Stefano Campostrini[1], **David V. McQueen**[2]

[1] *Department of Applied Statistics and Economics, University of Pavia, Italy*
[2] *National Center of Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA, USA*

Institutionalization of social and behavioral risk factor surveillance as a learning system

Summary

Objectives: This paper extends further the systematic nature of social and behavioral risk factor surveillance by showing how it becomes a learning system. The authors argue that such a systematic approach will lead to a better informed public health practice and a better understanding of key public health concerns such as obesity.
Methods: This paper is based on methods developed over some 25 years of experience in designing and operating behavioral risk factor surveillance systems. Measurement issues still remain a key concern and special attention is paid to the role of time as a critical variable in developing a surveillance system.
Results: The paper lays out five critical areas for attention if a system is to be considered a learning system. Examples are given from ongoing surveillance systems that have developed a base for a learning system and notes how these are institutionalized.
Conclusions: In the ideal case a learning system based on socio-behavioral surveillance is dependent on data being collected continuously with careful attention paid to the analytical challenges of such complex data. This is an ideal systematic approach that has not been realized in most surveillance systems now available.

Keywords: **Time – Surveillance – Analysis – Learning system.**

By the end of the 20th century a number of survey-based approaches to assessing behavioral risk factors existed across the globe (McQueen & Puska 2003). However, if one searches for comprehensive surveys on behavioral risk that were carried out in a surveillance mode, the number declines rapidly. In 1968 the World Health Organization (WHO) defined surveillance as the "systematic collection and use of epidemiological information for the planning, implementation, and assessment of disease control" (Teutsch & Churchill 2000: 72). Teutsch and Churchill consider the definitional area of surveillance and also take up the broader term monitoring. They write, instructively, "…some prefer to use the term monitoring to refer to the entire set of efforts to gather information about a population or its environment for health-related purposes and to restrict the use of the term surveillance to rapidly identifying and monitoring individuals and problems that pose an immediate threat to the health of the public."

Controversy continues to surround the definition of these terms and the discussion takes on slightly different overtones when one considers social and behavioral risk factor surveillance. In this paper it is important at the outset to clarify our perspective on surveillance for social and behavioral risk factors. This requires the attention to the other word of importance, namely "system". For our purposes a system can be defined simply as a group of independent but interrelated elements comprising a unified whole. Finally, from our perspective, time is of critical importance in surveillance.

What is meant by "time" in the context of social and behavioral risk factor surveillance? First, data should be collected over time as close to continuously as possible (McQueen et al. 1992). That is, in an ideal situation data would be collected constantly on a 24 hour basis seven days a week. Such a scenario would create a continuous stream of data. In a "true" surveillance at no time would the data collection machine or instrument be shut off. In the reality of practice, however, this scenario is not feasible. Even security cameras rarely use continuous photography, but rather a series of photos are taken at equal intervals of time. And, so called streaming video or a motion picture is really a series of

Soz.- Präventivmed. 50 Suppl 1 (2005) S9–S15
0303-8408/05/010S9–7
DOI 10.1007/s00038-005-4104-4

single photographs that are time bound, and, in theory, an "event" could occur in the interspace between photo stills. However, with behavioral survey techniques, it is feasible to collect data on a continuous basis by clearly defining the parameters of the unit of time and then systematically repeating that unit of time over a long duration. The U.S. Behavioral Risk Factor Surveillance System (BRFSS) is an excellent model of such a systematic surveillance approach because it collects data over many years with a one month time frame (http://www.cdc.gov/brfss/). We would assert that this comes remarkably close to a "true" continuous behavioral risk factor surveillance system.

What are the basic components of social and behavioral risk factor surveillance? First one must have the necessary knowledge base to understand the appropriate field of inquiry and the ability to construct reliable and valid instruments for data collection. Fortunately this area is highly developed with an extraordinary literature base from years of survey research and surveillance practice. For example, in the collection of data on smoking behavior there is an extensive compilation from many countries and many surveys of how to assess validly and reliably smoking practice through self-reported questionnaires. Similarly, in the field of sociology there is a long established tradition of how to understand and assess socio-economic status through questionnaires. That is not to say that an unquestioned field of inquiry exists. On the contrary, one reason for discussing social and behavioral risk factors for disease is that there is a considerable range of appropriate topics for inquiry. In the beginning of the BRFSS the emphasis was largely epidemiologically based. This emphasis reflected the state of the field at the time and also the epidemiological origins of surveillance in public health. Not surprisingly, the field of relevant behaviors largely reflected the behavioral causes of morbidity and mortality, namely factors such as smoking, alcohol abuse, poor dietary practices, and lack of physical activity. The field of inquiry has broadened greatly in recent years to include more social-cultural variables such as quality of life and social capital. Questionnaires used over the years have developed and followed this expanding conceptual base for surveillance of risk factors. More importantly, questionnaires have increasingly been driven by demands by end users of the system, a point to which we will return later.

A second key component of surveillance is the data. As with questionnaire construction, the attention given in the scientific literature to data type and its collection has been enormous. Extensive attention has been given to measurement issues, particularly in those areas that are challenging because of the dynamic nature of the behavior being measured (cf. Dean 1993). For example, an international effort has been underway to adequately measure physical activity. The challenge has been to derive a measure that adequately reflects the total physical activity of an individual throughout the day. No simple solution has been found, but multiple measurement attempts through appropriate questions have led to a widely accepted set of questions to ask (e.g. the International Physical Activity Questionnaire, (IPAQ) http://www.ipaq.ki.se/). The other key piece related to the collection of data has been on the subject of sampling. Despite general agreement on the need for probability samples, actually collecting such samples presents many challenges, especially in developing countries. In recent years the sampling issue has become more complicated because of the changing technology. Nonetheless, the primary issue is to produce good enough quality data on which to base decisions. Sample size, particularly the sample size produced for smaller time intervals demanded by continuous data collection, represents challenges to traditional sampling methodologies. However, a number of sampling challenges may be addressed by better analytical efforts following data collection.

The third key component of surveillance is the area of analysis. In our view this is an area that needs further attention. Despite the considerable statistical literature on analyzing data, this area suffers from three serious drawbacks. The first is the relatively few analytical techniques that have been applied routinely to surveillance data. We possess meters of shelf space occupied by reports that have applied only descriptive statistics to the data. Thus, the richness of multivariate statistical analyses has not been sufficiently applied to surveillance data. The deficit is compounded by the use of data analysis techniques designed for static, single survey data, which are not only potentially inappropriate for the analysis of time-related data but also fail to take into account the richness of the data available (Campostrini 2003). Finally, most unfortunate is the failure to take up new and challenging data analysis techniques arising notably from physics and other disciplines concerned with chaotic and nonlinear data. The reasons for this shallow use of analytic techniques are, like surveillance data itself, complex. Certainly, it is a mixture of training, resources and personnel. Innovation has not been a characteristic of the analysis of surveillance data to date, with the possible exception of the use of geographical and spatial analyses, which have offered some spectacular results.

A fourth key component of surveillance is interpretation. This has usually been seen primarily as a statistical problem. However this is a narrow approach. Interpreting data as important or significant simply by pointing out the result of a statistical test or that a certain probability of significance has been reached is simply not adequate in our view. To begin

with, one must understand that notions such as level of significance are themselves simply an historical product of a moment in the development of statistics and adopted by biostatistics and public health practitioners. The fact that such notions became rote expectations for the interpretation of data is most unfortunate particularly with behavioral data. Inadequate or poor interpretation of data cannot be easily separated from problems of appropriate analysis because they go hand in hand. Interpretation of social and behavioral data requires both statistical and theoretical sophistication. The theoretical part is the component that takes data from mere description to interpretation. We can only reiterate the saying that we have a "vast stock-pile of almost surgically clean data untouched by human thought." (Attributed to Alwyn Smith (1985) of Manchester, England.)

The fifth and final key component in social and behavioral risk factor surveillance is perhaps the most important and historically the most neglected. That component is data use. This concept goes beyond interpreting the output of surveillance systems and gets to the very heart of current issues in contemporary health promotion and public health practice. Use of surveillance data has been one of the main themes in two special conferences on behavioral risk factor surveillance, one in Savannah, Georgia, in 2000, and more recently at the Third International Monitoring Conference in Noosa, Australia. Much of the background on this is well covered in a number of recent publications and current thinking is summarized by Ottoson and Wilson (2003), paraphrasing Canadian work on guiding principles for enhancing utilization. They note that concern about use of surveillance data needs to shift from an afterthought to a forethought. This changes the use question from: "Now, that we have these data, how can we get people to use them" to a new question: "If surveillance data are collected, who will use them and how?" This relates to a guiding principle to involve potential users in decisions about the design, content, and interpretation of surveillance systems. Such involvement not only helps sharpen the relevance of data and findings but also encourages an ownership in the system that facilitates use. Related to this is the recognition that surveillance data are not the only influence on decisions related to community health interventions. This is not to diminish the importance of good quality data, rather to recognize that it is one of many influences on decision making.

From components to a system

All five components – knowledge base, data, analysis, interpretation and use – are critical for surveillance that constitutes a system for surveillance. However there are several characteristics of a surveillance system that make it unique. Most importantly is a fundamental understanding of a dynamic approach to each of the five components; four of these components are illustrated dynamically (see Fig. 1). There is nothing static about this approach. For example, in a static approach, one would first design a questionnaire, plan and draw a sample, collect the data, analyze the data, and then publish a final report along with academic articles. Each of these steps could be carried out independently and in a linear fashion. That is, once the data were collected, that phase would end and often the same staff would move on to analyze the collected data. This approach represents a linear process that is not dynamic. In a dynamic surveillance system, all the components are occurring simultaneously, and they are also highly related to each other. Furthermore, such a dynamic approach is easily conceptualized as a learning system because each component feeds into and relates to another component. Thus, the whole system "learns" about its operations.

A system can be either closed or open. A closed system is an isolated system having no interaction with the larger environment in which it resides (Von Bertalanffy 1968: 3). A surveillance system can operate chiefly as a closed system, and a critique to date may be that this has been a limitation. However, it is our position that when a surveillance system moves to an open system, it allows for its institutionalization into the larger public health institution because its learning system is emphasized. Thus it becomes part of a complex system with all the attendant characteristics such as the possession of a number of components, a diversity of components, linkages between the components, and a synergistic relationship among the components, which is nonlinear and complex.

The organizational learning perspective

An interesting approach that could make clearer the role of behavioral surveillance and how it could work as a system is that of organizational learning. The basic idea is that for an organization to carry out its mission, it should learn from

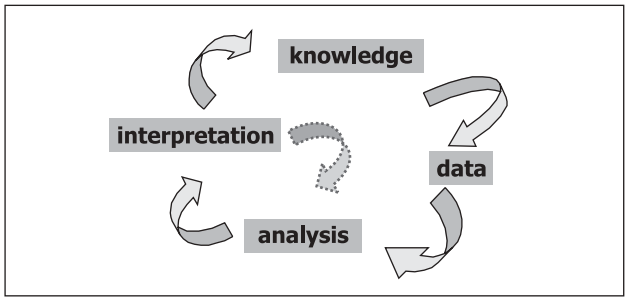

Figure 1 The surveillance components in a dynamic approach

Soz.- Präventivmed. 50 Suppl 1 (2005) S9–S15

innovations that arise and that attention should be paid to the mechanism for an organization to learn. "In order to become organizational, the learning that results from organizational inquiry must become embedded in the images of organization held in its members' mind and/or in the epistemological artifacts embedded in the organizational environment" (Argyrols & Scion 1996). Taking this perspective would help enhance the effectiveness of both the surveillance systems and the other systems that relate to surveillance. Before presenting how the concept of organizational learning could be applied to surveillance systems, it is necessary to discuss the role and the components within a surveillance system and the "other systems" that relate to it.

The system, the links with other systems, and the learning system perspective

Behavioral risk factor surveillance systems can be seen as information systems (IS), because the data they produce can be translated into information. Although it is clear that raw, unanalyzed data files are simply collections of numbers, they can be transformed through appropriate analysis and interpretation into information for policy and decision-making processes. This whole process can be interpreted as an information system. However, it is important to realize that an information system is not independent of the sociopolitical and organizational context in which it resides. Considering this, the immediate questions are "which context" and "which organization structure" should the behavioral risk factor surveillance support as data are prepared for use? To answer these questions we should consider the different inputs and outputs that a behavioral risk factor surveillance system can receive and produce.

The first macro-organization, that is the primary system to which a behavioral risk factor surveillance system is linked, (see Fig. 2) is that of the infrastructure for public health. In this regard, behavioral risk factor surveillance systems function to inform public health and support public health decision making. At the same time, following the idea of a learning system, behavioral risk factor surveillance systems get the knowledge (theory, rationales, but also hypothesis and questions) from the public health system itself.

Let us simply define the context in which public health practice takes place as a public health system. Of course it is not the only system that a behavioral risk factor surveillance system relates to. Data (and information) produced by behavioral risk factor surveillance systems can be used directly by the broader society, as well as other subsystems of that society. The recent impact of obesity trend data from the U.S. BRFSS on the media is a salient example of the BRFSS information going into another broad system of information [1]. As shown in the Figure 2, the BRFSS is not only embedded in the other systems to which it is related, but also from which it originated. From an organizational learning perspective, this embedment can be turned into a position of advocacy for risk factor surveillance.

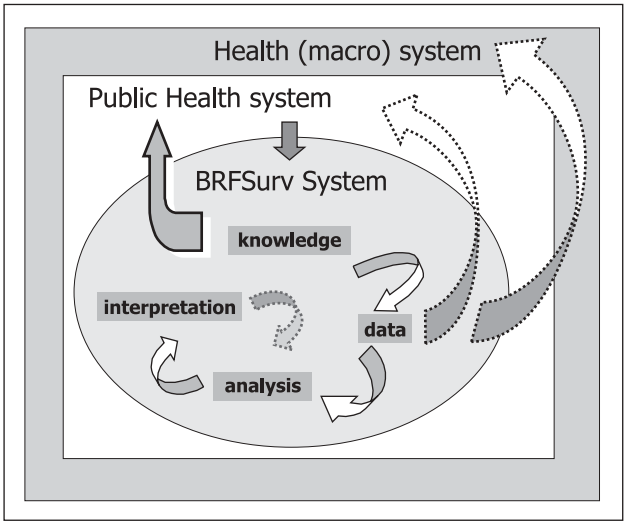

Figure 2 The surveillance system and the link with the other systems

Why advocate for a risk factor surveillance system? First, behavioral risk factor surveillance systems represent an "innovation" that can help public health, and the broader health system to "learn" from the information they produce. It should be noted that the fact of being embedded with the other systems is essential for them to learn and proceed. As already pointed out, to be effective the learning that results from the innovation should become embedded in the organization itself (Argyrols & Scion 1996). This learning is feasible only if the link among these systems – these macro-organizations – is well defined. On the other hand, a behavioral risk factor surveillance system can be considered an organizational entity that can, or should, learn from the knowledge it produces or from the knowledge that comes from the other systems of which it is part. This process of communication, knowledge and information exchange is crucial to the innovative learning component of the surveillance system.

The need for institutionalization

Our position is that behavioral risk factor surveillance systems must be a part of the public health system if they are to produce informative effects and provide a continuous learning process. However, it is important to emphasize that a continuous learning process implies that time, as an element, is critical. In contrast to the static nature of traditional surveys,

it is the time component of surveillance that makes the continuous learning process possible. It is also important to understand that isolated point-in-time surveys, even if routinely conducted at intervals of time that appear approximately equal, are not good models for a learning system. A learning system should be built into real experienced time if it is to provide rapid feedback and be of most use to the end users.

Besides the critical importance of a time component, one needs to consider the vitally important relationships with the other systems to which the behavioral risk factor surveillance system belongs. And because it is not possible to link two systems if both are not settled and stable, the need for the institutionalization of surveillance is evident. To institutionalize means to settle the system in a way that it has the capacity to sustain itself over a long period of time and have a future. Institutionalization is often seen as essentially a problem of financial resources, but such resources are only sustained once a system is institutionalized and legitimized as part and parcel of the underpinning expectations of a public health system. It is a matter of developing relationships, networking, and, most critically, of facilitating the use of data and information.

Quite often in the past too little attention has been paid to these aspects of institutionalization, and the dissemination of information from surveillance ended after the production of an official report. Now we recognize that if we want a surveillance system to be effective, it must include elements such as networking, social marketing, and information brokering. In general, a surveillance system must include a marketing strategy that is designed to proactively inform consumers. In fact a marketing strategy should be embedded in the surveillance system.

After nearly three decades of behavioral risk factor surveillance experience what has emerged is a clear image that a surveillance system is in many ways broader than a string of traditional "surveys" limited to a technical effort to produce data as its chief justification for existence. In the new conceptualization of surveillance systems many people are involved, not just the "data making" people. Among the personnel "newcomers" three groups are vitally important:

- *Data users* – the persons (practitioners, journalists, the man on the street) – who use surveillance data for many purposes. These persons could be considered the "customers" of surveillance.
- *Policy makers* – the persons who are the major potential "client" of surveillance. These individuals ultimately influence the building of the resource infrastructure for surveillance systems.
- *Researchers* – those persons who are interested in the data analysis and the critical interpretation of findings.

An example of a (learning) surveillance system in practice: the U.S. BRFSS and the obesity problem

To better understand surveillance as a learning system, consider the following example. One of the major public health problems in the United States is obesity, particularly the rapid growth of this problem. Fortunately, the development and history of the "obesity epidemic" has been documented by the world's largest and most developed behavioral risk factor surveillance system, the U.S. BRFSS (http://www.cdc. gov/brfss). Some factual background:

a) It is well known that obesity is an increasing problem in the United States: the national objective for the year 2010 is that of lowering obesity to the level of 15% in the adult population (present estimates range between 22% and 23% and the estimates are increasing [2]).
b) BRFSS time trend analysis estimates for the year 2010 a level of obesity that ranges from 19 to 35%, depending on population subgroups considered in the analysis [3]. Although starting from different levels, each population subgroup (grouping by gender, education, race, and income) shows a steady and stable increase in BMI during the past several years. If this trend is maintained none of these subgroups will fulfill the national health objectives for the year 2010.
c) More information is needed to identify the mechanisms of obesity, such as which behaviors and social-structural conditions may "cause" obesity, and to develop effective strategies for public health interventions.

These "facts" can be re-interpreted, re-assembled and viewed in light of a learning surveillance system perspective.

Knowledge coming from the public health system: obesity is a problem.

Data use: obesity is increasing.

Data analysis: obesity will increase in all population subgroups.

These components are the inputs and the outputs of the surveillance system and at this point the surveillance system could stop: nothing more, apparently could be said. But, if we allow the system to function as a learning process, it can be operated to gain new knowledge about the obesity behavioral mechanism, specifically it can respond by addressing three needs:

- *Need for new analyses.* By analyzing the relationships between the several behaviors and environmental conditions that underlie the obesity trend, some important new dimensions could better characterize those behaviors and socio-behavioral contexts that favor an observed increase in obesity.

– *Need for further interpretation.* Results from deeper analyses will need better interpretation by trying to link (conceptually and not statistically) what relates to current theories about obesity and people's behavior.
– *Need for new data.* Theory, analysis, and interpretation could then point out the lack of data collected by the present BRFSS so that data about other types of information can be identified and collected, such as measuring the effects of urban sprawl or non leisure time physical activities, or by developing more effective ways of measuring expenditure of caloric energy.

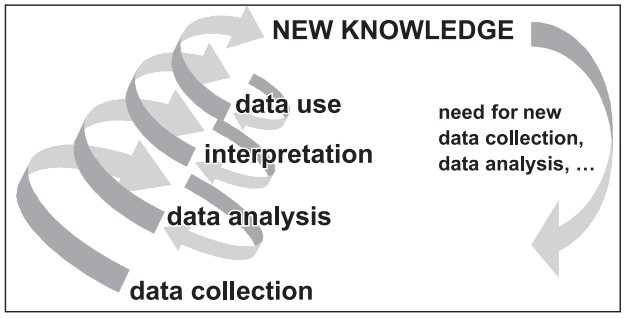

Figure 3 The 'SPIRAL' of the surveillance system

Final comments: the spiral of surveillance

Following the brief example presented above, we can recognize several important characteristics of an ideal surveillance system. First, it can learn from the knowledge produced by it and by the systems it is linked to. Second, the knowledge that is produced is always incremental and occurs in a time-bound structure. Third, the knowledge occurs in a spiraling way starting with the data collection, continuing with data analysis, which involves interpretation, and finally producing new knowledge (Fig. 3).

As new knowledge is produced it may be seen as the starting point for the loop illustrated in Figure 3, but this new starting point is above the previous loop, because we are starting at a higher level of knowledge. That is why selected the spiral to illustrate the underlying concept of this surveillance system: something that grows in a circular way, going forward but moving always to a higher position. This movement also represents the learning component, and it is characteristic of a dynamic system that is in continuous growth, which is why continuous data collection is such a fundamental part of a learning system approach to surveillance.

Starting from some acquired definitions and classical rationales for surveillance, we have argued for a relatively new perspective – the learning system. We have emphasized the systemic aspects related to the surveillance by discussing the reasons why it is important, crucial, and essential, to pay attention, not only to the core of the system (data collection and analysis) but also to the other elements that constitute this system. We believe that the learning system perspective could help in defining the role and the structure of surveillance by seeing the system as an essentially dynamic process that fully uses the advantages of continuous data collection.

NOTES

[1] A good example of how important a communication strategy can be on data use comes from the U.S. BRFSS. Over many years, the BRFSS used different media (written reports, academic articles, websites) to release data and analyses, but not much attention was given to the obesity figures. However, a new understanding of obesity occurred when these data were transformed into a series of maps in which the color of U.S. States changed following the increase in the percentage of the obese population. With this simple graphical help, the same data that had been reported in numerous written reports became more visible and more informative. In fact, once the data were released in the pictorial format, the news media paid a lot of attention to the information.

[2] In 2002 BRFSS estimated 22.1 % among the adult population as obese (respondents 18 and older who report that their body mass index (BMI) is 30.0 or more) in the year 2002 – source: BRFSS web site: http://www.cdc.gov/brfss/

[3] Analyses made by the authors on the data published in the quoted BRFSS web site.

References

Argyrols C, Scion D (1996). Organizational learning II: theory, method and practice. Reading, Mass: Addison Wesley.

Campostrini S (2003). Surveillance systems and data analysis. In: McQueen DV, Puska P, eds. Global behavioral risk factor surveillance. New York: Kluwer: 47–55.

CDC, BRFSS web site: http://www.cdc.gov/brfss/

Dean K (1993). Population health research: linking theory and methods. London: Sage.

McQueen DV, Uitenbroek DG, Campostrini S (1992). Implementation and maintenance of a continuous population survey by CATI: proceedings of the 1992 Bureau Census Research Conference, Washington, D.C.

McQueen DV, Puska P (eds.) (2003). Global behavioral risk factor surveillance. New York: Kluwer

Ottoson JM, Wilson DH (2003). Did they use it? Beyond the collection of surveillance information. In: McQueen DV, Puska P, eds. Global behavioral risk factor surveillance. New York: Kluwer: 119–131.

Soz.- Präventivmed. 50 Suppl 1 (2005) S9–S15

Smith A (1985). The epidemiological basis of community medicine. In: Smith A, ed. Recent advances in community medicine 3. Edinburgh: Churchill Livingstone: 1–10.

Teutsch SM, Churchill RE (eds.) (2000). Principles and practice of public health surveillance. 2nd ed. Oxford: Oxford University Press.

Von Bertalanffy L (1968). General systems theory: foundations, development, applications. New York: Braziller.

Adress for correspondence

Prof. Stefano Campostrini
Department of Applied Statistics and Economics
University of Pavia
Corso Strada Nuova, 65
I-27100 Pavia
stefano.campostrini@unipv.it