

CONSIDERATIONS IN DEFINING EVIDENCE FOR PUBLIC HEALTH

The European Advisory Committee on Health Research World Health Organization Regional Office for Europe*

H. David Banta

Netherlands Organization for Applied Scientific Research

The issue of evidence for decision-making, including formal policy-making and public health policy, has increased rapidly in prominence during the past few years. The term “evidence-based” was first used in health care in 1992, when Sackett and others coined the term “evidence-based medicine” (42). More recently, many have asked if public health activities are based on *evidence* (2;8;13;19;30).

But why has *evidence* become so fashionable? The main reason is because there is a clear need to document the underlying links between a potential intervention (including a program or a policy) and health, the ultimate outcome. Health care and public health are pragmatic sciences. One implication of the scientific character of these activities is that they evaluate the consequences of their decisions and actions by systematic research. The

***Members of the Committee**

Dr. William Dab: Chaire Hygiène et Sécurité du CNAM, 2 rue Conté, 75141 Paris Cedex 03

Professor Emil S. Gabrielyan: Vice-President, The National Academy of Sciences of Armenia, and Director Drug and Medical Technologies Agency, Moskovyan 15. 375001 Yerevan

Professor Egon Jonsson: Director, Swedish Council on Technology Assessment in Health Care (SBU), Karolinska Institute, P.O. Box 5650, 11486 Stockholm

Professor José Maria Martín-Moreno: Director, Spanish Agency for Health Technology Assessment, Instituto de Salud ‘Carlos III’, Calle Sinesio Delgado 6, E-28029 Madrid

Professor Martin McKee: Head, WHO Collaborating Centre for Health of Societies in Transition, London School of Hygiene & Tropical Medicine (LSHTM), Keppel Street, London WC1E 7HT

Professor Jorma H. Rantanen: Director-General, Institute of Occupational Health, Topeliuksenkatu 41A A, SF-00250 Helsinki

Professor Hans-Heiner Raspe: Direktor, Medizinische Universität zu Lübeck, Institut für Sozialmedizin, Beckergrube 43-47, D-23552 Lübeck

Professor Per-Gunnar Svensson: Director General, International Hospital Federation, 46-48 Grosvenor Gardens, London SW1W 0EB, United Kingdom

Dr. Tessa Edejer: Coordinator a.i., Effectiveness, Quality, Cost, Gender and Ethics, Global Programme on Evidence for Health Policy, World Health Organization, Geneva

Dr. Alan Lopez: Global Programme on Evidence for Health Policy (GPE/EBD), World Health Organization, Geneva

Dr. Elena Varavikova: Evidence and Information for Policy Cluster, Department of Health Services Delivery, FSP, World Health Organization, Geneva

results give reasons as well as justification for their statements, proposals, programs, and interventions. The results must be presented in an unambiguous comprehensive manner. This forms a basis of democratic discussions in open societies.

The point is that to fulfill its mission of promoting health and preventing diseases through organized efforts by the society, public health cannot be a matter merely of belief or just be based on a sort of fundamentalist ideology. Those working in the field need to know that a proposed strategy or technology has a reasonable chance of improving health without undue risk or other serious problems. Thus, all the activities now under the rubric "evidence-based" have developed to systematically gather the practical information arising from the available research and objective information within the field, rather than simply relying on unstructured expert opinions.

A key issue is that the concept of health has been broadened and deepened during the past few years beyond the absence of disease. The field of public health, which may also be referred to under the term "population health," has correspondingly expanded, responding to the growing body of evidence linking health to social, economic, and environmental determinants of health. The "new public health paradigm" emphasizes such factors and seeks to intervene in such areas to improve human health, in the broadest sense (48;49). This shift has produced many complexities (27). In particular, new approaches are needed that go beyond the traditional relationship of a discrete exposure and specific outcome, illustrated by, for example, the link between smoking and lung cancer. These approaches must take account of the complex determinants of exposure to risk factors, such as the constraints people face when making healthy choices, and the host response, including the growing evidence of how risk factors interact and the influence of genetic factors. They must also address the full range of policy responses, which will involve many sectors that have not traditionally been thought of as playing a role in health. To add further to the challenges, they must confront the actions of vested interests that sow confusion in an already complex field as a means of preventing effective action, most clearly illustrated by the actions of the tobacco industry (1).

Evidence must be related to these new determinants. Because evidence is a broad concept, the quality of available evidence is a key consideration. Ultimately, the goal is that evidence should be appropriately used in decision-making in public health, to the benefit of the people served by public health activities (31). The Committee believed strongly that such use of evidence is part of the professionalism of all those working to promote population health. The very concept of expertise is based on the idea that the professional "knows" that he/she can back up advice and proposed action with evidence. To use such evidence is part of the ethical responsibility of every professional.

Another important issue when trying to acquire and use evidence is the context: context of acquiring evidence and context of the setting that uses the evidence. Many public health interventions are diffuse and may be hard to define. There is then a problem in generalizing the evidence to other settings. Context may be seen as a nuisance or as the main issue to be studied and understood. Research publications are dominated by work from North America and Northern Europe; how is this evidence to be applied in other settings? How can the evidence be put within a local context? Public health systems are conditioned by local contexts, including laws, regulations, and payment systems, that vary from country to country. Feasibility may be determined by deep-rooted political and cultural assumptions. Acceptance and diffusion, then, are not fundamentally rational processes, but depend on the context.

Therefore, the purpose of this study is to review the concept of evidence, including its concept, importance, limitations, and challenges. The Committee hopes through this review to provide a potentially useful basis for development of public health policies.

SOME PRELIMINARY CONCEPTS

Literature review has shown that there is no common definition for “evidence” or “evidence-based.” Moreover, the term “evidence” has been commonly used in law and judiciary contexts before being used within the medical/health context. Thus, in discussing the notion of “evidence”, the Committee believed that the complexity of the dimensions embedded in its underlying concept made it advisable to avoid giving a simple definition of evidence, instead examining and defining key aspects that could go into a definition.

The Concept of Knowledge

The characterization of “evidence” is inextricably linked with the meaning attributed to “knowledge.” This has been the subject of philosophical enquiry for over two millennia, so what follows is inevitably an extremely superficial summary of two of the main sets of issues in how we understand the world. This analysis draws extensively from a recent review by Fulop et al. (9).

The first issue is the difference between, on the one hand, positivism, objectivity, or realism and, on the other, idealism, interactionism, interpretivism, or subjectivity. The former is based on the idea that there is a single reality that can be discovered. The latter involves the observer drawing on a particular conceptual paradigm, or view of the world, to interpret what is observed. In this second approach, there is no single reality to be discovered, giving rise to a set of ideas termed postmodernism, which see knowledge as fundamentally fragmented so that it is impossible to reach a universal understanding of phenomena (16).

These different approaches create a challenge for policy-makers seeking to act on evidence. An interpretivist approach might argue that a set of observations are so dependent on the context in which they are made or the theoretical framework applied by the researcher that they cannot be generalized. On the other hand, a positivist approach may fail to recognize the implicit assumptions underlying the observations, which might not be shared by other observers.

The approach labeled “subtle realism” has been proposed as a means of transcending these different paradigms. This concept takes the view that what is observed has an independent existence but that it may be represented in different ways by observers coming from different disciplinary and theoretical backgrounds (15). However, as there is a single reality, it is expected that the different representations will be congruent.

A second issue is the difference between deductive and inductive approaches to generating knowledge. Deduction proceeds from theory, which is used to generate a hypothesis that can then be tested. The strength of the evidence so generated is a function of the rigor with which the hypothesis has been tested to see whether it can be falsified (37). This is the central scientific method in classic science. In contrast, an inductive approach proceeds from a series of observations that are then used to generate theories, which in turn can be subject to hypothesis generation and testing.

Several implications arise from the existence of these different approaches. The most important is that, in any dialogue between researchers and policy-makers, it is essential for each to understand the approach to knowledge adopted by the other. Does one group believe in the existence of a universal truth that can be applied in very different situations, whereas the other sees the same results as being so context dependent that they cease to be applicable even in situations that appear outwardly the same?

Policy-makers must also be clear about when one approach is preferable to another. A positivist approach may be the best way of assessing the effectiveness of a drug, whereas an interpretivist approach may add to understanding of the meaning of any side effects

encountered. A broad understanding of a phenomenon will often draw on different approaches. This finding is a key truth for public health.

The Concept of Evidence

In relation to interventions to promote health, the term *evidence* brings along an underlying notion of practical information for action, including aspects such as:

- Information that is appropriate for answering questions about an intervention's effectiveness (does the intervention work?)
- Applicability of effectiveness data
- Intervention's positive or negative side effects
- Economic impact
- Barriers to implementation of interventions (44).

Evidence is the result of a search for useful knowledge. Part of this search involves classifying and appraising the different types of knowledge that are available. The knowledge needed depends on the specific outcome sought from an intervention. The most important evidence in public health documents that improved health status follows from a certain intervention, such as a preventive measure or a public health program. Here, it is important that a broad definition of health is used. Mortality or life expectancy is the most common health outcome measures used in research related to outcomes. But quality of life is a key issue as well and should be included in such evaluations. Therefore, every evaluation needs a clearly formulated objective stated in terms of health outcomes.

It is often difficult to use health as the outcome, or it may not be feasible. In that case, "surrogate measures" are often used. Therefore, if surrogate measures are used, one needs good indications that such surrogates are actually related to improved health. Ideally, the links between surrogate measures and health should be established empirically, that is, by research (22). The annex gives examples of statements relevant to public health and the type of evidence that is required to document their validity.

The Concept of Evidence—Different Views

A. Evidence is any useful information that serves as a basis for making decisions. In the real-life situation of public health, experimentally based information is often hard to obtain, for a variety of reasons, some of which are described below. Action is often based on the opinions of only a few people, sometimes only an important decision-maker such as the minister of health. In other words, decisions are made explicitly by someone who might be accountable to some constituency on specific terms of reference. The criteria of relevance and pertinence are basic within this process. Taking these facts into consideration, is information based on the experience of the minister to be called evidence or not?

On one side, some argue that unsupported anecdote or experience or information based on evidence, but undocumented, cannot be called evidence. On the other side, some have said that it may be more fruitful to refer to formal knowledge resulting from research and informal knowledge resulting from experience in working in public health (20).

B. Evidence is scientific evidence. Science is generally aimed at establishing and explaining links between observed phenomena. The scientific method is the process by which scientists, collectively and over time, endeavor to construct an accurate (that is, reliable, consistent, and nonarbitrary) representation of the world. Recognizing that personal and cultural beliefs influence both perceptions and interpretations of natural phenomena,

Box 1. Some perspectives on validity

Kvale (21) differentiates between communicative and pragmatic validity for qualitative research:

- Communicative validity refers to that the researcher's interpretations are communicated and negotiated with individuals who are involved in or acquainted with the studied context.
- Pragmatic validity refers to, literally, that the reader of the analysis made by the researcher is convinced of its trustworthiness.

Maxwell (25) finds five types of validity that may be used in qualitative studies:

Descriptive, theoretical, interpretative, generalizability, and evaluative.

the aim is to minimize those influences when developing a theory. In summary, the scientific method attempts to minimize the influence of bias or prejudice in the experimenter when testing a hypothesis or a theory or when assessing the effectiveness of a public health intervention. As already mentioned, the process of scientific inquiry is based on testing hypotheses.

Fundamentally, science is experimental. However, the scientist may initiate the experiment or use a natural experiment, so that the test of the hypothesis may be made using either observational or experimental methods. The observational and experimental methods can often be complementary (12).

Science is also often interpreted as quantitative in its methods or results. This view is false. Science may be either qualitative or quantitative or both. Science is now defined more broadly and includes social science. In social science, methods are often more qualitative (5). Nonetheless, unfortunately, qualitative research is often seen as "subjective" (50).

The essence of science is seeking to prove a hypothesis false. As Descartes stated, "science is scepticism". All hypotheses can only be tentatively accepted. There is no certainty in science. This concept raises difficulties in relating to policy-makers, who are usually seeking a degree of certainty. Moreover, there are circumstances in which one cannot isolate the phenomena of interest or when one cannot repeat the measure over and over again or systematize the different steps of the scientific method. In such cases, the results may depend in part on the history or the context of a situation. This situation often occurs in social interactions between people and has to be taken as one of the challenges of scientifically based evidence.

C. Evidence is "research-based evidence". This wording was the favored use of the term in the evidence-based medicine movement from the beginnings (17). However, recently, those working in the field believe that this position was too rigid and that experience also produces information that can be called evidence (43).

Aside from unsupported opinion, most evidence is acquired through a process of research. Research, carried out by any discipline, is based on common factors that need to be considered. Research strives for generalizability (9). All fields of research have standards for judging the methods of the research and, thereby, the outcomes of the research. A key question in all areas of research concerns internal versus external validity. Internal validity is the extent to which an observation measures what it is intended to measure. External validity concerns the generalizability to other settings (see Box 1).

The Concept of Research

Research is an unbiased and objective process of enquiry that produces knowledge (12;17). Research is often divided into primary research, which is systematic empirical enquiry, and secondary research, which is the accumulation of research findings into a robust body of knowledge through analysis of primary research studies, through such processes as meta-analysis and quantitative synthesis (26). A key strategy in secondary research is the systematic review, which has been defined as "an overview of primary studies that contains

an explicit statement of objectives, materials, and methods and has been conducted according to explicit and reproducible methodology” (14). Systematic review is a clear concept that has been used quite a lot for quantitative studies, but its application to studies of different designs and outcomes has not been much considered. Specifically, there is no accepted and definitive method of synthesizing the results of qualitative research or of combining the results of qualitative and quantitative studies. From another perspective, secondary research goals are mainly oriented toward making practical or handy what is already known (for example in the case of clinical practice guidelines).

Quantitative research and **qualitative research** both have their uses, and they can also be used in a complementary manner. Qualitative research tends to study activities in more naturalistic settings and does not result in quantitative outcomes. Quantitative research is associated with research under experimental settings and typically depends on numerical outcomes. Typically, qualitative research in public health is carried out by such social science disciplines as sociology and anthropology and complements quantitative research carried out through epidemiological and statistical approaches.

Another common distinction is between **basic research** and **applied research**. Basic research usually refers to research into the biological mechanisms that underlie the normal functioning of the human body and its malfunction in disease. However, the term *basic research* is now often used in a broader sense, as “epidemiology is the basic research field of public health” (12). Applied research is said to be an activity that draws on basic research information to create solutions to problems in prevention, treatment, or cure of disease. In short, basic research seeks understanding, whereas applied research seeks to control vital processes.

When applied research has demonstrated that a particular new intervention is useful, it generally will be implemented. There is a growing body of research on organizational factors that influence the implementation of interventions (9).

In the past, research has not often been guided by strategic concerns, such as how to re-organize health care services or how to implement a prevention policy. However, today, *researchers are sometimes being called on to produce research relevant to public health concerns* (3; and W. Solesbury, manuscript submitted for publication). In research policy, the wish for strategically important research has led, in some countries, to more emphasis on “top-down” initiated research and less emphasis on investigator-initiated research.

The Concept of “Quality” of Evidence

It is commonly accepted that information may have different levels of reliability and validity. For example, information from high-quality research, if relevant to the specific question, should have a high degree of reliability and validity.

Research appraisal is an important step. Although notions such as “bias,” “confounding,” and “validity” have been components of epidemiology during the past decades, the discussion about research appraisal has gained increasing visibility during recent years, as part of health technology assessment and evidence-based medicine, for example. Gray proposed this method of research appraisal (12):

- What is the question? (Define intervention and desired outcome)
- Is this the best type of research method for this question?
- Is the research of adequate quality?
- What is the size of the beneficial effect and of the adverse effect?
- Is the research generalizable to the whole population from which the research sample was drawn?
- Are the results of the research applicable to the local population? To the local situation?

Research appraisal is now standard in epidemiological, quantitative research. Such methods have not been developed in many other areas of research. Because it is so often mentioned as a key method in public health research, the paradigm of epidemiological research will be reviewed, followed by a brief discussion of other research methods.

Epidemiology, studies of the distribution and determinants of disease, has furnished a variety of tools that can be applied in seeking public health evidence. Indeed, Gray has asserted that epidemiological evidence is the most important type of evidence in public health and that “epidemiology is the principal basic science on which public health has developed” (12;13).

Epidemiology can be either observational or experimental research. In fact, the results of observational and experimental studies often complement each other (12).

Epidemiological methods are particularly suited to evaluating the efficacy and safety of specific interventions that are aimed at directly altering health status, such as medical treatment and certain prevention interventions (46). The strongest evidence is considered to come from randomized controlled trials (RCTs). RCTs are a kind of cohort study, but are experimental. In some areas, such as efficacy of clinical interventions, quantitatively oriented research can shed important light on the issues of concern. In these cases, systematic review of research findings is typically done to ascertain the extent and nature of available knowledge. However, as already stated, methods of systematic review have not been developed to the same extent for other types of research, such as qualitative, social science research. In short, the type of evidence and its interpretation vary a great deal from one area of public health to another (see Box 1).

Observational studies may be designed as cohort or case control studies, and they can be very useful in addressing particular questions. Cross-sectional studies, including basic surveys, are also useful for some informative purposes.

However, one must be cautious in interpreting epidemiological research, for several reasons. As with all research, the method must be congruent with the sought-after results. For example, if inappropriate parameters have been used, the results may be useless or even potentially harmful in attempting to solve a particular problem.

As mentioned above, research and its outcomes may be appraised for quality and applicability. Such research appraisal is probably most useful in dealing with quantitative research such as RCTs. The different epidemiological methods may be ranked in the often-used hierarchy of evidence and research design.

The paradigm of epidemiology has been presented in some detail here because it is given such prominence in the literature on research in public health. However, there are many other methods used in public health research. Familiar to most, and used often in conjunction with epidemiology, is economic research (6;39;45). Other methods that may not be so familiar include moral and ethical analysis, organizational psychology, sociology and other social sciences, political science, policy analysis, historical research, and action research (35;36).

Therefore, a set of disciplines and their methods can be applied to public health concerns. It is important to point out that the paradigm of epidemiology can be very useful but does not always apply perfectly to certain areas of research, for reasons developed throughout this paper. Moreover, each field has its own standards of excellence. At the same time, it may not be feasible to develop a hierarchy of evidence that applies to more than one field of research.

It also seems self-evident that the research method should be fitted to the question. Taking into account that evidence is intended to influence a certain specific area of public health, the type of evidence must be fit to the problem at hand. Box 2 gives some examples of areas of public health and some fields of research that may produce evidence related to the particular challenge (see also Annex 1).

Box 2. Examples of types of research that may be related to different stages in the implementation of a program

Diagnosis of the problem	- descriptive epidemiology, anthropology
Prognosis (without intervention)	- descriptive epidemiology, sociology
Norms, values, influences	- sociology, political
Possible interventions	- epidemiology, economics
Implementation	- sociology, action research, management sciences
Barriers	- sociology, action research
Monitoring outcomes (and feedback)	- descriptive epidemiology, sociology

USING EVIDENCE IN PUBLIC HEALTH

One interested in improving public health cannot just say that it is necessary to wait until research-based evidence is available. Furthermore, a legitimate concern is that research in many areas of the “new public health” aims at actions that are difficult to evaluate, such as those in health promotion (18;38). For example, what is a “healthy city” and what are the general and specific outcomes sought? Because of these difficulties, decisions that are mainly determined by good evidence of effectiveness would favor interventions with a medical rather than a social focus, those that target individuals rather than communities or populations, and those that focus on the influence of proximal rather than distal determinants of health (40;41). This approach would clearly be unsatisfactory for population health activities.

Furthermore, finding definitive evidence in many areas, especially the newer social concerns in public health, may often be very difficult, if not impossible. Databases concerned with clinical research are far-advanced and research findings, as well as systematic reviews, are relatively easy to find. This is not true in other areas of research, such as the social sciences. Databases are not established to the same degree, and research results are not readily available in electronic form (for example by means of the Internet). Even in areas where a great deal of research has been done, controversies may not be settled (Box 3).

Some types of research can surely shed light on the relationship between the intervention and health outcome, and every effort should be made to see that such research is done and used. It is also important to promote research studies to produce such evidence. However, it should also be acknowledged that all aspects of the problem cannot be examined and that validated evidence may only be applicable in limited aspects of some policy areas. Therefore, it is often fruitful to examine all types of information, although some may be more highly valued than others (for example, based on proven validity of the knowledge

Box 3. How much evidence is enough?

Seven randomized controlled clinical (RCTs) of mammography screening have been carried out, four in Sweden. A large number of meta-analyses and syntheses of these trials have been carried out, especially in countries considering a national policy to encourage mammography screening (51). Overall, these studies indicate a reduction in breast cancer mortality over time of around 15 percent in women over the age of 50. Based on these findings, several countries, including the United Kingdom, France, the Netherlands, and Sweden, have developed nationwide breast cancer screening programs. However, such screening has continued to be controversial, and many countries have decided not to implement national screening programs. In 2001, two Danish investigators re-analyzed the data from the seven RCTs, concluding, on methodological grounds, “there is no reliable evidence that screening for breast cancer reduces mortality” (34). At least two articles have been published since the Danish study answering and attempting to refute its conclusions (29;32). Indications are that all countries interested in this issue are re-examining their positions, and some are further delaying any decision awaiting further evidence. The controversy seems certain to go on for years.

Box 4. Research outcomes that seem to give conflicting results

A policy-maker is told by one researcher that she has demonstrated that anti-depressive medication is overused in the national health services. Another researcher tells the policy-maker that there are many depressed people who are not being treated when efficacious drugs are available.

produced). Stating the goal of the activity clearly and specifically is highly important, especially in more complex settings such as public health services and management, because otherwise the person seeking for the research-based evidence may be overwhelmed with many different types of research, much of which may be irrelevant to the real question (47). Often, because of different goals of research, studies may appear to contradict each other. This finding underlies the importance of seeking evidence that is truly relevant to the problem (Box 4.)

Implementation of public health programs obviously requires attention to other issues in addition to effectiveness of the proposed intervention. Consideration must be given to such issues as resources, social justice, community participation, cultural appropriateness, and political considerations. Policy makers must take many such factors into account, including the desires of the population for certain products that may be useless or harmful, such as herbal cancer remedies and dietary supplements.

A key issue in using research evidence is its excellence. Overall, good research acknowledges the possibility of error. The research process follows procedures that must then minimize the effect such errors may have (33). In a more specific sense, research must be judged within its own disciplinary context. Multidisciplinary research raises particular challenges in this regard.

A particular problem needing much more attention is that evidence of effective interventions often does not lead to effective action (11) (see Box 5), and the flip side, what to do when evidence is lacking or weak? (see Box 6). Relatively passive methods of disseminating and implementing evidence, such as publication in professional journals or mailing to targeted audiences, rarely lead to changes in behavior. To maximize the likelihood that the recommendation is used requires coherent dissemination and implementation strategies to capitalize on known positive factors and to deal with obstacles that have been identified. In the future, a key challenge is to develop a public health policy that seeks to identify and validate effective interventions and then use the research as the basis for action.

In any situation, opinions and observations may be the only evidence available and it is not realistic to refrain from attempting to intervene in areas of public health lacking definitive evidence (see Box 5). In time, hopefully, better evidence will become available.

Box 5. When evidence does not lead to effective action

1. It has been clearly demonstrated that treatment for diabetes mellitus can prevent premature mortality and morbidity. However, the majority of diabetics are not adequately treated, despite this clear evidence. Several programs have attempted to address this problem, but no system of organizing screening and treatment of diabetics has been shown to be better than any alternative.
2. The example of tobacco and health clearly shows the heterogeneity in handling evidence (interpreting the health effects and the effectiveness of control interventions) and the unequal implementation of strategies for controlling this problem (aspects of social participation, social groups, professionals in leadership position—for example doctors, priority among political choices, conflict of interests arising from economic interests). The attempts of the tobacco industry to discredit the large amount of evidence available are particularly instructive.

The adoption of speed limits and the reduction of mortality and injuries in traffic accidents is another good example that shows the influence of dimensions such as perception of risk, conflict among different social benefits, economic development, and health.

Box 6. What to do when evidence is lacking or weak?—The example of environmental health

Research is often lacking to support environmental health interventions. There are several reasons for this lack, including the difficulties of the research, the complexity of the interventions, and the need for long-term follow-up to demonstrate harm from the environmental exposure or the prevention of harm after the intervention. Because of the special nature of environmental exposures, when the population often has no choice in accepting the exposure, programs are implemented without firm evidence of harm. This is the so-called “precautionary principle” in widespread use in this field.

In summary, actions must often be taken in the absence of validated evidence to prevent possible health problems in the population.

Research results in conclusions, which can be the basis of recommendations. One can judge the basis of recommendations to adopt a certain policy following a framework similar to this (23):

- Supported by systematic, empirical evidence
- Supported by cogent argument
- Scale of likely health benefit
- Likelihood that the policy would bring benefits other than health benefits
- Fit with existing or proposed government policy
- Possibility that the policy might do harm
- Ease of implementation
- Cost of implementation

This framework would obviously require a much broader range of evidence than experimental, quantitative evidence (4).

SOME ADDITIONAL CHALLENGES CONCERNING EVIDENCE AND PUBLIC HEALTH

Review of several documents related to evidence for public health and population health (7;20; 28;40) points to the need for further investigation of some important issues, including:

- The transferability of research evidence, especially when the cultures of the researchers and the users are quite different; unquestionably, the context for application of the evidence is highly important;
- The adequacy of methods for demonstrating the links between complex interventions in the “new public health” and the sought-after outcome;
- The adequacy of methods for judging utility and transferability of evidence, especially when a randomized trial or other experimental study is not feasible. In particular, the complexity of many public health strategies, cutting across disciplines and fields of action, may make simple and universal rules of evidence impossible;
- The adequacy of methods of summarizing evidence, especially when it consists of results of quantitative and qualitative research, as well as experience and opinion. A particular problem is the hierarchy of evidence, which may not apply to multidisciplinary research concerning complex public health strategies;
- The adequacy of methods for investigating the dissemination and implementation of evidence into public health practice.

Moreover, the impact and practical consequences of public health research need proper assessment (24).

CONCLUSION

In the context of population health, different statements require certain types of defense. Opinions of respected authorities are helpful but seldom sufficient. Evidence from appropriate empirical research is considered superior.

The best evidence rule expects the professional to use the best available evidence, not the best possible or best conceivable evidence. The best available evidence may not be good evidence. It can be biased, willingly or unintentionally, or even fabricated.

Studies produce data, not evidence. Evidence emerges only after a careful deliberation of the quality of the design and conduct of the study. Evidence is based on data, professional judgment, and common sense. Professional judgment can be guided by formalized procedures. Checklists for evaluating the quality of studies can be helpful, especially in the context of evidence reports, systematic reviews, meta-analyses, and health technology assessments. They are available for many types of studies, but generally not for qualitative research.

Convincing evidence is also a matter of presentation and rhetoric. This is one reason why uncertainty in judgments and decisions cannot be totally avoided. Another reason, more relevant, is the probabilistic character of any “evidence-based” statement. To successfully defend a statement does not mean that it is verified. It means only that it presently cannot be judged to be false.

Asking for, presenting, questioning, and accepting evidence occurs in social (interactive) situations, embedded in distinct social contexts, determined by history, traditions and culture, economic interests, political goals, public perceptions, first principles, emotions, and so forth. Therefore, evidence has to be contextualized.

From the scientific literature, it is clear that evidence is not very solid in many areas of public health and population health. The evidence base certainly needs to be improved. Most likely, the use of available evidence also needs to be improved. There does seem to be a consensus that evidence-based public health is less developed than evidence-based medicine (7;10). Developments in public health lead in the direction of facing this challenge (13).

The term “evidence” is not defined in this study. The primary issue in such a definition is whether knowledge gained from observation or experience can be termed evidence. The Committee agrees that the quality of the evidence is a critical consideration and that evidence from well-designed research is likely to be more valid than “evidence” from informal sources. In addition, the Committee agrees that attempts to determine the quality of the evidence are very important. Furthermore, the Committee agrees that improving methods of synthesis from different areas of research is an important challenge for the future. Therefore, the issue concerning whether the results of experience or observation is or is not evidence is not particularly important. The most important issue is to accept that available evidence based on research must be used in making decisions in the field of public health.

Quantitative, epidemiological research is highly important and very valuable for public health policy. Qualitative research is also necessary and should usually be used in a complementary manner with quantitative research.

Furthermore, it is important to recognize that the scientific evidence (including efficacy and effectiveness of the interventions) is only a component within the complex environment of the decision-making process. There are many factors that condition any decision, such as issues of a political nature, social demand and support, leadership, lobbies, correlation of influences, and power. There is also useful information about the context and the process of taking of decisions, and this evidence is usually generated by methods within the realm of the social and behavioral sciences.

An important challenge is how to implement interventions that have been scientifically demonstrated to have important potential benefits for health. A variety of approaches and a variety of evaluation methods are certainly essential here (30). There is also the need of fostering research where this is more needed (for example, priority-setting in health research).

In summary, the issue of evidence in public health is quite complex. No existing model is adequate to the task of answering all the important questions concerning evidence and public health. Enough is known, however, to make considerable progress without waiting for methodological developments.

REFERENCES

1. Bero LA, Montini T, Bryan-Jones K, Mangurian C. Science in regulatory policy making: case studies in the development of workplace smoking restrictions. *Tob Control*. 2001;10:329-336.
2. Brownson RC, Gurney JG, Land GH. Evidence-based decision making in public health. *J Public Health Manag Pract*. 1999;5:86-97.
3. Davies H, Laycock G, Nutley S, et al. A strategic approach to research and development. In: Davies H, Nutley S, Smith P, eds. *Evidence-based policy and practice in public services*. Bristol: Policy Press; 2000:229-250.
4. Davies H, Nutley S, Tilley N. Debates on the role of experimentation. In: Davies H, Nutley S, Smith P, eds. *What works? Evidence-based policy and practice in public services*. Bristol: Policy Press; 2000:251-275.
5. Davies P. Contributions from qualitative research. In: Davies H, Nutley S, Smith P, eds. *What works? Evidence-based policy and practice in public services*. Bristol: Policy Press; 2000.
6. Drummond M, O'Brien B, Stoddart G, Torrance G. *Methods for the economic evaluation of health care programmes*. New York: Oxford University Press; 1997.
7. Eriksson C. Learning and knowledge-production for public health: A review of approaches to evidence-based public health. *Scand J Public Health*. 2000;28:298-308.
8. Frank JW. Public health policy and the quality of epidemiological evidence: How good is good enough? *J Public Health Policy*. 1985;6:313-321.
9. Fulop N, Allen P, Clarke A, Black N. Issues in studying the organisation and delivery of health services. In: Fulop N, Allen P, Clarke A, Black N, eds. *Studying the organisation and delivery of health services*. London: Routledge; 2001:1-23.
10. Glasziou P, Longbottom H. Evidence-based public health practice. *Aust N Z J Public Health*. 1999;23:436-440.
11. Granados A, Jonsson E, Banta HD, et al. EUR-ASSESS project subgroup report on dissemination and impact. *Int J Technol Assess Health Care*. 1997;13:220-286.
12. Gray J. *Evidence-based healthcare*. London: Churchill Livingstone; 1997.
13. Gray M. Evidence-based public health—what level of competence is required? *J Public Health Med*. 1997;19:65-58.
14. Greenhalgh T, Taylor R. Papers that go beyond numbers (qualitative research). *BMJ*. 1997;315:740-743.
15. Hammersley M, Atkinson P. *Ethnography: Principles in practice*. London: Routledge; 1995.
16. Hatch M. *Organisation theory: Modern symbolic and postmodern perspectives*. Oxford: Oxford University Press; 1997.
17. Haynes RB. What kind of evidence is it that evidence-based medicine advocates want health care providers and consumers to pay attention to? *BMC Health Serv Res*. 2002;2:3.
18. International Union for Health Promotion (IUHP). *Evidence of health promotion effectiveness. Shaping public health in a new Europe*. Vanves, France: IUHP; 1999.
19. Jenicek M. Epidemiology, evidenced-based medicine, and evidence-based public health. *J Epidemiol*. 1997;7:187-197.
20. Kamper-Jorgensen F. Knowledge-base, evidence and evaluation in public health. *Scand J Public Health*. 2000;28: 241-243.
21. Kvale S. Introduction. In: Kvale S, ed. *Issues of validity in qualitative research*. Lund, Sweden: University of Lund (Studentlitteratur); 1989.

22. Liberati A, Sheldon TA, Banta HD. EUR-ASSESS Project Subgroup report on methodology. Methodological guidance for the conduct of health technology assessment. *Int J Technol Assess Health Care*. 1997;13:186-219.
23. Macintyre S, Chalmers I, Horton R, Smith R. Using evidence to inform health policy: case study. *BMJ*. 2001;322:222-225.
24. Martin-Moreno J. The rating of public health research: a complex but necessary exercise (editorial). *J Epidemiol Biostat* 2001;6:415-416.
25. Maxwell J. Understanding and validity in qualitative research. *Harv Educ Rev*. 1992;62:279-300.
26. Mays N, Roberts E, Popay J, eds. *Synthesising research evidence*. London: Routledge; 2001.
27. McKee M. Epidemiology in the 21st century. The challenges ahead. *Eur J Public Health*. 2001;11:241-242.
28. McQueen DV. Strengthening the evidence base for health promotion. *Health Promot Int*. 2001;16:261-268.
29. Miettinen OS, Henschke CI, Pasmantier MW, Smith JP, Libby DM, Yankelevitz DF. Mammographic screening: no reliable supporting evidence? *Lancet*. 2002;359:404-405.
30. Neuberger J. Where's the evidence? Making the case for public health. *J Epidemiol Community Health*. 2001;55:77-78.
31. Nutley S, Webb J. Evidence and the policy process. In: Davies H, Nutley S, Smith P, eds. *What works? Evidence-based policy and practice in public services*. Bristol: Policy Press; 2000:13-41.
32. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet*. 2002;359:909-919.
33. Oakley A. *Experiments in knowing: gender and method in social sciences*. Cambridge: Policy Press; 2000.
34. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet*. 2001;358:1340-1342.
35. Ovretveit J. *Action evaluation of health programmes and changes*. Aabingdon, UK: Radcliffe Medical Press; 2002.
36. Ovretveit J. Medical managers can make research-based management decisions. *J Manag Med*. 1998;12:391-397, 322.
37. Popper K. *The logic of scientific discovery*. London: Hutchinson; 1959.
38. Rootman I, Goodstadt M, McQueen, et al. *Evaluation in health promotion: principles and perspective*. Copenhagen: WHO Regional Office for Europe; 2001.
39. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276:1172-1177.
40. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*. 2002;56:119-127.
41. Rychetnik L. In defense of flexible qualitative researchers. *Aust NZ J Public Health*. 2001;25:471; discussion 471-472.
42. Sackett D, Richardson W, Rosenberg W, Haynes R. *Evidence-based medicine: How to practice and teach EBM*. London: Churchill Livingstone; 1997.
43. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. *BMJ*. 1996;312:71-72.
44. Supplement to American Journal of Preventive Medicine (SAJPM). Introducing the guide to community preventive services: Methods. First recommendations and expert commentary. *Am J Prev Med*. 2000;18:35-43.
45. Task Force on Principles for Economic Analysis of Health Care Technology. Economic analysis of health care technology: a report on principles. *Ann Intern Med*. 1995;122:61-70.
46. U.S. Preventive Services Task Force. Report. Guide to clinical preventive services, 2nd edition. Baltimore: Williams and Wilkins; 1996.
47. Walshe K, Rundall TG. Evidence-based management: from theory to practice in health care. *Milbank Q*. 2001;79:429-457, IV-V.
48. WHO Regional Office for Europe. *Developing public health in the European Region*. Copenhagen: WHO Regional Office for Europe; 1998.

49. WHO Regional Office for Europe. *Regional Committee paper. Strengthening public health in the European Region*. Copenhagen: WHO Regional Office for Europe; 1998.
50. Wodak A. What value is public health? *Aust N Z J Public Health*. 2000;24:114-115.
51. Woolf SH, Henshall C. Health technology assessment in the United Kingdom. *Int J Technol Assess Health Care*. 2000;16:591-625.

ANNEX 1

Types of statements made in public health and the empirical evidence to corroborate them. *Descriptive/comparative statements on health states/events*

“In X, the point prevalence of back pain is 40 percent”, “it is 10 percent higher than in Y”

Statements as such are to be defended by results from valid health and morbidity surveys.

Statements on social forces and existing regulations (Benjamin et al., 1995)

“It is contrary to political declarations, violates local assumptions, poses severe problems to health insurances and pension funds”

To be defended by document analyses, expert interviews, and qualitative research (sometimes a survey may be helpful).

Diagnostic statements

“Community X shows all signs of social deprivation”

To be defended by results from diagnostic studies based on valid taxonomies, criteria, tests, and rules.

Prognostic statements

“If nothing happens the situation will deteriorate within the next 3 years”

To be defended by results from observational or analytic cohort studies with relevant outcomes.

Causal statements

“The situation is caused by a mixture of unfavourable working conditions, high unemployment rate, recent sociolegal changes, and the influence of mass media”

To be defended by results from (mostly) observational analytic epidemiological research (cohort, case control studies) with sufficient internal and external validity.

Intervention statements

“A mass media campaign with TV spots, articles in newspaper and public advertisements will reduce disabling back pain prevalence by 5%, this will be cost-efficient”

To be defended by results from experimental epidemiological studies (field studies) or population based before-after studies (and analogy), cost analyses.

Implementation statements

“The successful Australian intervention will not work in our community/region”

To be defended by “cogent arguments” based on earlier experiences, analogies, results from qualitative research on barriers/supportive factors.

Outcomes statements

“To our surprise, our goals have been reached”

To be supported by descriptive/comparative research (see above).