Weighing the evidence: the Canadian experience\(^1,2\)

Richard B Goldbloom

**ABSTRACT** The Canadian Task Force on the Periodic Health Examination created a hierarchy of evidence that has been used for the past 18 y to evaluate the scientific evidence for and against the preventability of each condition reviewed. The methodology developed by the task force may be applicable to study of the preventive aspects of dietary sodium and health. This paper describes the history and modus operandi of the task force, *Am J Clin Nutr* 1997;65(suppl):584S–6S.

**KEY WORDS** Canadian Task Force on the Periodic Health Examination, public health policy, dietary sodium, health assessment

**HISTORY OF THE CANADIAN TASK FORCE ON THE PERIODIC HEALTH EXAMINATION**

In the early 1970s, Federal Minister of Health Marc Lalonde (1) published a paper titled *A New Perspective on the Health of Canadians*. The central theme of this work was that most conditions adversely affecting the health of Canadian citizens then and in years to come were what Lalonde termed “diseases of lifestyle”: problems such as substance abuse, accidents, violence, poverty, family dysfunction, poor diet, indolence, and unwanted pregnancy.

One consequence of the Lalonde Report occurred in 1976, when the Conference of Deputy Ministers of Health established the Canadian Task Force on the Periodic Health Examination. The task force was mandated “to determine how the periodic health examination might enhance or protect the health of the population... [and] to recommend a plan for a lifetime program of periodic health assessments for all persons living in Canada.” Although the principal focus of the task force was the clinical encounter between asymptomatic patient and caregiver, we recognized that for many conditions this focus widely overlapped issues of public health policy. The task force methodology is equally relevant as a basis for public health policy as for the clinical encounter. When we consider the possibility of modifying a nation’s sodium intake, the focus is, like that of the task force, on asymptomatic citizens.

I had the privilege of joining the task force at its inception in 1976 and have served as its chair for the past decade or so. When I joined the task force, I expected that we would immediately begin a systematic examination of many conditions. As it turned out, we spent the first 2 y of our existence developing and refining a methodology before starting to tackle a single potentially preventable condition. This seemed an interminable, frustrating exercise at the time, but with the wisdom of hindsight I realize it was the most valuable time that we spent. The task force made a conscious, fundamental policy decision that evidence was to take precedence over opinion (or consensus). For this to happen, we created a hierarchy of scientific evidence that we have used ever since to evaluate the preventability of every condition under review.

From the beginning, being part of the Canadian task force has been an incredible learning experience, especially for a pediatrician who had no particular background in epidemiology, adult preventive medicine, or public health. If, for example, you had asked me in 1976 to explain how efficacy, effectiveness, and efficiency are different, I would undoubtedly have stammered and tried to fake my way through. These terms and their practical implications have come to hold very serious meaning, and I realize now that in evaluating any potentially preventive intervention, the distinction is of fundamental importance.

Three criteria guided our choices of conditions to be assessed: 1) the current burden of suffering; for example, the number of people affected, the degree of disability, and mortality; 2) the validity and acceptability of the detection maneuver; and 3) the effectiveness of the resulting intervention (effectiveness = efficacy ± compliance).

In speaking to medical audiences about various preventive measures I have sometimes tried to illustrate a point by asking how many people in the audience have had a sigmoidoscopy examination. It is not unusual to get a show of 20–30 hands. I then ask how many have had a second sigmoidoscopy. Typically, the response is one or two. As you know, some authorities have suggested periodic screening by sigmoidoscopy as a preventive measure for colon cancer. However, if most of the asymptomatic customers cannot be enticed to return after their first experience, the procedure can hardly be recommended as effective for widespread or universal screening, aside from such issues as cost or the number of examinations required to benefit a single individual. Acceptability of the intervention is an essential prerequisite to widespread implementation.

We have come to realize how the application of many widely recommended preventive health care measures has been based on extrapolation far beyond what was justified by the available evidence. The power of wishful thinking was epitomized some time ago in a cartoon in the *New Yorker*, showing two angels

\(^1\) From the Department of Pediatrics, Dalhousie University, Halifax, Canada.

\(^2\) Address reprint requests to RB Goldbloom, 1 Walton Killam Children’s Hospital, 5850 University Avenue, Halifax B3J 3G9, Nova Scotia, Canada.
standing on a cloud, one saying to the other: “to play it safe, I still take one aspirin every other day.”

The Canadian task force created a hierarchy of evidence that has been used for the past 18 y to evaluate the scientific evidence for and against the preventability of each condition reviewed (Table 1). Note that we give the lowest evidential status to opinions of respected authorities. In clinical practice, doctors caring for individual patients must make binary decisions on most issues: ie, do it or do not do it. However, the scientific evidence on the effectiveness of many preventive interventions does not allow for simplistic all-or-none recommendations. Modifying factors may include age, sex, risk status, and environmental factors. We therefore developed a classification of recommendations that was bidirectional and of graded strength (Table 2).

In 1979, the Canadian task force first published its methodology and its evaluations of the preventability of 76 conditions. A central recommendation was that the previously undefined annual checkup be abandoned and replaced by a lifetime sequence of age- and sex-specific preventive packages of proven effective interventions. Finally, it was recommended that most preventive interventions be incorporated into health visits for other purposes.

The effect of our initial report was amazing. We received > 40 000 requests for reprints, and several years later when the US Preventive Services Task Force was established it adopted the Canadian methodology virtually unchanged. Since then, the Canadian and US task forces have worked in a tight, productive collaboration, attending each others meetings and often combining forces in analyzing various conditions. The Canadian Guide to Clinical Preventive Health Care, includes among its 81 chapters some 20 originally prepared by US task force members and adapted by Canadian task force members to the Canadian context (2). The US guide to clinical preventive services will contain several chapters originally written by members of the Canadian task force and adapted to the US context (3).

The combined experience of the two task forces has taught us several important lessons. First, we learned that when we weigh the evidence rigorously and consistently, then, with a few notable contentious exceptions (eg, mammography, cholesterol screening, and prostate cancer screening), the direction and strength of recommendations emanating from most analyses are usually readily apparent and do not require endless debate or a decision by consensus.

Second, an invaluable side effect of this rigorous analytic process is the automatic production of a set of unanswered questions about each condition under review. This research agenda is the essential basis for developing clearer or stronger recommendations in the future.

Third, we have been impressed by the relatively small number of conditions for which the evidence is sufficiently strong to support an A recommendation based on type 1 evidence. This underlines the immaturity of the science and practice of preventive medicine. It is perhaps not surprising that most of the preventive interventions that do merit an A recommendation based on type 1 evidence (Table 1) are those that are applied prenatally or soon after the beginning of life, such as newborn screening for hypothyroidism and phenylketonuria and childhood immunizations against various infectious diseases.

Aside from interventions such as smoking cessation and screening for hypertension, the payoff of preventive interventions tends understandably to diminish the later in the life cycle they are applied. Clearly, even recommendations that are currently solidly based on evidence (or lack thereof) must be reviewed periodically as new evidence emerges or to take into account different populations and environmental changes.

**TWO FUNDAMENTAL PRINCIPLES**

Our experience suggests that two principles should be fundamental to any consideration of modifying the sodium intake of the entire population. 1) Intermediate outcomes must not be used as surrogates for health outcomes (eg, a significant reduction in average blood pressure is not significant unless it can be shown to cause a clinically significant reduction in hypertensive morbidity or mortality without significant harm). 2) All potential outcomes of the intervention must be considered, not only those involving the morbidity and mortality of the target condition, but also outcomes that may involve morbidity and mortality of unrelated conditions.

The great cholesterol debate has important lessons to teach us on this score. If, as some clinical trials have suggested, screening for serum cholesterol or interventions to lower the population’s serum cholesterol concentrations can be shown to reduce the number of coronary events but do not affect all-cause mortality (or if all-cause mortality increases), then one may correctly question the net gain to society.

**WHEN EMOTION TRIUMPHS OVER REASON**

There is an additional important consideration. The recent history of preventive health care contains several telling examples in which emotion has triumphed over reason and in which legislators or disease-oriented associations were unwilling to wait for solid evidence of effectiveness before translating pre-
liminary evidence into action in the sacred name of prevention. Let me cite two recent examples. Some 13 states in the United States have legally mandated screening of schoolchildren for scoliosis. However, a careful review of the evidence by the Canadian and US task forces concluded that there was insufficient evidence to indicate that scoliosis screening was effective or ineffective. Similarly, some jurisdictions have mandated preschool developmental screening with the Denver Developmental Screening Test, although the single randomized clinical trial of such a program showed that such screening provides no detectable benefit and may cause harm by generating unnecessary anxiety in parents.

Thus, although we are often ready to condemn the communications media for premature announcements of promised cures or preventative items, health policymakers can be equally guilty of implementing interventions before the evidence for effectiveness is sufficiently conclusive. In 1986, Marshall Becker (4), then chairman of the Department of Health Behavior and Health Education at the University of Michigan, addressed the annual meeting of the US Public Health Association on the tyranny of health promotion. He pointed out that "by suggesting that almost everything we consume, do, or interact with in our environment is dangerous, we trivialize both our efforts at health promotion in general and those recommendations which, if implemented, might truly exert important positive influences on health" (4).

At least four questions about dietary sodium need to be addressed. 1) Are there good population-based randomized clinical trials of reduced dietary sodium intake? 2) If good population-based randomized clinical trials have been done, were intermediate outcomes (eg, lowering blood pressure) or ultimate health outcomes (eg, mortality and morbidity) determined? 3) What are the proven or potential adverse effects of reducing dietary sodium intake in the entire population (eg, hypertensive and normotensive individuals, individuals in hot climates or occupations, and growing children). 4) Do individuals with high or low sodium intakes compensate significantly (sodium through increased sweat or renal losses) and, if so, to what extent does this negate the effect of modifying intake?

PUBLIC HEALTH POLICY

Because of the numbers of people who will be affected by authoritative pronouncements, the enunciation of public health policy carries an even weightier responsibility than does the development of policy affecting individual patients. Nowadays, we must be particularly sensitive to the need to avoid further confusing an already thoroughly confused public. Angell and Kassirer (5) commented on this issue in the New England Journal of Medicine. They noted that we "are serving a public that believes passionately that the more we can learn about what to eat or how to live, the longer we will live." They concluded that "although we would all like to believe that [such] changes... can greatly improve our health, the likelihood is that, with a few exceptions such as smoking cessation, many if not most such changes will produce only small effects" (5).

Perhaps the most eloquent enunciation of this scientifically cautious philosophy was written by the late Lewis Thomas (6) in his essay "On Magic in Medicine." Thomas used as his theme the "seven healthy life habits" then widely promoted by a major US health insurance company. He said the following:

Nobody can say an unfriendly word against the sheer goodness of keeping fit, but we should go carefully with the promises.

There is a bifurcated ideological appeal contained in the seven-life-habits doctrine, quite apart from the subliminal notion of good luck in the numbers (7 come 11). Both ends of the political spectrum can find congenial items. At the further right, it is attractive to hear that the individual, the good old freestanding, free-enterprising American citizen, is responsible for his own health and when things go wrong it is his own damn fault for smoking and drinking and living wrong (and he can jolly well pay for it). On the other hand, at the left, it is nice to be told that all our health problems, including dying, are caused by failure of the community to bring up its members to live properly, and if you really want to improve the health of the people, research is not the answer; you should upheave the present society and invent a better one. At either end, you can't lose.

In between, the skeptics in medicine have a hard time of it. It is much more difficult to be convincing about ignorance concerning disease mechanisms than it is to make claims for full comprehension, especially when the comprehension leads, logically or not, to some sort of action. When it comes to serious illness, the public tends, understandably, to be more skeptical about the skeptics, more willing to believe the true believers. It is medicine's oldest dilemma, not to be settled by candor or by any kind of rhetoric; what it needs is a lot of time and patience, waiting for science to come in, as it has in the past, with the solid facts.

REFERENCES