as sociologic descriptions of attitudes of specific groups of people. In the Philippines, an attitude called ba-hala na connotes a lack of capacity or will to control one’s fate. A near equivalent would go something like “let’s just wait and see, there’s really nothing much we can do about the situation.” This external locus of control may have an adverse effect on patient compliance. In our scenario, we don’t expect patient compliance to be a problem since we give streptokinase intravenously as a single dose.

Are There Important Differences in Provider Compliance That Might Diminish the Safety and Efficacy of the Treatment?

In this guide, provider compliance refers to a host of diagnostic tests, monitoring equipment, intervention requirements, and other technical specifications that clinicians must satisfy to safely and effectively administer a treatment. Financial conditions in a health care center, access to equipment, technological expertise, and availability and skill of health personnel may influence treatment effectiveness. For instance, while carotid endarterectomy may benefit low-risk patients when surgery-associated stroke is low, the net effect for such patients in centers with higher surgery-associated stroke rates may be an increase in adverse outcomes.

In less industrialized countries, many hospitals and clinics do not have easy access to sophisticated equipment, so problems of provider compliance are common. For example, while rheumatic atrial fibrillation remains a common problem in Asian countries, few laboratories in rural areas perform the tests necessary for titration of warfarin dose. This limitation is likely to reverse the critical balance between effectiveness and safety of treatment.

Sources of Evidence

Because of experience regarding availability of equipment, laboratory tests, and health personnel resources, practitioners themselves are a good source of information regarding feasibility interventions. Clinicians’ assessments can be supplemented by formal quality-of-care assessments and postmarketing surveillance of adverse effects. Whatever the source of information, a thorough understanding of the technical requirements for safe and effective administration should guide decisions regarding the ability to comply.

Administration of streptokinase carries potential hazards, foremost of which is catastrophic bleeding. Facilities for emergency administration of cryoprecipitate, fresh frozen plasma, or whole blood must be available. In hospitals without efficient blood banking systems, it may be difficult to cope with bleeding emergencies. This increases the potential hazards of treatment and may tip the balance between benefit and harm.

EPIDEMIOLOGIC ISSUES

When satisfied that biologic, social, or economic differences do not compromise applicability, the clinician must examine the patient’s characteristics that can influence either the magnitude of the benefit or the risks of treatment (and thus, the trade-off between the 2). The last 2 guides address these issues.

Do My Patients Have Comorbid Conditions That Significantly Alter the Potential Benefits and Risks of the Treatment?

The presence of other conditions in a particular locality may affect treatment efficiency in 2 possible ways: competing diagnostic possibilities or competing causes of outcome. The management of pneumonia in developing countries provides an example of a competing diagnostic possibility.

The acute respiratory tract infection management protocol includes a symptom-driven algorithm for differentiating pneumonia from nonpneumonia. This protocol identifies children who need antibiotics and has proven effective in reducing mortality from pneumonia among children younger than 5 years. Unfortunately, similarities exist in the clinical presentation of pneumonia and malaria. In malaria endemic areas, clinicians may expect an increase in false-positive “pneumonias.” These patients with false-positive diagnoses and their parents will not respond to antibiotics for pneumonia, and a delay in instituting antimalarial treatment may result. If the drop in accuracy is large enough, the balance between harm and benefit will change. To resolve this issue, investigators have initiated a study to determine if the acute respiratory tract infection protocol can maintain its effectiveness in malaria endemic areas (S. P. Lupisan, unpublished data, 1998).

Competing causes of target events may also affect the magnitude of benefit. An example comes from the management of acute MI in some Filipino hospitals. A recent study disclosed 30 in-hospital deaths in a cohort of 149 patients admitted to a charity hospital (ISIP Study Group, unpublished data, 1996). On the basis of results from the meta-analysis, clinicians might expect streptokinase to reduce this 20% death rate by 25%. However, a closer look at the local data shows a contrast with the original studies in which virtually all deaths were a direct result of cardiac ischemia. In the Philippine study, noncardiac causes (mostly pneumonia with sepsis) were responsible for 11 of the 30 deaths. Streptokinase will not reduce mortality in such patients. Adequate antibiotic coverage may result in a greater (and more economical) reduction in mortality for patients who develop pneumonia.

In addition to reducing benefit, other morbidity may affect the magnitude of risk. Surgical mortality may increase in malnourished patients, shifting the balance between benefit and risk. On occasion, other morbidity can also work in the opposite direction—increasing efficiency. For example, a patient with a large infant, in whom the clinician is considering warfarin, may also have atrial fibrillation. Since anticoagulation reduces stroke risk in such patients, the presence of atrial fibrillation strengthens the indication or treatment.

Sources of Evidence

Cohort studies provide the most reliable information on comorbid conditions. In the MI scenario, we used data from the local study of 149 patients to evaluate the impact of other morbid conditions. As we noted, we can expect streptokinase to prevent around 5 of 19 cardiac deaths (but none of those from other causes), and the absolute reduction in all-cause mortality is a decline from 30 (20.1%) of 149 to 25 (16.8%) of 149.

Are There Important Differences in Untreated Patients’ Risk of Adverse Outcomes That Might Alter the Efficiency of Treatment?

In our Users’ Guide on therapy, we addressed the relationship between a patient’s risk of an adverse event and the magnitude of the treatment impact. Because the issue is so important in assessing applicability of trial results, we will review it in detail.
In the therapy Users’ Guide, we introduced the notion of number needed to treat (NNT). Thinking about NNT requires an understanding of the concepts of relative risk, relative risk reduction, and absolute risk reduction. Readers desiring a full discussion of these concepts can refer to the earlier article. But because it estimates the number of patients who need to receive treatment (with implications about the associated toxic effects and cost) to prevent an adverse event, clinicians can use the NNT to consider a treatment’s efficiency.

The NNT is the inverse of the absolute risk reduction resulting from a particular treatment in a particular group of patients. If a patient’s risk without treatment is 20%, then we expect 20 of 100 patients without treatment to experience an adverse event. When we administer a treatment with a relative risk reduction of 10%, only 18 treated patients will experience adverse events. Thus, for every 100 patients treated, we prevent 2 events, and the NNT is 50. If the expected event rate in untreated patients is cut by half, the NNT is 100 divided by 2.8 or approximately 36 patients. This reasoning, and much of what follows, assumes that relative risk reduction remains constant across subgroups. While testing this assumption can be difficult, there are situations in which the assumption will fail, and clinicians should be alert to this possibility. Fortunately, however, in most instances the assumption will not introduce important inaccuracies in the NNT.

One source of difference in expected event rates is country of origin and residence. Keys compared the 20-year incidence of coronary deaths in the United States, 5 European countries, and Japan. He found an extremely low incidence of coronary death in the Japanese cohort, despite correction for baseline differences in recognized risk factors. Similar results have been observed in preliminary reports of the ongoing Multinational Monitoring of Cardiovascular Disease and Their Determinants project. In this study, involving 39 centers from 26 countries, east Asians showed a much lower incidence of coronary death than their western counterparts. Age-standardized mortality rates for coronary heart disease were lowest among Japanese (40 of 100,000), and highest in North Ireland (414 of 100,000).

Thinking of the NNT, this 10-fold difference in incidence among the Japanese would translate to a 10-fold increase in the NNT for a drug preventing coronary deaths. This decrease in efficiency may warrant a reconsideration of applying the results of a trial to low-risk patients. We consider the issue of balancing costs and effects in our Users’ Guides for determining a level of recommendation.

**Sources of Evidence**

Cohort studies on the course of disease in untreated patients can provide excellent risk data, and such studies are even more useful when they define subsets of patients at varying risk. Of 424 Filipinos with MI who were eligible for streptokinase (but in whom the drug was not administered) and who participated in a cohort study conducted in 9 centers in metropolitan Manila, 37 (11.1%) suffered cardiac death. This provides a good estimate of the expected event rate. If streptokinase had been given, it would have prevented 25% of the deaths, reducing the absolute mortality rate to 8.3%. Thus, 2.8% of the those otherwise destined to die would have been spared (the absolute risk reduction), and the NNT is 100 divided by 2.8 or approximately 36 patients.

The expected event rates varied in the patient population. Young patients with small infarcts had a much lower expected mortality (and thus much larger NNTs) than old patients with large infarcts. Using prognostic information from these various subgroups, we constructed Table 2, which shows the expected mortality according to age and left ventricular wall involvement and the corresponding NNT to save 1 life in each group. As the table shows, NNT can range from 16 (when treatment is applied to patients with a poor prognosis) to as much as 179 (when treatment is applied to patients with a good prognosis).

**Table 2**—Baseline Mortality Rate Without Treatment and Estimated Number Needed to Treat or to Save 1 Life Using Streptokinase in Filipinos With Acute Myocardial Infarction, Tabulated According to Age and Wall Involvement

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Age $&lt;$ 60 y</th>
<th>Age $\geq$ 60 y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mortality Rate</td>
<td>No. Needed to Treat</td>
</tr>
<tr>
<td>Wall involvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infarction</td>
<td>0.02</td>
<td>179</td>
</tr>
<tr>
<td>Non-Q-wave myocardial</td>
<td>0.04</td>
<td>89</td>
</tr>
<tr>
<td>Anterolateral wall infarct</td>
<td>0.05</td>
<td>71</td>
</tr>
<tr>
<td>Massive anterior wall infarct</td>
<td>0.14</td>
<td>20</td>
</tr>
</tbody>
</table>

Varying patient risk will affect benefit of treatment no matter what the environment in which you practice. Even if you work in the Western tertiary care environment in which investigators conducted their original studies, you will still face high- and low-risk patients. The critical trade-off between risk and benefit may vary in these patient groups, mandating a different treatment decision.

**COMMENT**

These guides address the task of applying the results of clinical trials done on restricted, specially selected populations to other groups. Although inspired by the predicament in less industrialized countries, the guides are relevant to all situations where clinicians must make decisions regarding applicability. By breaking down the problem into specific questions, we have provided guides for busy clinicians who make daily attempts to strike a balance between making “unjustifiably broad generalizations and being too conservative in one’s conclusions.”

When clinicians suspect limited applicability (ie, when a response of “yes” is encountered for any of the questions), what can they do? This will depend on whether the anticipated differences are important, and if important, whether they are remediable. For example, differences in disease pathophysiology (guide 1) do not always mean that applicability is limited. Management of a cataract, for instance, will probably be the same regardless of the cause. Differences in treatment response (guide 2) can sometimes be accommodated by altering administration of a treatment (such as adjusting the dose of a drug). Education, training, provision of necessary equipment, and other attempts at optimizing compliance may address problems in patient and provider compliance (guides 3 and 4).

For differences in comorbid conditions or expected target event rates (guides 5 and 6) the clinician’s response will depend on the differences observed. If an increase in efficiency is anticipated (as when disease prognosis is worse or the incidence of an adverse outcome is greater), a recommendation to treat can be more easily accepted. A decrease in efficiency, on the other hand, should lead clinicians to be more cautious in accepting a treatment recommendation.

When the answer to 1 or more of the guide questions is “yes,” and the differences noted are important and not easily remediable, clinicians should not assume that the trial results can be readily applied. In these instances, an additional RCT may be warranted.
RESOLUTION OF THE SCENARIO

What should we recommend regarding thrombolytic use for the Filipino patient admitted for acute MI? There is no reason to believe that Filipinos have a different disease pathogenesis or a different response to treatment with thrombolytics (guides 1 and 2). Patient compliance will not be an important issue since the drug is given intravenously as a single dose (guide 3). The technical requirements for administration are often, but not always, available, and when they are not, the risks may outweigh the benefits of thrombolytic administration (guide 4).

Two issues remain to be resolved, both dealing with the magnitude of treatment impact. Pneumonia is an important comorbid condition, accounting for one third of deaths, at least in some charity hospitals (guide 5). However, rates of cardiac death are still sufficiently high (11.1%) that the relative risk reduction we can achieve with streptokinase (28%) will result in an NNT of 32 for the overall population (guide 6). For subgroups of patients, however, the NNT will range from 16 to 179, depending on the age and the size of the infarct (Table 2).

Should we recommend the routine use of streptokinase among Filipinos presenting with acute MI? The guides have brought us closer to an answer. We have confirmed applicability of the thrombolytic data on the effectiveness of streptokinase, but only in centers with adequate blood banking facilities. We have also refined estimates of treatment impact, based on knowledge of the course of disease among Filipinos. However, the cost of the drug is approximately $250 per treatment and in the Philippines the average annual per capita income is only $1300 (National Statistics Office, unpublished data, 1994). These figures highlight the difficult economic trade-off associated with administering streptokinase.

The judgment about whether to give streptokinase will depend on who pays for the treatment (in the Philippines, usually the patients themselves), patient and family values, what resources are available (usually limited in our charity hospital setting), and the competing needs (for example, the need for antibiotics because of a high incidence of pneumonia, in turn a result of overcrowding in the hospital wards). For equally applicable treatments, our final decision may differ for a much less costly, but equally effective and applicable treatment, such as aspirin for our MI patient.

References
5. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? JAMA. 1990;263:2998-2991.
6. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, B: what were the results and will they help me in caring for my patients? JAMA. 1994;271:69.