Lessons Learned from Prevention Programs: Different Endpoints Should Be Used in Secondary and Tertiary Prevention

Franz Porzsolt

Abstract It is mandatory to compare cost and consequences of healthcare services if public support is requested. This request will apply to all healthcare services including prevention. As the demand for health care will always exceed the available resources, methods that make it possible to select the “best” programs for implementation have to be developed. The selection of the “best” programs is not easy because there exist so far no generally accepted quality criteria that can be used to identify the “best” prevention programs.

Based on a model on structural and functional properties of a disease, it is concluded that the traditional outcomes of treatment and prevention may be useful for the evaluation of tertiary prevention programs, but not of secondary prevention programs. Neither the traditional endpoints of treatment studies nor traditional surrogate parameters are useful for the evaluation of secondary prevention programs.

Using the assumptions of the model and a list of available data in secondary prevention programs we recommend to assess five indicators for description of the value of a secondary prevention program: quality of life, surrogates for life expectancy, the perspective of the assessor, the conditions of assessment, and finally the payment. As each of these five items offers two possible values prevention programs may be classified into 32 different groups.

2.1 Introduction

It has long been recognized that economic analyses of prevention programs should be completed (Broudy et al. 1979; Dalziel and Segal 2006; Foster et al. 2003; Holtgrave et al. 1996; Manau et al. 1987; Tager and Sondik 1985; Wang et al. 2006). The special problem of most prevention programs is the long time period between their intervention and outcome. This is probably the reason that most economic evaluations are confined to the description of costs (Foster et al. 2003; Manau et al. 1987), discuss the interventions that are necessary to reduce the risk of disease (Holtgrave et al. 1996 Mahoney et al), or discuss the uncertainty of the economic analysis of prevention programs (Dalziel and Segal 2006; Wang et al. 2006).

The structure and the process of most prevention programs are described in detail but
almost no information is available on the achievement of the finally intended outcomes. The expected outcomes are usually well defined and there are several surrogate parameters, which refer to promising results. The validity of these intermediate results depends on the mode of assessment (under experimental or under day-to-day) conditions. The results may be assessed by healthcare providers or healthcare users. Depending on the assessed topic and on the conditions of assessments, the meaning of the assessed outcome and the value of the prevention program may be quite different. The economic decisions may include several aspects such as financial resources, research manpower, the chance to succeed, and the reduction of ineffective programs applied to our patients.

We shall not be satisfied with preliminary data and with the fact that most prevention programs are dangling their final results. We should provide data which describe the information and the interventions are offered to the healthcare users, i.e., people at risk. We have to record the behavior of this group of people, i.e., the adherence to our recommendations and to document different types of comprehensible outcomes.

This paper describes a Gedanken experiment (Mach 1976; Popper 1968) that generated an algorithm, which can be used to evaluate various prevention programs according to the patients’ perspective.

### 2.2 Methods

The Gedanken experiment is based on three assumptions.

- First, a disease can be detected by structural and/or functional properties. The structures can be described by macroscopic, microscopic, or biomolecular features, i.e., laboratory tests or imaging methods. The functions of a disease can be detected by the effects of a disease on a persons’ quality of life or life expectancy.
- Second, functional properties of a disease are more important than structural properties. Examples are shown in Table 2.1. These examples demonstrate that functional properties of a disease are more important than structural properties: prostate cancer is frequently detected by structural properties, which are detected by digital rectal examination. Many of the detected prostate cancer affect neither the quality of life nor the life expectancy. In patients with cancer of unknown primary the absence of detectable structures does not rule out the disease but may have considerable effects on both quality and quantity of life.
- Third, diagnostic systems, which focus structural but no functional properties of diseases – like many of our staging systems – cannot be used for the assessment of secondary prevention programs.

<table>
<thead>
<tr>
<th>Structural properties of diseases</th>
<th>Functional properties of diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequently detected</td>
<td>Most forms of cancer</td>
</tr>
<tr>
<td>Rarely detected</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>Frequently detected</td>
<td>Cancer of primary unknown</td>
</tr>
<tr>
<td>Rarely detected</td>
<td>Death of unknown reason</td>
</tr>
</tbody>
</table>
These three assumptions provide the basis for the categorization of traditional outcomes. As these traditional outcomes are not really useful for the assessment of the value of secondary prevention programs we propose different criteria for secondary and tertiary prevention. These criteria have to be measurable, available, and interpretable.

2.3 Results

2.3.1 Traditional Outcomes of Treatment and of Prevention Programs

Prevention programs in this paper are structured according to the different phases in the natural history of a disease as described by Baselga and Senn (2008). The expected goals of these programs are the prevention of either the appearance of a disease (i.e., primary prevention). This aspect of prevention is not included in our analysis. Other important goals are the prevention of disease progression (i.e., secondary prevention including screening) or the prevention of the fatal consequences of a disease (i.e., tertiary prevention). When these final goals are critically analyzed six theoretical outcomes can be described (Fig. 2.1).

One of the most frequently observed final outcome of a prevention program is nonadherence due to other priorities. Although nonadherence is rarely quantified it is well known and several strategies were recommended to overcome this problem (Bosch-Capblanch et al. 2007; DeKosky 2006; DiMatteo 2003; Ebrahim 1998; Hiatt 1997; Rockson 2009; Stirratt and Gordon 2008).

The explicitly intended outcome of secondary prevention is cure. It is definitely not difficult to identify the failures of secondary prevention as these patients undergo palliative therapy and will either be able to survive with cancer or have to continue palliative therapy as long as its benefit outbalances its harm. Patients who survive with cancer belong to the group in Table 2.1 in which structures but no functions of the disease are detectable. Patients who need palliative therapy will finally succumb to functions of the disease either without or with detectable structures of the disease.

The identification of patients who will be cured even without prevention is more difficult because nearly all patients with a confirmed positive screening result will undergo treatment. These patients who are cured either with or without treatment belong to the group in Table 2.1 in which neither structures nor functions of the disease are detectable. Except from very few special situations there will be no population in which the “natural course” of the disease can be investigated. Such a “special situation” emerged when mass screening of women for breast cancer was introduced in Norway. Zahl et al. (2008) used a rather intelligent approach to identify the number of breast cancer cases that were detected with and without mammography. As there is positive evidence, which seems to exclude several possibilities to explain this observation (Kaplan and Porzsolt 2008), the most likely explanation favors the interpretation that about 20% of invasive breast cancers that are detected by screening mammography will regress spontaneously (Porzsolt and Hölzel 2009).

In addition to the final outcomes a large series of surrogate parameters is assessed following secondary prevention programs. The assessed surrogates are related to the investigated interventions or are obtained from special tests. Examples of interventions are lifestyle changes or treatments. Suggested lifestyle changes were exercise, weight management, healthy diet, moderate alcohol consumption, and fruit and vegetable intake (Cummings et al. 2009). The recommended treatments include oophorectomy (Metcalfe 2009), treatment with
phytochemicals (Adams and Chen 2009), selective estrogen receptor modulators (Li and Brown 2009; Howell et al. 2008; Powles 2008; Wickerham et al. 2009), retinoids (Bonanni and Lazzeroni 2009), bisphosphonates (Valachis et al. 2010), nonsteroidal anti-inflammatory drugs (Agrawal and Fentiman 2008), or vitamin D (Welsh 2007).

In other studies it is claimed that the risk of breast cancer can be predicted, e.g., by intraductal tests such as nipple aspiration fluid, ductal lavage, mammary ductoscopy, or periareolar fine needle aspiration (Cazzaniga et al. 2009).

In summary, the traditional endpoints of treatment may be useful in tertiary prevention programs but are not helpful for secondary prevention when the success following screening and the subsequent treatment has to be described. None of the assessed surrogate parameters could reliably predict the success of secondary prevention.
prevention. Therefore, it is necessary to search for new indicators of successful secondary prevention programs.

### 2.3.2 Proposal for the Evaluation of Prevention Programs

The traditional endpoints of treatment studies such as cure or survival will only rarely be reported following secondary and tertiary prevention and are therefore considered as theoretical outcomes of prevention programs. From a scientific point of view, it is often not possible to differentiate between spontaneous cure, treatment-related cure, or survival with disease. Therefore, the traditional endpoints are no ideal candidates when measurable endpoints are needed for comparative evaluations especially of secondary prevention programs.

On the other hand there is sufficient information in most secondary and tertiary prevention programs on aspects such as practicability, demand, compliance, and side effects. These criteria are correlated with the final endpoints of secondary and tertiary prevention programs because they all influence the adherence to the program and are signs of program success or failure.

Prevention is usually demanded by patients because it induces hope and confidence. The demand will be correlated at least with short-term adherence, and this can be quantified. In some programs adherence was quantified and may even predict survival (Bauer et al. 2010; Martín-López and Hernández-Barrera 2010), a result which has to be interpreted with care.

The effects of prevention are sometimes assessed under day-to-day conditions and sometimes only under the experimental conditions of a study. Even this information on the condition can be used as indicator for the value of a prevention program.

It can also be assessed if the effects of a program are self-assessed by the patients or are assessed by proxy raters and finally the payer of a prevention program, private or public, can be identified.

This information can be used for the evaluation of prevention programs because it is frequently available, easy to record, and describes at least some aspects of the value of prevention programs.

Considering this information, which is available in most prevention programs, and the lessons that were derived from the three above assumptions, five levels of outcomes that consider

<table>
<thead>
<tr>
<th>Level</th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
<th>(D)</th>
<th>(E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct (1) or no direct (0) effect on quality of life</td>
<td>(1/0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect (1) or no effect (0) on surrogate of life expectancy</td>
<td></td>
<td>(1/0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritized effect assessed under day-to-day conditions (1) or under experimental (0) conditions</td>
<td></td>
<td></td>
<td>(1/0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritized effect is self-assessed (1) or proxy rated (0)</td>
<td></td>
<td></td>
<td></td>
<td>(1/0)</td>
<td></td>
</tr>
<tr>
<td>Prevention is financed by private (1) or public (0) resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1/0)</td>
</tr>
</tbody>
</table>
functional aspects of disease management are proposed (Table 2.2). These aspects are quality of life (A), life expectancy (B), the perspective of the assessor (C), the conditions of assessment (D), and finally the payment (E).

Quality of life (A) refers not only to aspects that are recorded by quality of life instruments but also to any behavior that can be explained as action to optimize the quality of life (Magai et al. 2007). An example for differences in optimizing the quality of life may be the two groups of women who accept or do not accept mammography. Although we cannot be sure that the sender and the receiver of information will perceive the same message at least one of the two target populations, i.e., women who accept mammography, obviously consider the benefit of mammography higher than the psychological stress and discomfort associated with mammography while women who refuse mammography interpret the identical information differently. The acceptance rate of a prevention program may therefore be a useful indicator for the effect of a prevention program on the quality of life. Well-accepted prevention programs are expected to have a more favorable effect on the quality of life than rarely accepted programs.

The possibility to influence the own life expectancy (B) is always an important aspect for decision making. The strong request of screening supports the assumption that patients wish to contribute to improve their life expectancy. As we have begun to understand the factors that influence these decisions (Misono et al. 2008), the number of appropriate surrogate parameters, which indicate positive effects on the patients’ life expectation, will increase. Therefore, the indicators of life expectancy were included in the list of criteria for evaluation of prevention programs.

The perspective of the person who makes the assessment (C) has a considerable influence on the described outcome. Quality of life assessment is an example for the huge differences that can be observed when quality of life is self-assessed or assessed by a proxy rater. Self-assessments are higher rated than proxy ratings also in the evaluation of prevention programs.

The important influence of assessment conditions (D) are not always considered carefully enough. In the scientific literature efficacy and effectiveness are differentiated. Efficacious programs are assessed by healthcare providers under ideal but artificial conditions. Effective programs are assessed by healthcare providers under day-to-day conditions. Although these definitions are used by some groups (Ernst and Canter 2005; Porzsolt et al. 2010) they are not commonly accepted. Effects that are observed under day-to-day conditions are considered more important than effects observed under ideal but artificial conditions.

Finally, we consider the payer (E) of the prevention program. Prevention programs that require public resources are considered less valuable than programs that are privately financed.

Among the five criteria for assessment of the value the measurable effects on quality of life were estimated higher than surrogate parameters for the life expectancy. The effects on these two outcomes were followed by the condition under which the data were generated and recorded because this information will be more often contributed to the evaluation than the information on the rater of the effect. In addition to the information on the rater the information on the payer was considered the least important aspect for description of the value of a prevention program.

2.4 Discussion

2.4.1 The Risk of Interpretation in Prevention Programs

It is not easy to interpret data from prevention programs as the true reasons of the observed effects may be generally expected. So it is commonly
assumed that strong adherence to guidelines or to recommendations of prevention programs will result in extension of life expectancy. There are even grants, which are provided for projects to demonstrate a correlation of adherence to recommendations and increased life expectancy.

The correct interpretation may be quite complicated as it should be excluded that patients with less complications and a less problematic course of the disease will be those who can follow the recommendations more closely than patients with complications or side effects of incompatibilities. In other words, especially in secondary prevention programs there are several possible confounders, which have to be excluded to avoid misinterpretation of the obtained data.

The value of prevention programs from the patients’ perspectives is not identical to the perspectives of other stakeholders. Patients as well as persons at risks of health problems are satisfied if a prevention program restores the lost health or prevents the loss of health. Other stakeholders such as doctors, hospital managers, health insurance companies, or politicians strive for additional goals. None of these partners will be satisfied if only the patients’ goals will be achieved by prevention. The endeavor to reach additional goals is absolutely justified. The emerging risk is associated with a potential imbalance between the patients’ goals and the goals of any other stakeholder. This risk can be avoided if we make sure that the patients’ goals will be the primary goals of any healthcare prevention program and that at least part of the patients’ goals are achieved to justify the expenditures for prevention. The described proposal describes a strategy for the assessment of the patients’ value of prevention programs.

2.4.2 The Resulting Need for Evaluation of Preventive Medicine

The increasing demand for healthcare resources is justified by progress and innovations in health care. This progress and innovations have to be financed. Unless additional value is generated these investments inherit the risk of an economic hazard. The critical question therefore is what should be accepted as “added value” and for whom. These criteria will be different for treatments and tests. It is probably not justified to qualify a test as “added value,” which just identifies a new target population without providing any further evidence, e.g., on the risk reduction that can be achieved in this newly identified population. A corresponding problem applies to any new preventive treatment, which generates promising surrogate parameters but no additional information on the final success of the prevention program. Our recently published strategy (Porzsolt et al. 2009) proposes a classifcation of innovations. This classification is based on the amount of available qualitative information. The more information is available, which quantifies the “added value,” the higher is the recommendation for the public support of the investigated prevention program.

The strategy was developed to provide a solution for the general problem of healthcare prevention programs. This general problem emerges because our societies consider health care a public good without having commonly accepted indicators to evaluate the quality of this good.

For evaluating the quality of prevention programs the benefits and risks of prevention programs have to be defined. We have summarized a wide spectrum of different preventive treatments and a huge variety of surrogate parameters, which are applied to justify further research. These surrogate parameters are specific for the used treatment and are, therefore, difficult to compare. Two aspects should be considered in a scientifically sound analysis. First, the specificity of the applied treatment is more important than the specificity of the surrogate parameter. Second, the surrogate parameter but not the treatment has to be correlated with the intended final outcome. In other words, surrogate parameters cannot be used to confirm the specificity of
a prevention program but may be used to demonstrate the effectiveness of the program if the correlation of the surrogate with the final end-point has been confirmed.

2.4.3 Recommendations for Future Preventive Research

It should be considered that the public demand for prevention programs is rather strong but the reasons of this strong demand are not well understood. It may be that screening programs for fatal diseases generate perceived safety, which can be quantified (Ursula Rochau, Thesis Medical Faculty University of Ulm 2009, Larissa Gampert, Thesis Medical Faculty University of Ulm 2010, Andreas Knie, Thesis Medical Faculty University of Ulm 2010) and is probably one of the most important aspects of quality of life. So far, most preventive research has been addressing only traditional endpoints, which are rarely achieved, or surrogates of these endpoints, which are not reliable indicators of the quality of a prevention program.

The distinction of structural and functional properties of a disease may be a helpful concept to accept additional markers for description of the quality of prevention programs. It is important in the future to identify functionally inert diseases such as most prostatic cancers, which can be diagnosed due to structural markers but do not need preventive treatment as these forms of cancer cause no harmful functions. This differentiated concept has been discussed in several research areas such as ophthalmology (Brown 2008), nutrition (Paukov 2007), pulmonary disease (Wedzicha and Hurst 2007), and in basic biomolecular research (Malavaki et al. 2008).

The specificity of a prevention program is another problem, which is barely discussed. Although this problem is one of the most complicated research problems and requires ambitious research strategies it should urgently be addressed. It may even be worth the establishment of a special task force. The related problem of “nondisease” is known for almost 50 years (Meador 1965) and has been supported by more recent work on the existence of nondisease genes (Osada et al. 2009). For those who are following these more basic considerations, it is not surprising that the US Prevention Services Task Force (USPSTF) revised their recommendations (regular breast cancer screening only at the age of 50, screening only every other year, neither encouraging nor teaching breast self-examination and no mammography beyond 75). The new findings and the recommendations of a task force should be sufficient to trigger a discussion, which is directly related to the effectiveness and efficiency of today’s medicine.

If preventive medicine wants to keep up with the progress in other areas of medicine it will be mandatory to establish standard rules for basic preventive research. These rules may include the description of the risk profiles of the target population, the investigated interventions, as well as the reported outcomes, which are discussed in this chapter. The resulting structure, which offers the distinction of 32 different values for secondary prevention programs, may serve as proposal for further discussions.

References

USPSTF. http://www.breastcancer.org/symptoms/testing/new_research


Broudy DW, Swint JM, Lairson DR (1979) Prospective economic evaluation of lead poisoning prevention programs. J Community Health 4:291–301


Ghosh PF, AK KRM (2009) Qualitative assessment of innovations in healthcare provision. BMC Health Serv Res 9:50


Clinical Cancer Prevention
Senn, H.-J.; Otto, F. (Eds.)
2011, X, 162 p. 7 illus., 4 illus. in color., Hardcover
ISBN: 978-3-642-10856-3