Screening as a Cancer Control Strategy

Rastreamento como Estratégia de Controle do Câncer
Cribado como Estratégia de Control del Cáncer

Moysés Szklo¹; Liz Maria de Almeida²

INTRODUCTION

Cancer control has been defined by the USA National Cancer Institute as:

The conduct of basic and applied research in the behavioral, social, and population sciences to create or enhance interventions that, independently or in combination with biomedical approaches, reduce cancer risk, incidence, morbidity and mortality, and improve quality of life¹.

In addition to primary prevention, screening (early detection) is an important strategy for cancer control. In this paper, we summarize the major tenets of screening in general and use prostate and colorectal cancer as examples of the application of these tenets.

In the natural history of cancer, early detection is a secondary prevention approach that takes place within the detectable pre-clinical phase (DPCP) (Figure 1), and is based on either removal of precancerous lesions, (e.g., uterine cervix and colorectal), or early detection (cervix, colorectal and breast). The DPCP, which begins when the earliest possible detection is possible and ends when clinical disease is diagnosed based on signs or symptoms, also contains the so-called lead time. Lead time, which is the period that begins with actual early detection and ends with clinical disease, refers to how much early diagnosis can be advanced. Thus, the maximum lead time is the DPCP. Both the lead time and the DPCP can be estimated², and are important variables when the objective is to determine periodicity of screening.

Early detection can be population-based (screening) or opportunistic (case finding) – the latter based on offering screening in the context of an individual medical encounter (“case-by-case” basis). As Rose has aptly demonstrated³, population-based prevention strategies

---

¹ The Johns Hopkins Bloomberg School of Public Health Epidemiology Department. Division of Population Research, National Cancer Institute José Alencar Gomes da Silva (INCA). E-mail: mszklo1@jhu.edu. Orclid ID: https://orcid.org/0000-0001-9433-6366
²Division of Population Research, INCA. E-mail: lalmeida@inca.gov.br. Orclid ID: https://orcid.org/0000-0002-6132-9358

---

Figure 1. Cancer control is based on the natural history of the disease

[Diagram showing the natural history of cancer control]
are more effective than those based on individual level approaches. Thus, screening is often more effective than case finding.

Cancer screening’s key principles are that (1) the test should have a reasonably high validity and be acceptable to the target population, (2) early detection and treatment lead to better outcome than detection based on symptoms, (3) there is an adequately long DPCP, which allows identification of preclinical disease at regular intervals, (4) prevalence is high and, thus, false positivity is minimized, (5) facilities for diagnostic confirmation and treatment should be readily available, (6) screening should be cost-effective vis-à-vis the total expenditures related to health and (7) without treatment, most cases in the pre-clinical phase progress to a clinical phase (a principle that may not be true for certain cancers, e.g., prostate and breast).

It should be emphasized that, although highly sensitive and specific tests are a necessary condition for screening, as they allow detection of the disease in the DPCP, the ultimate utility of a screening program is the extent to which it decreases the risk of the disease outcome.

EVALUATION OF SCREENING

Evaluation of screening is conducted by studies of process and studies of outcome. Studies of process include, for example, the proportion of eligible persons in a given population that undergo screening procedures and proportion of false positives. Evaluation of outcomes relates to effectiveness of the screening process; its main types are the comparison of case-fatality rates (or their complement, cumulative survival) between screened and non-screened patients with the disease of interest, and comparison of mortality in all individuals (not just patients) according to whether they were assigned to the screened or the control group (Figure 2). Because of the possibility of lead time bias (see the next section), the latter is the ideal form of evaluation of screening programs. Other outcomes in screening evaluation include recurrence rate, quality of life and temporal trends in patients found to have early lesions.

BIASES IN SCREENING EVALUATION

The following biases may occur when evaluating the effectiveness of a screening program: selection bias, which includes referral/volunteer bias and length-biased sampling, lead time bias and overdiagnosis bias.

Referral/volunteer bias may occur when the selection of people to receive or not receive the screening procedures is not based on random allocation. Since individuals at a higher risk of a given outcome may be more likely to self-select (e.g., women with a family history of breast cancer), volunteer bias may occur. This bias can be prevented by the conduct of a randomized trial.

Length biased sampling occurs when individuals identified by screening (in a periodic screening program) are compared with those whose diagnosis is made between screening exams (interval cases). Because interval cases usually have a more rapid progression than cases diagnosed by the screening exam, the latter look like they have a better prognosis (Figure 3). Prevention of this bias is based

Figure 2. Two strategies for evaluation of screening effectiveness
on comparing mortality for all individuals allocated to the screening program – regardless of whether they are identified by the screening procedure(s) with the mortality in individuals in the control group.

Lead time bias occurs when survival (or case-fatality) is estimated in patients from the time of early diagnosis. As individuals who undergo screening procedures are likely to be diagnosed earlier, the overall observed survival is influenced by lead time and thus, even if there is no difference in survival between screened and non-screened individuals, a longer survival for screened individuals is observed, as it is counted from the date of early diagnosis (Table 1). Two solutions for preventing this bias are possible: (1) estimation of lead time for the disease under evaluation, which is then subtracted from the survival of the screened group; for example, if the lead time is 2 years and the survival is 8 years from early diagnosis, the actual survival for those who are screened is 6 years. and (2) use of mortality in all screened and non-screened individuals as the main outcome to evaluate effectiveness of screening; because mortality is not calculated from the date of diagnosis and this type of evaluation is not based only on patients, lead time bias is not a consideration and, thus, this type of bias does not occur.

Finally, overdiagnosis bias may result from the inclusion of false positives in the evaluation of screening. As these false positives have a better survival than those who actually have the disease, this bias tends to artifactually increase survival in individuals subjected to screening.

**Translating Knowledge about Screening to a Screening Program**

The process of translating knowledge about screening to a screening program starts with a review (hopefully systematic) of the literature or at least one well designed randomized trial, which leads to evaluation of levels of evidence and programmatic options with or without sensitivity analysis. Based on this evaluation, a cost-effectiveness analysis is carried out, which results in recommendations for the implementation of evidence-based policies. There is usually tension between evidence and obstacles, which can be of an ethical, political or resource-based nature.

**Levels of Evidence**

Decisions about implementation of a screening program (or any other program) should be made based on levels of evidence. Exhibit 1 shows the main levels of evidence. For all levels, it is assumed that the intervention

---

Table 1. Lead time bias: two patients with exactly the same survival from (biological) disease onset

<table>
<thead>
<tr>
<th>Patient</th>
<th>Onset of cancer</th>
<th>Early diagnosis</th>
<th>Clinical diagnosis*</th>
<th>Death</th>
<th>Survival from diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient A</td>
<td>January 2004</td>
<td>2005</td>
<td>January 2015</td>
<td>Patient A survival = 10 years</td>
<td></td>
</tr>
<tr>
<td>Patient B</td>
<td>January 2004</td>
<td>Not screened</td>
<td>January 2015</td>
<td>Patient B survival = 7 years</td>
<td></td>
</tr>
</tbody>
</table>

No gain when adding lead time to the survival of patient A: [Patient A survival - lead time] = Patient B survival = 10 - 3 = 7 years

<table>
<thead>
<tr>
<th>Patient</th>
<th>Onset of cancer</th>
<th>Early diagnosis</th>
<th>Clinical diagnosis*</th>
<th>Death</th>
<th>Survival from diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient A</td>
<td>January 2004</td>
<td>2005</td>
<td>January 2020</td>
<td>Patient A survival = 15 years</td>
<td></td>
</tr>
<tr>
<td>Patient B</td>
<td>January 2004</td>
<td>Not screened</td>
<td>January 2015</td>
<td>Patient B survival = 7 years</td>
<td></td>
</tr>
</tbody>
</table>

Patient A survival is greater than that of Patient B survival because [Patient A survival – lead time] > Patient B survival = 15 - 3 = 12 > 7 years

*Based on symptoms and signs

does more good than harm. The highest level is the result of a systematic review of the literature or a high quality randomized trial. Well-designed observational studies (cohort and case-control) constitute the next level. The following level is the presence of dramatic results in uncontrolled experiments, such as the reduction in cervical cancer mortality after the introduction of Pap testing. Finally, the lowest level is recommendation from experts – not based on systematic evidence - who are convinced that the policy, program or intervention is effective.

These levels are usually discussed by a task force, such as the United States Preventive Services (USPSTF) and the Canada’s Periodic Health Examination task forces, which assigns grades to express their recommendations as to whether the program has a net benefit and, if so, whether it should be implemented (Exhibit 1). More details about this process will be provided as follows in the real-life examples of prostate and colorectal cancers.

THE EXAMPLE OF PROSTATE CANCER

Excluding non-melanoma skin cancer, prostate cancer is the most common type of cancer among men in countries of the American Continent and in parts of Europe, Africa and Oceania. The main risk factor associated with prostate cancer is ageing. Clinical examination and PSA test in combination may suggest the presence of the disease, but the histopathological analysis of the prostate tissue is needed to confirm the diagnosis. In addition, Gleason’s histologic grade complements the information needed to determine the best treatment for the patient. The extent of the disease at time of diagnosis is the main prognostic factor related to 5-year survival rate which, in the USA, varies from 100% for local and regional stages to 29% for the distant stage.

The prostate-specific antigen (PSA) has been traditionally used to diagnose early prostate cancer. There are different types of PSA tests, but for the purposes of this example, we will use a value of PSA ≥ 4 ng/ml to indicate a positive test result. Sensitivity and specificity of a positive PSA test have been given variously as 35-71% and 63-91%, respectively; as a result, percent of false positives is relatively high, having varied from study to study from about 20% to almost 70%. Transient causes of false positivity include prostatitis, urethral endoscopy and some medications (e.g., finasteride). Long term false positive tests results from benign prostatic hyperplasia.

Notwithstanding the relatively high false positive rate, a positive test usually leads to biopsy which, in addition to cost, results in complications such as severe pain in about ¼ of patients, hematuria and hematospermia in approximately ½ and infection in 3-4%. If biopsy indicates presence of cancer, surgery, radiation therapy or active surveillance is recommended. If surgery or radiation therapy is conducted, complications include urinary incontinence, urethral stricture and sexual impotence. The latter is particularly common, occurring in about 10-18% of patients undergoing surgery and 3-8% of those who undergo radiation therapy. Active surveillance is, however, becoming more common and it is recommended to patients at very low or low risk; Exhibit 2 shows the classification of the prostate cancer risk profile suggested by the Johns Hopkins, Department Urology.

Exhibit 1. Translational and Implementation of policy, programs or interventions

<table>
<thead>
<tr>
<th>Levels</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Systematic review or at least one well designed randomized controlled trial has shown that the intervention does more good than harm</td>
</tr>
<tr>
<td>II-1</td>
<td>Well designed cohort or case - control analytic studies (preferably multi-center) suggest that the intervention does more good than harm</td>
</tr>
<tr>
<td>II-2</td>
<td>Dramatic results in uncontrolled experiments (natural experiments) suggest that the intervention does more good than harm</td>
</tr>
<tr>
<td>III</td>
<td>Authoritative and respected experts in the field are convinced of the value or lack of value of the intervention</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Translation?</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>There is high certainty that the net benefit is substantial</td>
<td>Design and offer/provide this intervention (or program/policy)</td>
</tr>
<tr>
<td>B</td>
<td>There is high certainty that the net benefit is moderate or moderate-to-substantial</td>
<td>Design and offer/provide the intervention only if other considerations support offering or providing the intervention on a case by case basis. Case-finding is recommended</td>
</tr>
<tr>
<td>C</td>
<td>There is moderate or high certainty that the net benefit is small</td>
<td>Discourage the use of this intervention</td>
</tr>
<tr>
<td>D</td>
<td>There is moderate or high certainty that there is no net benefit or that the harms outweigh the benefits</td>
<td>If the intervention is offered, individuals should understand the uncertainty about the balance of benefits and harms. Case-finding is recommended</td>
</tr>
<tr>
<td>I</td>
<td>The current evidence is lacking, of poor quality or conflicting</td>
<td></td>
</tr>
</tbody>
</table>


Exhibit 2. Classification of the prostate cancer risk profile suggested by the Johns Hopkins, Department Urology.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Low risk</td>
</tr>
<tr>
<td>II</td>
<td>Intermediate risk</td>
</tr>
<tr>
<td>III</td>
<td>High risk</td>
</tr>
</tbody>
</table>

Exhibit 3. The classification and management of the prostate cancer risk profile suggested by the Johns Hopkins, Department Urology.
Effectiveness of prostate cancer screening

Two of the best randomized trials on the effect of screening on mortality from prostate cancer were conducted in the USA and Europe, respectively9,10. In the USA trial, cumulative prostate cancer mortality after about 9 years was greater in the screening than in the control group. After the same follow-up period, no difference in prostate cancer mortality in the European trial was found between the groups. Reflecting these trends, in 2012 the US Preventive Services Task Force assigned a grade of evidence (see Exhibit 1), reflecting a moderate/high certainty that no net benefit could be expected from screening and that, therefore, implementation of PSA testing should be discouraged. However, further follow-up of the European trial showed a significantly lower mortality in the PSA than in the control group11. This positive result prompted the American Cancer Society (ACS)7 to assign a grade C of evidence in 2016, which poses that the level of certainty is moderate or high that the expected benefit is small; consequently, its recommendation was for a “case-by-case” approach to PSA testing; in other words, suggesting that implementation should be based on “case finding”. The “case by case” approach was specified by the ACS for different age groups (Exhibit 3). It is useful to reproduce literally the recommendation from the ACS:

Exhibit 2. Active surveillance: criteria and recommendations from the Johns Hopkins, Department of Urology, based on prostate cancer patient’s risk profile

<table>
<thead>
<tr>
<th>RISK PROFILE</th>
<th>CRITERIA</th>
<th>RECOMMENDATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low</td>
<td>PSA &lt;10</td>
<td>Gleason score &lt;7* and stage T1c** and PSA density &lt;0.15*** and unilateral disease regardless of percent core involvement</td>
</tr>
<tr>
<td></td>
<td>PSA 10-20</td>
<td>Gleason score &lt;7 and stage T1c and PSA density &lt;0.10 and unilateral disease with &lt;3 cores containing cancer regardless of percent core involvement</td>
</tr>
<tr>
<td>Low</td>
<td>Stage T1c or T2a**** and Gleason score &lt;7 and PSA density &lt;10</td>
<td>Age &gt;65 Preferred if life expectancy &lt;10 yrs</td>
</tr>
<tr>
<td>Intermediate</td>
<td>T2 or PSA 10-20 or Gleason score 3+4</td>
<td>Life expectancy &lt;10 years</td>
</tr>
<tr>
<td>High</td>
<td>Stage T3 or Gleason score 3+4 or PSA&gt;20</td>
<td>Not recommended</td>
</tr>
</tbody>
</table>


*The cells are well differentiated and look like healthy cells
**Tumor found during needle biopsy, usually because of elevated PSA
****PSA number ÷ prostate volume
*****Tumor involves ½ of 1 side of the prostate

Although in the updated European study the prostate cancer mortality was significantly lower in the PSA than in the control group, a C grade of evidence, rather than A or B (Exhibit 1), was selected because the authors could not find a difference in overall (all-cause) mortality between the groups.

In addition to the trials summarized in the previous paragraph, other clinical trials have been conducted to examine the effectiveness of a positive PSA test. In these trials, extensively and systematically reviewed by Fenton et al12, with the exception of one trial, prostate cancer mortality was lower in the screened than in the non-screened group, with effectiveness varying widely from 4.0-42.0%. It should, however, be noted that – as for the European trial – the relative risk for all-cause mortality was close to 1.0 in all trials; that is, no benefit was seen for overall mortality. Based on Fenton’s systematic review, the USPSTF recommended a grade C for men aged 55-69 years, and a grade of D for those ages 70 and older. It can be
hypothesized that, as for the ACS, the USPSTF decided to assign a grade C (rather than A or B) for those aged 55-69 years because overall mortality was not decreased with PSA screening in all randomized trials conducted heretofore.

The conundrum of defining false positivity in prostate cancer

There is a consensus that a relatively high proportion of patients with prostate cancer do not die from prostate cancer. In the USA-based Surveillance, Epidemiology, and End Results (SEER) Program, for example, of about 221,000 incident cases occurring in the United States every year from 1975 through 2011, only approximately 27,500 yearly deaths with prostate cancer as underlying cause have been observed. This corresponds to an annual case-fatality of around 12.5%. Thus, prostate cancer is very likely not invasive in a large proportion of patients, which means that, using lethal cancers as true cases, an expanded definition of false positives would include not only those with a positive PSA without the disease, but also those with the disease that does not become invasive. There are current efforts to identify biomarkers that would allow prediction of invasiveness of prostate cancer. In the meantime, as mentioned previously, active surveillance has been recommended for individuals at low and very low risk (Exhibit 2).

Screening and primary prevention are both important: the example of colorectal cancer

Colorectal cancer is the third most incident and lethal type of cancer, with 1,849,518 new cases and 880,792 deaths worldwide. The most recent USPSTF guidelines for colon cancer screening are from June 2016, which recommends that screening for colorectal cancer should start at age 50 years and continue through age 75 years. The Task Force suggests a combination of 3 tests: fecal occult blood test (FOBT) or fecal immunological test (FIT) every 3 years, flexible sigmoidoscopy every 5 years and colonoscopy every 10 years. This recommendation is based on strong evidence (Grade A) of effectiveness, however, and particularly for screening in developing countries, less invasive options should be found for colorectal cancer, as colonoscopy or even sigmoidoscopy may not be acceptable to most people and is an expensive procedure that requires well-trained professionals.

For adults aged 76-85 years, the recommendation is based on a “case-by-case” basis and considers the individual’s overall health and prior screening history (evidence Grade C, denoting moderate or high level of certainty).

As effective as colorectal screening is, particularly if novel, more acceptable, yet highly sensitive and specific strategies are found, primary prevention cannot be neglected. As estimated by Platz et al, if everyone in the population had optimal levels of factors associated with colorectal cancer, 71% of colorectal cancer would be preventable. These optimal levels include body mass index (kg/m²) <25, ≥75 minutes/week of vigorous exercise or ≥150 minutes/week of moderate plus vigorous exercise, not smoking, alcohol <15 g/day, red meat intake <2 servings/week and >100 µg consumption of folic acid supplement/week.

Conclusion

Although primary prevention, whenever possible, is the best strategy, screening is also an important approach for cancer control. Assessing the effectiveness of cancer screening programs as well as the validity of new tools for the early diagnosis of specific cancer types are important for health managers’ decision making. Thus, guidelines must be reviewed periodically.

The examples of prostate and colorectal cancers show that the decisions to plan and implement population-wide cancer screening are not trivial, and must be carried out taking into consideration the evidence resulting from well-designed studies. In addition, a careful assessment of risks and benefits involved in diagnostic and therapeutic procedures should be conducted.

References


