The Need for Perspective in Evidence-Based Medicine

Steven H. Woolf


http://jama.ama-assn.org/cgi/content/full/282/24/2358

Correction Contact me if this article is corrected.
Citations This article has been cited 48 times. Contact me when this article is cited.
Topic collections Medical Practice; Health Policy; Quality of Care; Evidence-Based Medicine Contact me when new articles are published in these topic areas.
Related Letters Risk Factors for Violent Death in Children
The Need for Perspective in Evidence-Based Medicine

Steven H. Woolf, MD, MPH

Aproximately 400,000 Americans die each year because of tobacco use. This dramatic statistic prompts questions about other risk factors: How many deaths are due to physical inactivity, elevated serum lipid levels, hypertension, and obesity? Knowing the answers would help gauge the relative impact of different preventive strategies. For instance, if lowering saturated fat and alcohol intake can prevent breast cancer, would doing so save more lives than secondary prevention (eg, mammography screening) or treatment of existing disease (eg, adjuvant tamoxifen)? Nearly 900,000 Americans die each year of heart disease and stroke. Nearly 900,000 Americans die each year of heart disease and stroke. Nearly 900,000 Americans die each year of heart disease and stroke. Nearly 900,000 Americans die each year of heart disease and stroke. Nearly 900,000 Americans die each year of heart disease and stroke.

β-Blockers, aspirin, thrombolysis, angiotensin-converting enzyme inhibitors, and warfarin can reduce mortality rates in these patients by 21% to 33%. Are efforts to properly administer these drugs more or less likely to prevent cardiovascular deaths than getting smokers to quit or controlling hypertension?

This article discusses the importance of systematically comparing the relative effectiveness of interventions in preventing disease outcomes. An information repository for such comparisons is lacking in the United States, and decision makers at the population, clinical, and individual levels often must choose between interventions for a disease without knowing which works best. This article proposes a national program to regularly evaluate and report this information.

Priorities in Decision Making

Some might argue that determining which interventions work best is unnecessary because all effective interventions should be implemented. However, limited resources force clinicians, policy makers, and patients to emphasize some interventions more than others. A mismatch between what is emphasized and what works best necessarily compromises health outcomes. However, it is difficult to identify misplaced priorities without a reference standard that indicates how interventions would be prioritized if health were the only concern.

Optimizing health is surely not the only concern in making choices; costs, feasibility, and context also matter. Which priorities prevail depends on who is making the decision. Payers and politicians prefer services that are cost-effective and popular. Clinicians respond to their patients’ chief complaints. Patients pursue tests and treatments promoted by family, news media, and advertisements. It is fair to set priorities for these reasons, but it is unwise to do so without knowing whether (and by how much) health out-

Research advances are generating a growing body of clinical trial and other data on the effects of tests and treatments on outcomes, but there is no information resource within the health care system that systematically puts that information in perspective. Policy makers, clinicians, and individuals lack a ready means to compare the relative effectiveness of various interventions in prolonging survival or preventing the occurrence or complications of a disease: information that is critical in setting priorities. A crude analysis of preventable deaths suggests that evidence-based primary prevention (getting the population to stop smoking, exercise, lower cholesterol levels, and control blood pressure) would prevent considerably more deaths per year than would various evidence-based treatments for cardiovascular disease.

Examining evidence from this perspective calls attention to mismatched priorities—most health care expenditures in the United States go toward treatment of diseases and their late-stage complications and relatively few resources are devoted to primary prevention and health promotion. Similar analyses at the individual level can help patients put personal options in perspective. This article proposes a bibliographic evidence-collection center and simulation modeling program to estimate potential benefits and harms of competing interventions for populations and individuals. Such evidence-based projections would enable policy makers, clinicians, and patients to judge whether they give due priority to the interventions most likely to improve health. With the steady growth in research data, the need for a system that enables society and individuals to put evidence in perspective will become progressively more urgent.

JAMA. 1999;282:2358-2365
comes are compromised. Although this information will not necessarily alter choices, it is everyone’s right to know it. Indeed, policy makers have a duty to consider how their priorities affect the public’s health.

Unfortunately, valid data for such comparisons are lacking. The literature includes countless head-to-head trials of drugs or procedures, but it provides no “big picture” comparison of the net health effect of competing medical and public health strategies. Such information is important to populations, clinicians, and individuals.

At the population level, society cannot thoughtfully determine which measures to emphasize without data on their relative health effects. Which messages to broadcast to the public, which tests and treatments physicians should offer, which services insurers should cover, which performance indicators quality improvement organizations should track, and which questions deserve research attention cannot be prioritized in accordance with public need without knowing their relative impact on population health. For instance, a treatment that may be 10 times less effective than another treatment may receive 10 times as much funding or publicity.

At the clinical level, the shared goal of all physicians—to optimize patients’ health—cannot be realized without knowing which interventions are most likely to achieve that goal. A variety of treatments and systems of care can improve outcomes. Clinicians, pressed by health plans and government to improve the quality of care for numerous diseases, cannot optimize everything at once. To know which quality improvements a practice should address first, clinicians need data on relative benefits to order priorities according to patient needs.

At the individual level, healthy persons need information about relative effectiveness to rank recommended lifestyle changes, screening tests, drugs, and alternative therapies that they learn about from physicians, family, the media, and the Internet. Patients with disease, especially those faced with multiple treatment options, want to know the relative likelihood of benefits and harms to set personal priorities. Should a person with diabetes focus first on glycemic control or on controlling blood pressure and weight? Glycemic control lowers the 10-year risk of microvascular complications from 10% to 8%. By how much does cardiovascular risk modification reduce risk?

The irony of the information age and of evidence-based medicine is that the means to readily answer such questions are lacking. There is no shortage of primary data—thousands of clinical trials are published each year, summarized in systematic reviews and meta-analyses—but these reports quantify the effects of treatments on selected end points. They tend not to examine data from the reverse perspective: to take an important health outcome and compare the relative effects of various strategies for preventing it. For example, meta-analyses can estimate by how much mammography and adjuvant tamoxifen each lower breast cancer mortality, but no information repository ranks the potential impact of these interventions on breast cancer deaths at the population level or on an individual woman’s risk of dying from breast cancer.

Previous Work

Although the burden of disease in the population is well documented, current methods for determining which interventions are most effective in ameliorating those diseases are rudimentary. The best priority-setting methods ask whether a technology satisfies criteria on a checklist (eg, burden of disease, effectiveness, cost-effectiveness), but tools are lacking to rank interventions based on relative effects on health outcomes. Studies of cost-effectiveness contrast the economic value of interventions, not which interventions improve health the most.

Only a handful of studies have attempted such comparisons. In the early 1990s, several reports estimated the number of deaths attributable to modifiable risk factors. One analysis concluded that the leading causes of death in the United States (and their annual mortality figures) were tobacco (400,000), diet and inactivity (300,000), and alcohol (100,000). (A current effort ranks clinical preventive services based on estimated gains in quality-adjusted life years [Ashley Coffield, Partnership for Prevention, written communication, 1999].) These analyses concern primary and secondary prevention, thus providing no comparisons with treatments for existing disease.

A few reports have compared the effects of prevention and treatment on life expectancy, often yielding surprising results. One analysis estimated that getting a 35-year-old man to quit smoking would improve his life expectancy by 10 months, but performing bone marrow transplantation on a patient with non-Hodgkin lymphoma would add 72 months. The seeming superiority of transplantation says more about how life expectancy is calculated than about the treatment’s relative effectiveness or its impact on public health. Gains in life expectancy refer to differences in the area under survival curves, not to the extra time a patient will live after treatment. When benefits are averaged across individuals, measures that dramatically reduce mortality for populations seem to have a more trivial impact than those that markedly increase survival for a small subgroup. To many, the more practical question is which intervention is most likely to prevent the occurrence or complications of a disease or untimely death.

An Examination of Preventable Deaths

What is the relative impact of primary prevention, screening, and treatment on the number of deaths each year in the United States? For instance, helping people to stop smoking, exercise, lower cholesterol levels by 10 percentage points, and control blood pressure would prevent an estimated 328,000, 178,000, 133,000, and 68,000 deaths, respectively (Table 1). Pneumococcal vaccination and screening for breast, cervical, and colorectal cancer would each
prevent an estimated 4000 to 10 000 deaths annually. Cardiovascular treatments (β-blockers and aspirin for acute myocardial infarction, angiotensin-converting enzyme inhibitors for congestive heart failure, warfarin for atrial fibrillation) are estimated to prevent approximately 3000 to 17 000 annual deaths.

Similar comparisons on an individual level put personal options in perspective. Suppose that a 45-year-old woman, advised by her physician to stop smoking, to exercise regularly, and to get a mammogram, makes getting a mammogram her priority because she believes that early cancer detection is most likely to prevent her untimely death. As

<table>
<thead>
<tr>
<th>Intervention</th>
<th>RRR</th>
<th>Mortality Subcategory</th>
<th>Proportion of Population at Risk (p)</th>
<th>Total Deaths (Z)</th>
<th>Preventable Deaths, No.</th>
<th>Source Data and Comments†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Prevention: Risk Factor Modification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking cessation</td>
<td>0.59, in men²⁶</td>
<td>All-cause</td>
<td>0.25</td>
<td>2 239 229</td>
<td>328 044</td>
<td>p = percentage of persons ≥ age 18 y reporting current smoking²⁷ Z = deaths from all causes among persons ≥ age 25 y in 1996²³</td>
</tr>
<tr>
<td>Physical activity (“moderate sports activity” as defined by Paffenbarger et al²⁶)</td>
<td>0.77, in men²⁶</td>
<td>All-cause</td>
<td>0.29</td>
<td>2 239 229</td>
<td>177 940</td>
<td>p = percentage of persons ≥ age 18 y reporting no leisure-time physical activity²⁹ Z = deaths from all causes among persons ≥ age 25 y in 1996²³</td>
</tr>
<tr>
<td>Lipid lowering (reduction of total cholesterol level by 10 percentage points)</td>
<td>0.89²⁹</td>
<td>All-cause</td>
<td>0.51</td>
<td>2 239 229</td>
<td>132 777</td>
<td>p = percentage of adult population with total cholesterol level ≥ 200 mg/dL,²⁰ Z = deaths from all causes among persons ≥ age 25 y in 1996²³</td>
</tr>
<tr>
<td>Blood pressure control (with high-dose diuretic therapy, as defined by Psaty et al³¹)</td>
<td>0.88³¹</td>
<td>All-cause</td>
<td>0.23</td>
<td>2 239 229</td>
<td>68 382</td>
<td>p = percentage of persons ≥ age 18 y with hypertension³² Z = deaths from all causes among persons ≥ age 25 y in 1996³²</td>
</tr>
<tr>
<td><strong>Primary Prevention: Immunizations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumococcal vaccine in persons ≥65 y</td>
<td>0.78‡</td>
<td>Pneumonia</td>
<td>0.55</td>
<td>74 349</td>
<td>9922</td>
<td>p = percentage of persons ≥ age 65 y who have not received pneumococcal vaccination³³ Z = deaths from pneumonia (ICD-9 codes 480-486) in persons ≥ age 65 y in 1996²³</td>
</tr>
<tr>
<td><strong>Secondary Prevention: Cancer Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammography every 1-2 years in women ≥40 y</td>
<td>0.73§</td>
<td>Breast cancer</td>
<td>0.31</td>
<td>43 500</td>
<td>4475</td>
<td>p = percentage of women ≥ age 40 y who had not had mammogram in last 2 y, 1996-1997³⁶ Z = estimated cancer-specific deaths for 1998³⁶</td>
</tr>
<tr>
<td>FOBT annually in persons ≥50 y</td>
<td>0.77³⁷</td>
<td>Colorectal cancer</td>
<td>0.69</td>
<td>56 500</td>
<td>9632</td>
<td>p = proportion of persons ≥ age 50 y who have not had FOBT in past year adjusted for sex distribution in this age group (45% male, 55% female)³⁷ Z = estimated cancer-specific deaths for 1998³⁷</td>
</tr>
<tr>
<td>Papanicolaou smears at least once every 3 y in women ≥20 y</td>
<td>0.09∥</td>
<td>Cervical cancer</td>
<td>0.29</td>
<td>4900</td>
<td>3644</td>
<td>p = percentage of women aged 21-65 y who had not had a Papanicolaou smear in the last 3 years, 1997³⁸ Z = estimated cancer-specific deaths for 1998³⁸</td>
</tr>
</tbody>
</table>
shown in Table 2, the patient faces a 1.8% probability of dying from breast cancer before age 75 years. The chance that mammography will prevent her death during that time is 0.5% (1 chance in 205), and the probability that other screening tests will do so is even lower. Her life is much more likely to be saved by primary prevention. Stopping smoking and becoming physically active would reduce her 30-year risk of dying by 10.9% (1 in 9) and 6.1% (1 in 16), respectively. Lipid and blood pressure control would offer similar benefits. Compared with these lifestyle changes, disease control would offer similar benefits. Comparative Lipid and blood pressure control would reduce her 30-year risk of dying by 10.9% (1 in 9) and 6.1% (1 in 16), respectively.

Yet in practice, the opposite occurs. Most dollars spent on health care in the United States go toward treating advanced disease, often late in life. Only 3% of health care expenditures is devoted to prevention. Health promotion and primary prevention receive relatively little attention in patient care, medical journals (S.H.W., unpublished data, 1999), health insurance coverage, research, and medical news reports. True to the “rule of rescue,” society focuses on individuals with diseases rather than on preventing those diseases.

Agreement with the data and policy ideology presented herein is less important than the underlying principle: a reliable method for comparing and contrasting the effectiveness of interventions in achieving health outcomes is needed to properly evaluate priorities. Lacking this, the most basic of questions—are we doing the things that help people the most?—cannot be answered. Clinical trials and meta-analyses do not, by themselves, provide the answer. Data such as those in Tables 1 and 2 begin to clarify the price paid for misplaced priorities in lost lives and health, but greater precision is necessary if such data are to inform the kind of information that the medical community currently lacks and how valuable it could be to policy and personal choices. For example, if these estimates are even roughly accurate, then prevailing priorities for the nation and individuals are markedly askew. To save the most lives, clinicians, health plans, insurers, and researchers should be concentrating their resources on helping people to stop smoking, exercise regularly, lower lipid levels, and control blood pressure.

### Policy Implications

The point behind these estimates is not to suggest that they are necessarily correct or even the best way to present the evidence; more rigorous analyses are needed for precise projections, as is detailed below. Rather, the intent is to illustrate the kind of information that the medical community currently lacks and how valuable it could be to policy and personal choices. For example, if these estimates are even roughly accurate, then prevailing priorities for the nation and individuals are markedly askew. To save the most lives, clinicians, health plans, insurers, and researchers should be concentrating their resources on helping people to stop smoking, exercise regularly, lower lipid levels, and control blood pressure.

Yet in practice, the opposite occurs. Most dollars spent on health care in the United States go toward treating advanced disease, often late in life. Only 3% of health care expenditures is devoted to prevention. Health promotion and primary prevention receive relatively little attention in patient care, medical journals (S.H.W., unpublished data, 1999), health insurance coverage, research, and medical news reports. True to the “rule of rescue,” society focuses on individuals with diseases rather than on preventing those diseases.

Agreement with the data and policy ideology presented herein is less important than the underlying principle: a reliable method for comparing and contrasting the effectiveness of interventions in achieving health outcomes is needed to properly evaluate priorities. Lacking this, the most basic of questions—are we doing the things that help people the most?—cannot be answered. Clinical trials and meta-analyses do not, by themselves, provide the answer. Data such as those in Tables 1 and 2 begin to clarify the price paid for misplaced priorities in lost lives and health, but greater precision is necessary if such data are to inform

---

**Table 1. Projected Outcomes From Population Perspective (cont)**

<table>
<thead>
<tr>
<th>Intervention</th>
<th>RRR</th>
<th>Mortality Subcategory</th>
<th>Proportion of Population at Risk (p)</th>
<th>Total Deaths (Z)</th>
<th>Preventable Deaths, No.</th>
<th>Source Data and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tertiary Prevention: Early Treatment of Cardiovascular Disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of ACE inhibitors in patients with congestive heart failure</td>
<td>0.69</td>
<td>Congestive heart failure</td>
<td>0.69</td>
<td>46 484</td>
<td>11 000</td>
<td>p1; Z = deaths from congestive heart failure (ICD-9 code 428), 1996</td>
</tr>
<tr>
<td>Use of β-blockers during/after acute myocardial infarction by patients who lack contraindications</td>
<td>0.75</td>
<td>Myocardial infarction</td>
<td>0.26</td>
<td>213 448</td>
<td>17 023</td>
<td>p = patients ≥ age 65 y without contraindications discharged from hospital after myocardial infarction without prescription for β-blockers; Z = deaths from acute myocardial infarction (ICD-9 code 410) in persons ≥ age 25 y</td>
</tr>
<tr>
<td>Use of aspirin during/after acute myocardial infarction by patients who lack contraindications</td>
<td>0.79</td>
<td>Myocardial infarction</td>
<td>0.19</td>
<td>213 448</td>
<td>10 365</td>
<td>p = percentage of “ideal” patients with acute myocardial infarction discharged from nonpilot hospitals without prescription for aspirin, 1995; Z = deaths from acute myocardial infarction (ICD-9 code 410) in persons ≥ age 25 y</td>
</tr>
<tr>
<td>Use of warfarin by patients with atrial fibrillation who lack contraindications</td>
<td>0.67</td>
<td>Cerebrovascular disease</td>
<td>0.60</td>
<td>14 983</td>
<td>3418</td>
<td>p2 (denominator is ambulatory visits by patients with atrial fibrillation without reported use of warfarin, 1996 National Ambulatory Medical Care Survey); Z = stroke deaths for 1996, multiplied by 0.094 (population-attributable risk of strokes due to atrial fibrillation)</td>
</tr>
</tbody>
</table>

---

*RRR indicates relative risk reduction associated with the intervention; CI, confidence interval; ICD-9, International Classification of Diseases, Ninth Revision; FOBT, fecal occult blood testing; and ACE, angiotensin-converting enzyme.
†Preventable deaths = Z (p) [1 − RRR]/(RRR + p [1 − RRR]); p indicates proportion of eligible population that has not received the intervention; Z, number of deaths potentially preventable by the intervention.
‡The RRR is derived from a pooled effect in a meta-analysis, which was not statistically significant (95% CI, 0.57-1.06).
§The RRR is derived for women aged ≥50 years (effect size for women 40-49 years is uncertain).
¶The RRR is derived for women aged 40-49 years from pooled observational data.
**The RRR is derived from a pooled effect on vascular deaths from antithrombotic therapy.**

©1999 American Medical Association. All rights reserved.
rather than mislead. Sophisticated and updated analyses (rather than the crude estimates in this and prior reports) and ready access to this information by policy makers, clinicians, and patients are required.

**Uncertainties Surrounding Estimates**

Any such effort would encounter formidable methodologic problems. To appreciate the breadth of the challenge, one need look no further than the methods used to generate the estimates in Tables 1 and 2. Population projections (Table 1) were derived by multiplying the number of deaths potentially preventable by smoking cessation by 0.59, for men, and the number of deaths potentially preventable by physical activity by 0.77, for men. Population projections (Table 2) were derived by multiplying the number of deaths potentially preventable by lipid lowering by 0.89, for men, and the number of deaths potentially preventable by blood pressure control by 0.88, for men.

| Table 2. Probability of Benefit From the Perspective of a 45-Year-Old Woman* |
|---------------------------------|----------------|----------------|----------------|----------|---------------------------|
| **Intervention** | **30-Year Probability of Death Without Intervention (Mortality Subcategory), %** | **RRR** | **30-Year Probability of Death With Intervention, %** | **Absolute Difference, %** | **NNT** | **Source Data and Comments** |
| Smoking cessation | 26.7 (from all causes) | 0.59, for men | 15.8 | 10.9 | 9 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data |
| Physical activity (“moderate sports activity” as defined by Paffenbarger et al26) | 26.7 (from all causes) | 0.77, for men | 20.6 | 6.1 | 16 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data |
| Lipid lowering (reduction of total cholesterol level by 10 percentage points) | 26.7 (from all causes) | 0.89, for men | 23.8 | 2.9 | 34 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data |
| Blood pressure control (with high-dose diuretic therapy, as defined by Psaty et al22) | 26.7 (from all causes) | 0.88, for men | 23.5 | 3.2 | 31 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data |
| Pneumococcal vaccine beginning at age 65 years | 0.6 (from pneumonia) | 0.78, for men | 0.5 | 0.1 | 716 | 30-year probability derived from reference 2, not adjusted for sex |
| Mammography every 1-2 years | 1.8 (from breast cancer) | 0.73, for men | 1.3 | 0.5 | 205 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data |
| FOBT annually beginning at age 50 years | 0.9 (from colorectal cancer) | 0.77, for men | 0.7 | 0.2 | 453 | 30-year probability developed from survival calculations using age-(by quintile) and sex-specific SEER data, 1991-199548 |
| Papanicolaou smear at least once every 3 years | 0.2 (from cervical cancer) | 0.09, for men | 0.0 | 0.2 | 578 | 30-year probability developed from survival calculations using age-(by quintile) and sex-specific SEER data, 1991-199549 |
| Use of β-blockers during/after acute myocardial infarction by women who lack contraindications | 3.3 (from heart disease) | 0.75, for men | 2.5 | 0.8 | 120 | 30-year probability developed from survival calculations using age-(by decile) and non–sex-specific data for acute myocardial infarction (ICD-9 codes 390-398, 402, 404-429) |
| Use of aspirin during/after acute myocardial infarction by women who lack contraindications | 3.3 (from heart disease) | 0.79, for men | 2.6 | 0.7 | 143 | 30-year probability developed from survival calculations using age-(by decile) and non–sex-specific data for acute myocardial infarction (ICD-9 codes 390-398, 402, 404-429) |
| Use of warfarin by women with atrial fibrillation who lack contraindications | 0.2 (from stroke) | 0.67, for men | 0.1 | 0.1 | 2014 | 30-year probability developed from survival calculations using age-(by decile) and sex-specific data for cerebrovascular disease (ICD-9 codes 430-438) |

* NNT indicates number needed to treat, the reciprocal of the absolute difference between the probabilities (NNT is the number of 45-year-old women who would need to undergo the intervention to prevent a death before age 75 years); RRR, relative risk reduction associated with the intervention; CI, confidence interval; FOBT, fecal occult blood testing; SEER, National Cancer Institute’s Surveillance, Epidemiology and End Results Cancer Registry Program; and ICD-9, International Classification of Diseases, Ninth Revision.

† See corresponding footnotes to Table 1.

‡ Value rounded up from 0.0497%.
the intervention by the preventable fraction. The preventable fraction is defined by the formula \( p(1-RRR)/(RRR + p(1-RRR)) \), where \( p \) is the proportion of the eligible population that has not yet received the intervention and RRR is the relative risk reduction associated with the intervention.

The probability that an intervention would prevent a 45-year-old woman from dying before age 75 years (Table 2) is the absolute difference between the 30-year cumulative probability of death with and without the intervention. The latter was derived from a standard survival calculation for a hypothetical cohort of 45-year-old women, using the annual mortality rates specified in Table 2 to estimate both the number of women entering successive years and cumulative death rates. The cumulative death rates were multiplied by the RRR values in Table 1 to derive the 30-year probability of death with the intervention.

This approach has its limitations. For example, the preventable fraction formula makes the statistical assumption that risk reductions occur equally and independently to all who receive the intervention and not at all to others.\(^{38}\) This is untrue for many interventions. Moreover, each component in the formula is imprecise.

**Relative Risk Reductions.** Although each RRR value was taken from a major meta-analysis or clinical trial, data from a particular study apply only to certain conditions, are surrounded by wide confidence intervals, and differ from other reports. Most data come from randomized trials, but some are from observational studies. The rates are incidence-based but are used cross-sectionally. The RRR is applied equally to all persons who die of the disease, although all deaths are not equally preventable. The analysis for the 45-year-old woman presumes that interventions are delivered and prevent death at the same rate for 30 years. Some rates are for all-cause mortality, while others are for disease-specific deaths. The model is binary, but health effects are continuous. Table 1 assumes, for example, that exercise prevents deaths only for sedentary persons, yet all persons benefit to some degree from more intense activity.\(^{39}\)

Estimates assume complete adherence to study conditions, and the projections for primary prevention assume that the population is 100% successful in changing behavior. Similarly, the calculation for the 45-year-old woman assumes complete adherence to treatment. Projections based on such optimistic assumptions give policy makers and individuals an upper boundary of what is possible but are unrealistic unless bounded by estimates using current compliance rates. Optimal trial conditions (efficacy) misrepresent the real world (effectiveness), where variations in clinician skills, the intensity and duration of interventions, patient adherence, and local resources influence outcomes.

**Proportion of Population at Risk (\( p \)).** Data on the proportion of persons who have not received interventions are also elusive. The rates vary by age, race, socioeconomic status, insurance coverage, and access to care. Estimates for some interventions (eg, cardiac drug therapy) derive from studies in selected settings, vary widely, and often measure adherence momentarily (eg, at hospital discharge) rather than over the long-term.

**Total Deaths (\( Z \)).** For interventions that reduce all-cause mortality, the model defines total deaths (\( Z \)) as the number of deaths in the United States among persons aged 25 years or older. For interventions that achieve a reduction in a specific type of death, the number of cause-specific deaths in the population that could potentially benefit from the intervention was used. Such choices force implicit assumptions about causality. For example, counting deaths in persons aged 25 years or older as potentially preventable by smoking cessation assumes that smoking causes death within a decade of exposure.

Using annual deaths to contrast interventions also is controversial. That an intervention prevents a death does not clarify at what age death occurs, for how long death is averted, or the quality of life remaining.\(^{24}\) Measuring only mortality ignores illness, injuries, and emotional disorders. To many, quality-adjusted life years or cost-effectiveness information is more meaningful. Furthermore, historical context is ignored by the model’s reliance on current death rates. For example, Papanicolaou screening appears to save relatively few lives, because the current death rate from cervical cancer is relatively low (3.4/100,000).\(^{3} \) Four decades ago, before Papanicolaou screening, the death rate was more than 10 in 100,000.\(^{60} \)

**Other Issues.** Interventions are defined inconsistently across data sources. For instance, lipid lowering can refer to diet or drug therapy, to different lipid fractions, and to different end points (average population level, relative reduction in levels, proportion of population below a fixed value). The best data for each component of the calculation may come from different populations. For the mammography calculations (Table 1), the values for total deaths (\( Z \)) encompass all age groups, the values for RRR come from women aged 50 to 69 years, and the values for proportion of the population at risk (\( p \)) come from women aged 40 years or older.

Table 1 makes the best case for treatments by examining the leading cause of death, cardiovascular disease, and by evaluating treatments that reduce mortality for that disease by large percentages (21%-33%). The number of lives saved by treatments for less common causes of death (eg, acquired immunodeficiency syndrome) would be lower. In Table 2, projected benefits of treatments, averaged across all 45-year-old women, are smaller than would be expected for women with extant disease. The chances that a 45-year-old woman will have an acute myocardial infarction before age 75 years and that β-blockers will prevent her death is 0.8% (1 in 120) (Table 2), but a patient who has had an acute myocardial infarction, who faces a 13% probability of dying within 1 year,\(^{43} \) has a 3.2% probability (1 in 31) of averting death if given a β-blocker when discharged from the hospital (assuming RRR = 0.75).\(^{3} \)

**Proposal for a National Program for the Synthesis of Evidence**

These daunting problems are solvable. Well-designed simulation models can in-
PERSPECTIVE IN EVIDENCE-BASED MEDICINE

corporate complex variables and permit users to examine projections and conduct sensitivity analyses from whatever perspective they wish. Each user can specify outcome measures (eg, deaths averted, quality-adjusted life expectancy, cost-effectiveness ratio), populations and risk profiles, type and intensity of interventions, magnitude of risk reduction, and degree of patient adherence. The model can format projections to accommodate different users, whether they be officials making decisions for populations or individuals (and their clinicians) making choices for themselves.

Mathematical models are imperfect: they cannot replace real observation and are always subject to criticism, but even approximations of relative effectiveness are better than guesswork. There are precedents for using simulation models in public policy (eg, weather forecasting, air traffic control). To date, their application in health care has been limited, but some work has been done with heart disease and in priority setting for US communities and European countries.

An evidence analysis service that projected the outcomes of competing health interventions based on the most recent data would have value to policy makers, clinicians, and individuals. Policy makers (eg, legislators, health plans, research funders) could use the projections for more rational resource allocation. The data could enlighten debates on priorities in government and managed care, testimony on Capitol Hill and in statehouses, and public discourse in press conferences and news reports. Including socioeconomic variables in the model would permit policy makers to contrast the benefits of investing in medical care with broader social policies affecting education, job security, housing, and the environment.

At the clinical level, data projections could be used in decision aids and software to help clinicians and patients compare the likely effects of interventions. At the individual level, projections could be presented in lay print media, broadcasts, and Web sites, enabling both healthy persons and patients to compare the relative importance of medical and self-care interventions.

To achieve credibility and viability, the data resource must be scientifically rigorous, impartial, sustainable, and user-friendly. To achieve scientific rigor, the effort must involve multiple disciplines (simulation modeling, operations research, systems analysis, epidemiology, statistics) to solve inherent methodologic problems. The simulation model would need to import data from an evidence warehouse, accommodate multivariate assumptions specified by the user, and export projections in formats suitable for the intended audience.

A bibliographic data warehouse to maintain and organize evidence for the model could map studies to a relational database that features links between outcomes (eg, diabetic retinopathy) and related interventions (eg, glycemic control, laser therapy). Users seeking evidence about the outcomes of an intervention or interventions that affect an outcome would be taken directly to the key studies at that linkage. Such a database would be more “intelligent” than conventional search tools (eg, MEDLINE), which find studies by matching words and index terms. The speed and specificity of such a system would have broad appeal; clinicians and policy makers have great need for quick access to current data on the effects of interventions.

Impartiality must be safeguarded if the data are to be trusted. Model projections will favor some interventions more than others, potentially threatening or enhancing some commercial and professional interests. Sponsors and investigators should have no perceived gain in downgrading costly services (eg, payers) or promoting products (eg, pharmaceutical companies). The methods, bibliographic data warehouse, and simulation model should be nonproprietary and accessible in the public domain (eg, unrestricted Internet access).

The project would require a sustainable infrastructure (governance rules, partner agreements) to remain viable. The costs of the effort would require a stable funding source committed to long-term operation, such as a foundation or government agency that lacked a vested interest in the outcome of the specific projections. Ultimately, the project must generate useful end products, with materials developed by different professionals for diverse audiences: policy and commercial analyses for legislatures and health plans, detailed data for medical journals, press releases for news media, office tools for busy clinicians, lay materials that present data understandably, and software and Internet services.

Such an ambitious project cannot be launched at once. It must be phased in, beginning with exploratory methods and work and a gradual expansion of analytic capabilities and infrastructure. Initial reviews should be limited in scope to certain health interventions (eg, preventive services) and outcome measures (eg, mortality) for pilot testing. But the need to begin is paramount. That such a system is lacking and that societal and personal priorities must be set without knowing the consequences is unsettling. Evidence-based medicine has provided a solid foundation—much more is known about the effects of individual interventions—but it is now time to establish systems to organize the wealth of evidence and put it in perspective. With the escalating growth in research, the need for better organization of evidence will only become more urgent.

Acknowledgment: This project would not have been possible without the encouragement and support of David W. Marsland, MD; I thank David Atkins, MD; Jeff Harris, MD, MPH; J. Michael McGinnis, MD, and Steven M. Teutsch, MD, MPH, for their helpful review comments on this article.

REFERENCES
5. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure [pub-
lished correction appears in JAMA. 1995;274:462].


PERSPECTIVE IN EVIDENCE-BASED MEDICINE

©1999 American Medical Association. All rights reserved.