COURSE: Evidence-Based Approaches to HPV Screening implementation

Module 2. Principles of screening

Authors

Laura Costas, MD, PhD

Medical Specialist in Preventive Medicine and Public Health, Cancer Epidemiology Research Programme, Catalan Institute of Oncology, Spain.

Paula Peremiquel-Trillas, MD, MPH

Medical Specialist in Preventive Medicine and Public Health, Cancer Epidemiology Research Programme, Catalan Institute of Oncology, Spain.

Reviewer

Eduardo L. Franco, Prof

Former Chairman (2011-23), Department of Oncology, and Former Director (1995-2024) of the Division of Cancer Epidemiology, McGill University, Montreal, Canada

This module is part of the training course: ***Evidence-Based Approaches to HPV Screening implementation,*** developed within the RISCC project, funded by *European Union's Horizon 2020 research and innovation programme under grant agreement No 847845*.

CONTENTS

INTRODUCTION AND LEARNING OBJECTIVES

In 2022, almost 20 million newly diagnosed cancer cases occurred worldwide, and more than 9.7 million patients died from cancer. Cancer is a major public health problem worldwide and is a leading cause of disease and death in many countries, yet a significant proportion of cases can be prevented, particularly those benefiting from early detection, precancerous detection or primary prevention actions.

As seen in the previous module, cervical cancer is a paradigm of cancer that is amenable to prevention through screening. Because of its natural history, we have a window of opportunity to detect and treat precancerous lesions before they can progress to malignant lesions. The screening process is complex, however, and its impact relies on many components. Definition of the target population, the performance of screening tests, and the management of subjects who test positive are key aspects that may define whether a screening programme succeeds or fails.

This module reviews the general aspects of cancer screening, including the criteria for a suitable screening test and the overall requirements for implementation of a screening programme. We will review the basic concepts of a traditional two-by-two table, which displays the distribution of positive or negative test results according to their presence or absence of disease. This simple table will allow us to extract basic screening concepts such as sensitivity or predictive value. However, its simplicity is deceiving, given that the same test may perform differently in different settings or populations, e.g., primary screening or a referral clinic and women living with or without HIV.

Although this module provides concepts largely applicable to other diseases, the module focuses on cervical cancer and its singularities.

At the conclusion of this course, participants will be able to:

- Understand the principles of cancer screening.
- Identify the types of screening programmes.
- Understand screening test characteristics.
- Know the overall requirements for screening programme implementation

UNIT 1. UNDERSTANDING CERVICAL CANCER PREVENTION

1.1 Natural history of the cervical cancer and types of cancer prevention

The natural history of a disease describes the factors that contribute to the switch from a healthy status to a disease status and its steps from preclinical to clinical to invasion. Understanding how a disease evolves is fundamental to learning how to prevent it from progressing.

Many diseases have certain, relatively well-defined, well-known stages in which different types of preventive measures can be applied. For instance, **Figure 1** shows the different well-known stages of the natural history of cervical cancer.

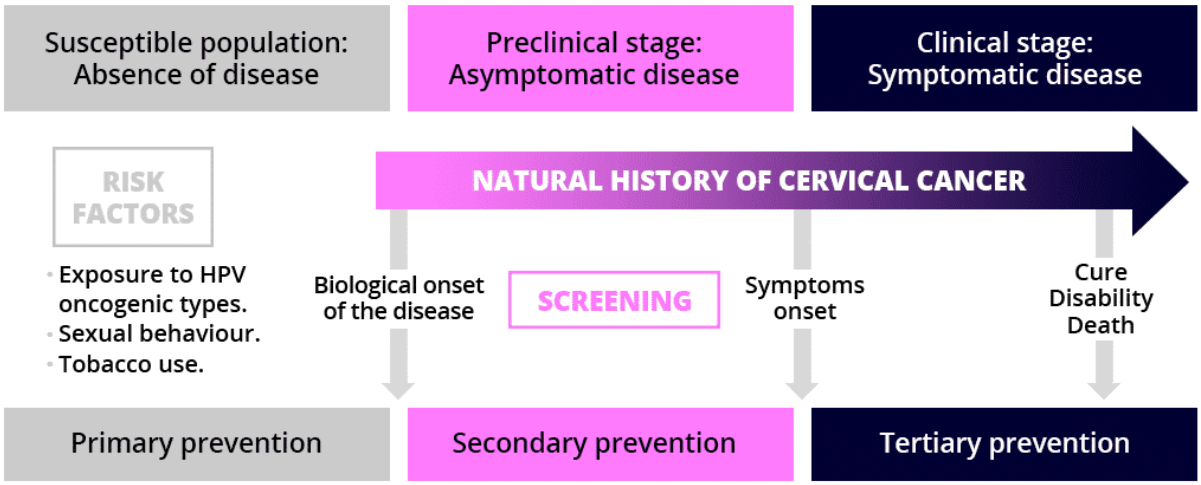


Figure 1. Natural history of cervical cancer and types of preventive measures

Primary prevention

The disease process begins with exposure to aetiological factors in a susceptible host (Principles of Epidemiology | Lesson 1 - Section 9, 2020).

At this stage, primary prevention focusses on protecting healthy individuals from the biological onset of disease via health promotion and health protection interventions (Eeles et al., 2018). Primary prevention strategies therefore aim to reduce the incidence of disease by eliminating exposure to risk factors or by increasing population resistance to these.

The main primary prevention action in cervical cancer is avoiding exposure to oncogenic HPV types – the necessary cause of cervical cancer - through vaccination.

Secondary prevention

In some individuals, the disease process is triggered and pathological changes begin to take place, normally without the individuals being aware (asymptomatic). This is known as the subclinical stage of the disease, or the latency period.

At this stage, secondary prevention or screening involves the systematic use of safe, easy-to-use and affordable tests to enable early diagnosis of disease followed by timely treatment (Eeles et al., 2018). Secondary prevention aims to identify health problems before they worsen, generally in asymptomatic populations, to provide treatment at an early stage and to improve disease prognosis. Therefore, screening aims to reduce disease mortality by decreasing its prevalence, shortening its duration, reducing the incidence of complications associated with the disease and increasing the quality of life of the patient.

In cervical cancer, it can take decades for an HPV infection to progress to cancer. Secondary prevention in this case can use cytology, HPV testing or a combination of both, to detect precancerous lesions caused by persistent HPV infection.

Tertiary prevention

At some point, the individual becomes symptomatic and seeks medical assistance. Ultimately, the disease process ends in recovery, disability or death.

Tertiary prevention aims to manage and decrease the impact of a long-term disease (avoiding complications and relapses) to reduce morbidity, disability and mortality among those diagnosed and undergoing treatment.

The presentation and course of cancer will vary in different individuals and contexts, even for the same disease. HPV infections may never progress to cervical cancer in most women while the process may result in severe or fatal illness for a minority. This diversity is called the spectrum of disease.

CERVICAL CANCER PREVENTION MEASURES

These are the different strategies or interventions being used to prevent exposure to HPV and progression to cervical cancer, and mitigation measures once cancer has been diagnosed.

Primary prevention	Vaccination against oncogenic HPV types Use of latex barriers such as condoms Health information and education about tobacco use
Secondary prevention	Detection of oncogenic HPV types (HPV testing) with or without detailed information on genotype Detection of cellular morphological lesions (cytology) Detection of biomarkers associated with HPV-driven carcinogenesis
Tertiary prevention	Tailored management Stage-specific and biomarker-based precision therapies Prognosis-based surveillance

In addition to the specific strategies to prevent cervical cancer, there are other more general strategies, such as the European Code Against Cancer (Schüz et al., 2015). It comprises a list of 12 recommendations developed by a group of experts after assessment of the available data on actions that individual European citizens can take to help prevent cancer:

1. Do not smoke. Do not use any form of tobacco.
2. Make your home smoke free. Support smoke-free policies in your workplace.
3. Take action to be a healthy body weight.
4. Be physically active in everyday life. Limit the time you spend sitting.
5. Have a healthy diet:
   - Eat plenty of whole grains, pulses, vegetables and fruits.
   - Limit high-calorie foods (foods high in sugar or fat) and avoid sugary drinks.
   - Avoid processed meat; limit red meat and foods high in salt.
6. If you drink alcohol of any type, limit your intake. Not drinking alcohol is better for cancer prevention.

7. Avoid too much sun, especially for children. Use sun protection. Do not use sunbeds.
8. In the workplace, protect yourself against cancer-causing substances by following health and safety instructions.
9. Find out if you are exposed to radiation from naturally high radon levels in your home. Take action to reduce high radon levels.
10. For women:
    - Breastfeeding reduces cancer risk. If you can, breastfeed your baby.
    - Hormone replacement therapy (HRT) increases the risk of certain cancers. Limit use of HRT.
11. Ensure your children take part in vaccination programmes for:
    - Hepatitis B (for newborns)
    - Human papillomavirus (HPV) (for girls) – *NOTE*: *The European Code Against Cancer was published in 2015. Since then, several vaccination programmes offer also HPV vaccination to boys. A revised version of the code will be available in 2025.*
12. Take part in organized cancer screening programmes for:
    - Bowel cancer (men and women)
    - Breast cancer (women)
    - Cervical cancer (women)

However, for cancer prevention to be successful, these individual actions must be supported by government policies and actions.

Find out more about the European Code Against Cancer at: http://cancer-code-europe.iarc.fr

1.2. Principles of screening

Screening is the action of actively searching for a disease in an asymptomatic population. However, not all diseases are suitable for screening and certain criteria must be met.

In 1968, Wilson and Jungner set out ten principles to be considered before undertaking a screening programme. These criteria are key in understanding the complexity and requirements of a screening programme (Andermann, 2008; Wilson & Jungner, 1968).

	Wilson and Jungner principles	Criteria applied to cervical cancer
1.	The condition sought should be an important health problem	Cervical cancer is the fourth most common cancer among women worldwide. In 2020, 605,000 women were diagnosed with cervical cancer and more than 340,000 died from this disease
2.	There should be an accepted treatment for patients with recognised disease	Precancerous lesions at an advanced stage (CIN2+) can be treated through LEEP or surgical procedures. In resource-limited settings, ablative methods can be used
3.	Facilities for diagnosis and treatment should be available	Follow-up procedures and potential treatment after a positive result need to be agreed. In resource- constrained regions with difficulties in follow-up procedures, screen-and-treat approaches can be useful
4.	There should be a recognisable latent or early symptomatic stage	It can take decades for an HPV infection to progress to cancer. During this latency period, both HPV infections and precancerous lesions can be detected
5.	There should be a suitable test or examination	Cervical cytology and oncogenic HPV detection are effective methods for detecting preclinical phases
6.	The test should be acceptable to the population	Cervical sampling by a clinician or by women themselves is generally well accepted by participating women
7.	The natural history of the condition, including development from latent to declared disease, should be adequately understood	The spectrum of disease from HPV infection to cervical cancer is well understood
8.	There should be an agreed policy on whom to treat as patients	In the case of abnormal tests, specific protocols for referral and treatment exist
9.	The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole	Cervical cancer screening programmes are generally cost-effective
10.	Case-finding should be a continuing process and not a "once and for all" project	Screening programmes require a systematic approach with specific multidisciplinary guidelines and quality control procedures

In addition to these core criteria, additional criteria relative to the programme itself are also critical:

- The screening programme should respond to a recognised need.
- The objectives of screening should be defined at the outset.
- There should be a defined target population.
- There should be scientific evidence of screening programme effectiveness.
- The programme should integrate education, testing, clinical services and programme management.
- There should be quality assurance, with mechanisms to minimise potential risks of screening.
- The programme should ensure informed choice, confidentiality and respect for autonomy.
- The programme should promote equity and access to screening for the entire target population.
- Programme evaluation should be planned from the outset.
- The overall benefits of screening should outweigh any harm.

Screening is a process, not just a test. Follow-up procedures and potential treatment for screen positives are ethical imperatives.

UNIT 2. TEST PARAMETERS AND REQUIREMENTS

2.1. Criteria for a good screening test

Screening tests are used on subjects without clinical symptoms of disease. Potential tests therefore need to be evaluated carefully to limit potential negative impacts on health outcomes (delay at diagnosis, false positives, etc.) and health systems expenditure.

Thus, a screening test needs to be follows (Gray, 2004):

- **Accurate:** the test detects and classifies correctly patients who are ill and those who are not.
- **Reliable/reproducible:** the test provides the same result consistently when repeated and when performed in different settings.
- **Affordable:** the test is affordable for the health system and is useful in reducing the costs associated with the disease (monetary and non-monetary).
- **Accessible:** Subjects at risk have access to the screening test and subsequent procedures. Also, the steps are clear and easy to follow.
- **Acceptable:** Participating subjects and providers tolerate the screening test well. A screening test must be acceptable to the population, easy to use and cause minimal discomfort.
- **Simple:** the test is easy to use, and management of results is easy to follow This is especially relevant for implementation in low-resources settings
- **Safe:** The screening procedure is safe and subjects who test positive have minimal adverse effects. The potential harmful effects of screening need to be considered, quantified and evaluated. Potential harm in cervical cancer screening includes physical harm (such as pain or bleeding due to the screening or diagnostic tests), psychological harm (such as anxiety caused by a positive result) or harm related to false positives (such as unnecessary harm and tests) and false-negatives (undiagnosed disease) (McGee, 2002).

2.2 Test accuracy and reliability

When a test is performed, an individual can be classified as positive or negative. These results are compared with the results of a reference test or gold standard – the most accurate test available – to confirm disease and obtain the test accuracy readings.

In cervical cancer prevention….

a positive screening result is generally followed by a confirmatory test. If both tests are positive, these are generally followed by histological confirmation through biopsy of the cervix to determine which treatment is best.

Test accuracy measures provide data on the ability of the test to:

- Discriminate health and disease status, i.e. distinguishing people who are ill and those who are not.
- Predict disease by estimating the probability of having or developing disease following a positive or negative result.

The principal measures of test accuracy are:

- Sensitivity and specificity
- Positive and negative predictive values (PPV/NPV)
- Likelihood ratios (LR)
- The area under the curve (AUC) in the receiver operating characteristic (ROC) curve
- Overall diagnostic accuracy

Tested patients are classified into four groups based on test results and disease status as displayed in **Table 1**.

Table 1. Test results according to disease status

		Gold standard / Reference test	
		Subjects with the disease	***Subjects without the disease***
Test	***Positive***	True positive (TP): ill subjects detected by the test	False positive (FP): healthy subjects whose tests are mistakenly positive
	Negative	False negative (FN): ill subjects **not** detected by the test	True negative (TN): healthy subjects correctly classified by the test

2.2.1 Sensitivity and specificity

Sensitivity and specificity are the measures used to discriminate health and disease status (Altman & Bland, 1994b).

- Sensitivity measures the percentage of diseased subjects that test positive.

Sensitivity = TP / (TP+FN)

- Specificity measures the percentage of disease-free subjects who test negative.

Specificity = TN / (FP+TN)

A perfect test would classify all subjects as either TP or TN (i.e., 100% sensitivity and 100% specificity). However, tests rarely classify all subjects correctly according to their disease status since there is an overlap in the distribution of patients with and without disease.

To adequately compare the accuracy of two screening tests, we need to understand, among other things, how positivity is defined.

When comparing HPV testing and cervical cytology, for example, it is important to consider that HPV tests can differ among them on the number of HPV types targeted for detection or the method used to detect them. On the other hand, a cytology is usually considered abnormal with an ASC-US result or worse but a higher threshold of LSIL or HSIL can also be used.

ACTIVITY

Use the information in the following tables to calculate sensitivity and specificity for high-risk HPV testing and conventional cytology in detecting a histologically proven high-grade lesion (Clavel et al., 2001). Which test has the highest sensitivity and which the highest specificity?

	Histologically proven high-grade lesion		
	Yes	*No*	*Total*
Positive HPV test	47	284	331
Negative HPV test	0	1950	1950
Total	47	2234	2281

	Histologically proven high-grade lesion		
	Yes	*No*	*Total*
Positive cytology	32	104	136
Negative cytology	15	2130	2145
Total	47	2234	2281

Answer:

The results for sensitivity and specificity are as follows:

- HPV test sensitivity = 47 / 47 = 1
- HPV test specificity = 1950 / 2234 = 0.87

- Cytology sensitivity = 32 / 47 = 0.68
- Cytology specificity = 2130 / 2234 = 0.95

Therefore, HPV test has the highest sensitivity whereas cytology has the highest specificity

The calculations are expressed as proportions but they are usually converted to percentages as in the text below. Both can be used when talking about measures of accuracy.

The HPV test is superior because it identifies all 47 high-grade lesions as positive while cytology misses 15 proven cases (100% vs. 68% sensitivity). However, the HPV test classifies 284 women with no evidence of a high-grade lesion as positive, compared to 104 women in the case of cytology, and therefore cytology has higher specificity (95% vs. 87%).

2.2.2 Predictive values

Clinicians are generally interested in the detection or absence of disease based on the results of the test being applied. Predictive values denote the probability that the test gives the correct diagnosis (Altman & Bland, 1994a).

Using the two-by-two table above (**Table 1**), we can compute predictive values as follows:

- PPV measures the probability of presence of disease when the test is positive (proportion of subjects with a positive test result who are correctly diagnosed). Low PPV values indicate a substantial number of unnecessary subsequent tests, which is an important consideration relative to invasive tests in false positive subjects.

$$\mathbf{PPV = TP / (TP+FP)}$$

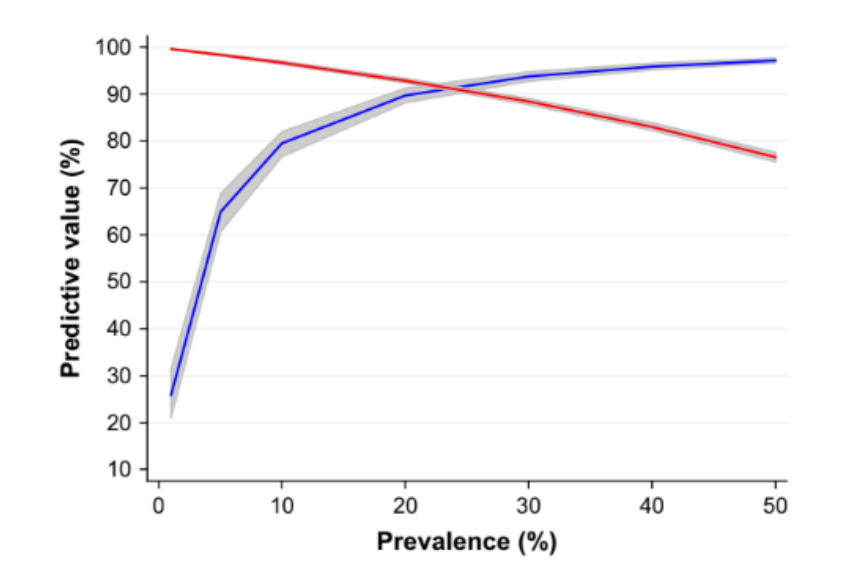
- NPV measures the probability of absence of disease when the test is negative, i.e. the proportion of subjects with a negative test result who are not ill.

$$\mathbf{NPV = TN / (TN+FN)}$$

Predictive values are affected by disease prevalence. As observed in the next figure, the PPV of a test increases with increasing prevalence, while the NPV decreases with increasing prevalence. Both increase with severity of disease (Brenner & Gefeller, 1997). Therefore, the PPV and NPV from one study should not be transferred to other settings with a different prevalence of disease.

EXAMPLE

The following figure reflects the changes in the predictive value for cytology performance in a hypothetical population of 10.000 women assuming a constant 70% sensitivity and 98% specificity. This simulated data shows how the PPV (blue line) of cytology increases as the prevalence of cervical lesions increases while the NPV (red line) decreases, in a simulated population of 10,000 women (Franco et al., 2009).



In cervical cancer prevention….

cytology has a higher PPV (a higher proportion of positive subjects actually have disease) while HPV tests have a higher NPV (we will miss fewer cases among those that tested negative). A high NPV provides a good assurance of a negative result over several years, which allows for an extended screening interval between screening tests and can convey cost-effectiveness to a screening program.

Sensitivity and specificity have been traditionally regarded as constant benchmarks of test performance and are frequently used to compare the diagnostic values of different tests and can be evaluated across multiple studies through meta-analyses or polling results. That said, sensitivity and specificity are also affected by disease prevalence (Brenner & Gefeller, 1997).

When disease prevalence is very low, the PPV will be very low even if sensitivity and specificity are high. This means that many subjects with a positive result who do not have the disease (false positives) will undergo subsequent testing unnecessarily.

In cervical cancer prevention….

the lower prevalence of HPV infections and disease due to an increased number of vaccinated women will affect the performance of HPV testing as primary screening test

Wilson and Jungner's first principle requires that the disease to be screened should be an important health problem. Screening for a disease with a very low disease prevalence would be very inefficient even with a good screening test.

NOTE: *For more information on the relationship between prevalence and HPV testing, please see* ***MODULE 3*** *and* ***MODULE 6****).*

ACTIVITY

Use the information in the following tables (Clavel et al., 2001) to calculate the PPV and the NPV for high-risk HPV testing and conventional cytology.

	Histologically proven high-grade lesion		
	Yes	*No*	*Total*
Positive HPV test	47	284	331
Negative HPV test	0	1950	1950
Total	47	2234	2281

	Histologically proven high-grade lesion		
	Yes	*No*	*Total*
Positive cytology	32	104	136
Negative cytology	15	2130	2145
Total	47	2234	2281

Answer:

To obtain the PPV of the HPV test we divide the number of positive results with disease (47) by the total number of positive results (331), i.e. 47 / 331 = 0.14. Similarly, NPV is calculated by dividing 1950 / 1950 = 1.

The values for cytology are as follows: PPV = 32 / 136 = 0.23 and NPV = 2130 / 2145 = 0.99.

2.2.3 Likelihood ratios

We can use sensitivity and specificity to obtain the likelihood ratios, which indicates how many times more (or less) likely it is that subjects with the disease have a particular result than subjects without the disease (Deeks & Altman, 2004).

- A positive likelihood ratio (LR+) indicates how many times more likely is it that a positive test occurs in subjects with the disease than in those without the disease. Values range from zero to infinity. The higher the value, the more likely the patient has the condition.

  As an example, let's say a positive test result has an LR of 9.2. This result is 9.2 times more likely to happen in a subject with the condition than it would in a subject without the condition.

  LR+ = Sensitivity / (1-Specificity)

- A negative likelihood ratio (LR-) indicates how much less likely it is that a negative test result occurs in a patient with the disease than in a healthy subject. When LR is below 0.1, it is considered to provide enough evidence to exclude diagnosis under most circumstances. The lower the LR- is, the stronger the evidence of absence of disease.

  LR- = (1-Sensitivity) / Specificity

Since sensitivity and specificity are used to calculate the LR, both LR+ and LR- once again depend on disease prevalence.

RULE OF THUMB

A rule of thumb for interpreting likelihood ratios (McGee, 2002)::

- *0 to 1: Decreased evidence for disease. Values closer to zero have a greater decrease in probability of disease. For example, an LR of 0.1 decreases probability by -45%, while a value of -0.5 decreases probability by -15%.*
- *1: No diagnostic value.*
- *Above 1: Increased evidence of disease. The farther away from 1, the more chance of disease. For example, an LR of 2 increases the probability by 15%, while an LR of 10 increases the probability by 45%. An LR over 10 is very strong evidence of disease.*

ACTIVITY

Use the results obtained in previous exercises and the information in the following table (Clavel et al., 2001) to calculate the LR+ and LR- for high-risk HPV testing.

How many times more likely a positive result will be observed in a subject with a high-grade lesion than in a subject without the condition?

	Histologically proven high-grade lesion		
	Yes	*No*	*Total*
Positive HPV test	47	284	331
Negative HPV test	0	1950	1950
Total	47	2234	2281

A. 7.69 times more likely to happen in a patient with a condition than in a subject without the condition.

B. 11.46 times more likely to happen in a patient with a condition than in a subject without the condition.

Answer:

The correct answer is A.

- Sensitivity = 47 / 47 = 1
- Specificity = 1950 / 2234 = 0.87
- HPV test LR+ = 1 / (1-0.87) = 7.69
- HPV test LR- = (1-1) / 0.87 = 0

The LR+ of 7.69 indicates that this result is 7.69 times more likely to happen in a patient with a condition than in a patient without the condition.

2.2.4 Defining the best cut-off value

Many diagnostic tests are quantitative and a cut-off point has to be established to distinguish health and disease status. Defining a positive and a negative result in a test requires establishing a criterion for what is a positive result.

In cervical cancer prevention….

HPV positivity is established at a defined signal intensity of the HPV assay that is associated with the presence of an underlying lesion.

Cytology abnormality is established at a defined severity of cellular abnormalities.

Ideally, because a screening test should have high sensitivity and specificity, the cut-off value should clearly distinguish those with the disease (positive test result) from those without the disease (negative test result) (**Figure 2A)**. However, this ideal scenario is not attainable, as there is usually some overlap between healthy subjects with high values and subjects with the disease and low values (**Figure 2B**).

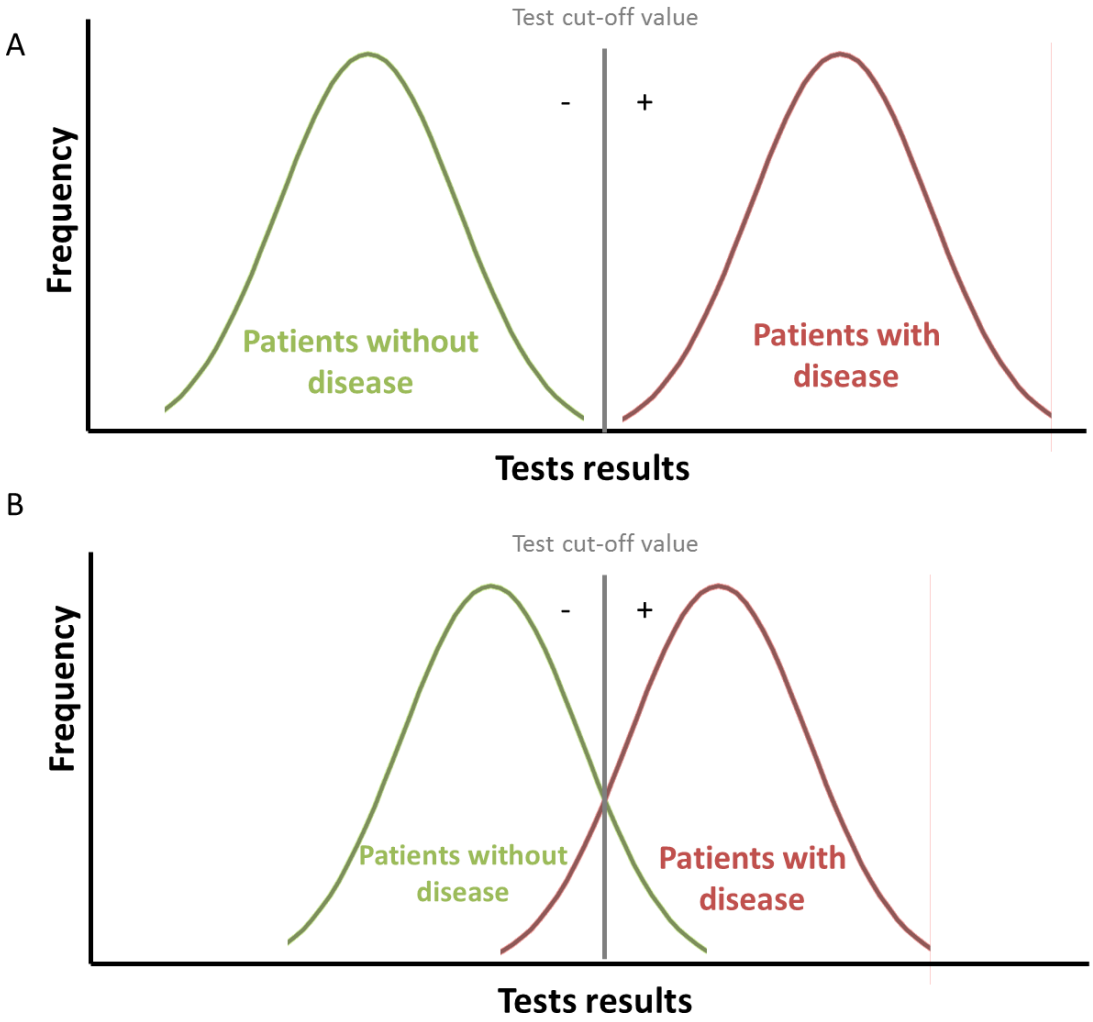


Figure 2. Cut-off values to define positive and negative in quantitative tests.

Usually when sensitivity increases specificity decreases, and vice versa. In **Figure 3**, if we choose cut-off A (100% sensitivity), we will detect all patients with a disease, but half of subjects without disease will be referred for unnecessary tests. The opposite occurs at cut-off B; by avoiding sending subjects without disease for unnecessary tests (100% specificity), we would miss half of patients with disease.

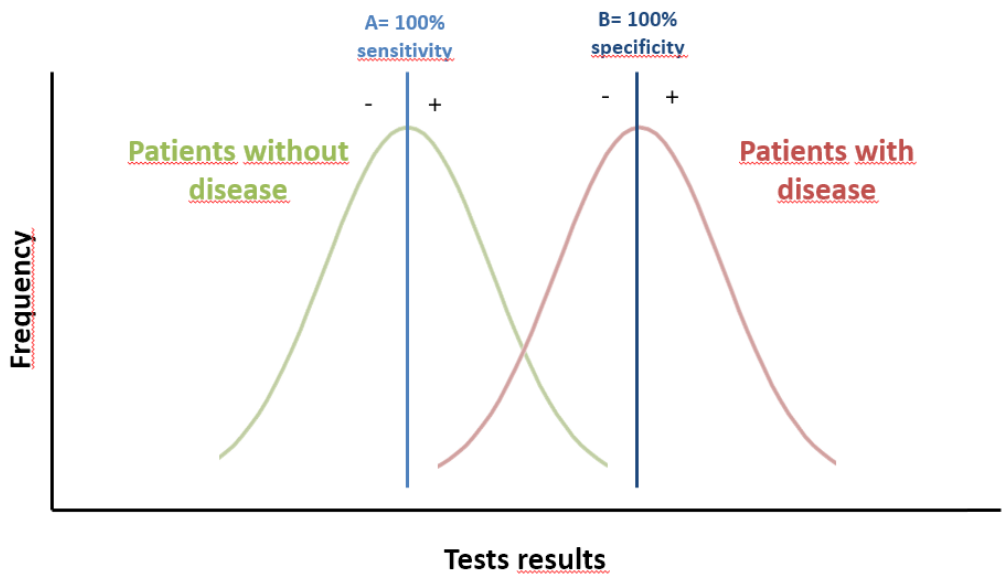


Figure 3. Distribution of disease and non-disease with diagnostic cut-off values and subsequent effects on sensitivity and specificity

In cervical cancer prevention….

HPV testing aims to detect persistent HPV infections, since these are the aetiological factor for cervical precancer and cancer. However, because we still cannot differentiate between acute and persistent infections, testing may sometimes identify an HPV infection that will clear spontaneously and not cause disease (false positive).

Receiver operating characteristic (ROC) curve

ROC curves are particularly useful for determining the cut-off points for a test that provides quantitative results as continuous data, such as the number of viral copies in an HPV test.

A ROC curve represents all pairs of sensitivity and specificity values for all cut-offs that could be established for the test in question, in which the x-axis represents 1-specificity (the false positive rate) while the y-axis represents sensitivity (true positive rate) (**Figure 4**).

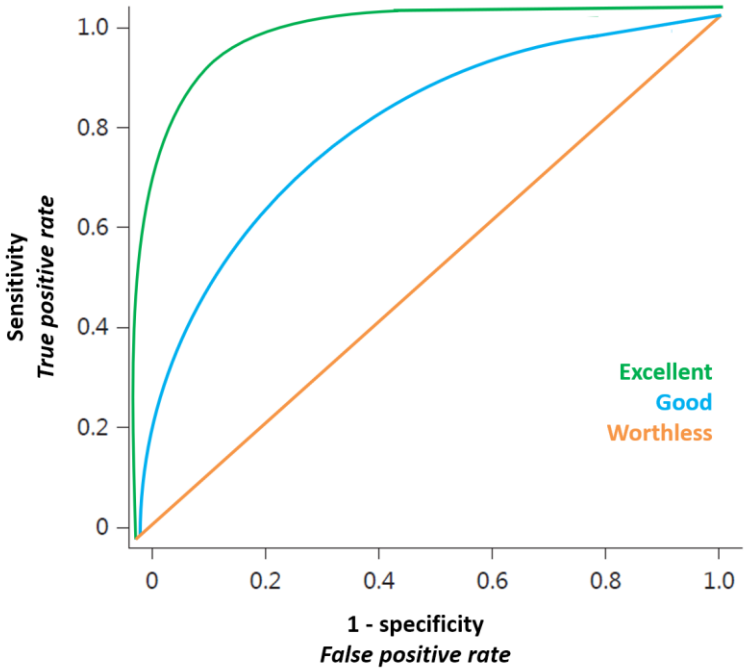


Figure 4. Example of ROC curves and AUC

The shape of the ROC curve and the area under the curve (AUC) enable an estimation of how high the discriminative power of a test is: the closer the curve is to the upper-left corner (highest sensitivity and specificity), the larger the AUC is and, therefore, the better the test is at discriminating health and disease:

RULE OF THUMB

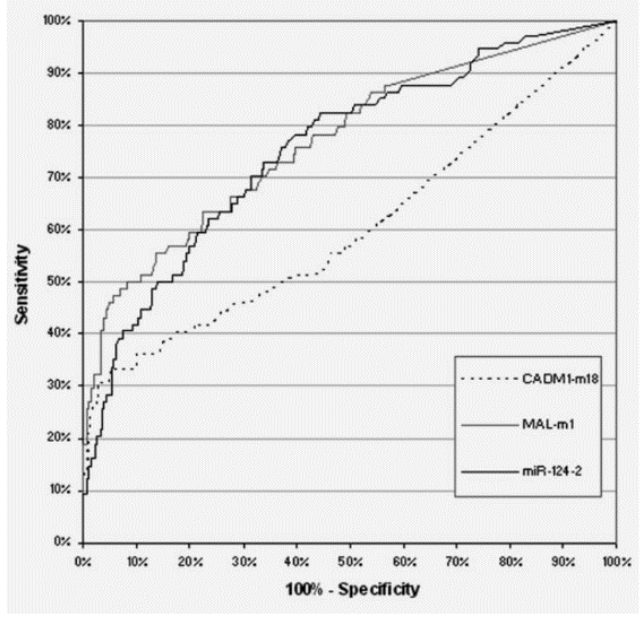
An AUC of 1 represents the perfect test, with 100% sensitivity and 100% specificity. An AUC of 0.5 represents a worthless test that does not allow discrimination, a test that irrespective of cut-off sensitivity plus specificity equals 1 (sensitivity 10% and specificity 90%, sensitivity 70% and specificity 30%, etc).

***NOTE:** A ROC curve should not be built using a unique cut-off value or threshold. The ROC curve is a graphical representation to assess a test's performance across different thresholds. A single unique cut-off value or threshold does not capture this trade-off and may not provide a comprehensive view of the test's performance.*

ROC curves are also useful to assess a test, such as cytology, whose abnormalities are graded in terms of severity, such as ASC-US, LSIL and HSIL.

EXAMPLE

An an example. the following figure shows the ROC curve for CIN3+ of different methylation markers according to different values of Ct ratios used to define each marker positivity (Hesselink et al., 2014).



As expected, at increasing Ct ratios the specificity improves but the sensitivity decreases.

For a first-line screening test, high sensitivity is preferred over high specificity to minimise the number of false negatives.

To reduce unnecessary referral for final diagnosis and/or treatment (false positives) in populations screened with a primary test with limited specificity, those that test positive are tested again with a more specific test. This step is called TRIAGE and may be embedded within the primary test itself. In this scenario, the primary test does not provide only a positive vs negative result but allows to further stratify those with a positive result. For example, a positive HPV test could provide genotyping data that would allow to stratify the risk.

2.2.5 Overall diagnostic accuracy

Lastly, the overall diagnostic accuracy of a test is a global measure of the proportion of subjects correctly identified among the total number of subjects:

Overall diagnostic accuracy = (TP + TN) / total subjects

This measure is also affected by disease prevalence; i.e with the same sensitivity and specificity, the overall diagnostic accuracy increases as disease prevalence decreases.

Of note, the overall diagnostic accuracy is sometimes used as one of the criteria to identify the best cut-off value of a test when the test result is given as quantitative data. However, it provides the same weight to the correct identification of true positives and true negatives and therefore may not be the most appropriate metric in situations where the costs or consequences of false positives and false negatives are significantly different. There are other mathematical approximations that can be used to assess the accuracy, like the Youden Index, but it also has been criticised for providing the same weight to sensitivity and specificity.

NOTE: *The results of all test accuracy measures (sensitivity, specificity, PPV, NPV, LR+, LR- and overall diagnostic accuracy), must be reported using uncertainty measures such as 95% confidence intervals.*

ACTIVITY

Which test (high-risk HPV testing or conventional cytology) would you use for primary screening and which one for triage if:

- High-risk HPV testing: sensitivity = 1, specificity = 0.87
- Cytology: Sensitivity = 0.68, specificity = 0.95

The correct answer is high-risk HPV testing for primary screening and cytology for triage

In high-resource countries, high-risk HPV testing with cytology triage is a standard procedure for early diagnosis of cervical cancer (Thomsen et al., 2020). High-risk HPV DNA testing has high sensitivity and moderate specificity, while cytology has low to moderate sensitivity but high specificity and is therefore used as a triage test. Due to the moderate specificity of high-risk HPV testing, if all positive subjects underwent colposcopy, this would result in an excess of women undergoing testing and the excess costs that this entails (low PPV). To increase the PPV, a high specificity test such as cytology is used in some settings for triage of women who test positive for HPV.

2.3. Potential biases in screening

Once screening tests have been validated and used in human populations, some biases need to be considered to assess the impact or effectiveness of screening programmes.

Lead time bias

Screening enables earlier detection of disease so that treatment can start at an early phase to boost the changes of longer survival. However, earlier detection of disease may be misleading sometimes as to the benefits it brings. A screening programme is not useful when there is an increase in survival without prolongation of life (i.e. patients die at the same time point irrespective of earlier diagnosis) (Eeles et al., 2018).

Consider two subjects with a similar natural history of disease (same onset and death date). Subject A is not screened whereas subject B is. Subject B appears to survive longer than subject A, but only because the disease was identified earlier (**Figure 4**).

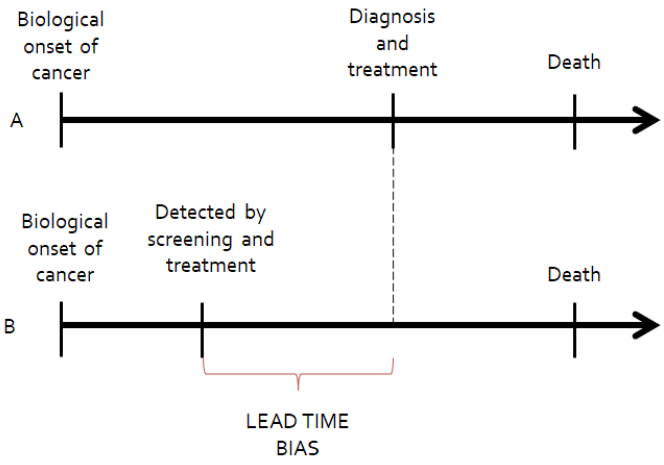


Figure 4. Lead time bias

Length time bias

Some cancers may have longer detectable preclinical phases while others may be more aggressive and progress rapidly (Eeles et al., 2018). Screening is likely to detect those cancers with longer detectable preclinical phases, which are more likely to be

more indolent than those not detected by screening (**Figure 5**). This could result in an increase in diagnoses (higher cost, potential harm and side effects of diagnosis and treatment) but not a reduction in mortality since these cancers would never have resulted in death.

Overdiagnosis bias is an extreme form of length time bias (Bunting, 2002). The detection of very indolent tumours in the screened group produces apparent increases in the number of cancer cases. Randomised trials with mortality rates as the endpoint can overcome lead and length time biases.

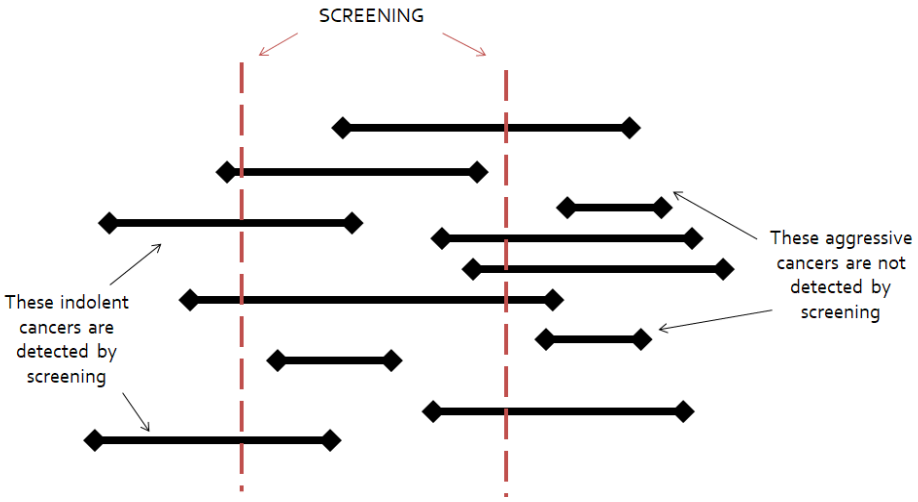


Figure 5. Length time bias

EXAMPLE

Erectile dysfunction and incontinence are common side effects of standard treatment for prostate cancer. The benefits of treatment outweigh these harms in aggressive forms of prostate cancer. However, a considerable proportion of men with prostate cancer have slow-growing tumours unlikely to be diagnosed or cause harm but which, if detected and treated, could result in the patient suffering these side effects unnecessarily (Bunting, 2002).

Selection bias

A self-selection bias may occur because participants in screening may be healthier and adhere better to cancer prevention recommendations and to treatment when they are cancer patients. Similarly, patients with higher cancer risk due to a family history of cancer may attend screening more frequently (Cox & Sneyd, 2013). The

outcomes associated with these populations may be better and therefore bias the apparent value of the screening programme.

Ascertainment or verification bias

Verification bias occurs when unequal verification of the presence of disease between test-positive and test- negative subjects is performed. For instance, participants who test positive in a screening test are further studied to confirm or discard disease, while participants who test negative are assumed to not have the disease and are no longer studied or in a smaller proportion. The calculation of test accuracy estimates (sensitivity, specificity, etc.) should take into account the possibility of this verification bias.

EXAMPLE

In a simulation exercise, Franco estimated the test accuracy for two studies conducted in a hypothetical population of 1000 women assuming 67% sensitivity and 75% specificity of a screening test under study. In one population, disease ascertainment was done for all subjects whereas in the other, ascertainment of disease was restricted to a random sample of 80% of those testing positive and 10% of those testing negative. In the population with incomplete ascertainment of disease the sensitivity was overestimated by 27% (94% instead of 67%) and the specificity was underestimated by 48% (27% instead of 75%). This example highlights the relevance of this measurement issue (Franco, 2000).

Did you know?

Despite the availability of some statistical methods to correct for the verification bias, in a recent pooled analysis, the authors pointed out that adjustment of the clinical performance of HPV testing for the ascertainment bias may overcorrect/underestimate its clinical performance in detecting truly precancerous abnormalities (i.e, the test sensitivity) due to overestimating the relevance of nonprogressive CIN2/3 (Castle et al., 2020).

UNIT 3. OPPORTUNISTIC VS. ORGANISED SCREENING PROGRAMMES

Screening is a complex public health strategy and must fulfil several criteria to be effective. Screening should be undertaken only:

- After efficacy, and ideally effectiveness, have been evaluated and established.
- When resources are available and sufficient to cover a large proportion of the intended target group.
- When it is possible to implement the required follow-up of screen-positive subjects to confirm disease and guarantee treatment options for those who are ill.
- When the disease is significant enough from a public health perspective to justify the effort and costs of screening.

Cancer screening programmes can be organised or opportunistic:

An **organised screening programme** is a population-based programme designed by a central public health structure (national or regional) to achieve the highest possible coverage equitably. All women in the target age group are invited to participate.

Organised programmes comprise systematic testing with a standardised and quality-assured test, a call and recall system involving a well-defined target population, delivery of test results and the results from additional studies, and treatment and follow-up care if necessary. The international Agency for Research on Cancer (IARC) defines an organised screening programme as one that has '*an explicit policy with specified age categories, method, and interval for screening; a defined target population; a management team responsible for implementation; a health-care team for decisions and care; a quality assurance structure; and a method for identifying cancer occurrence in the target population*' (IARC, 2022).

An **opportunistic or spontaneous screening programme** is conducted independently from an organised programme. In cervical cancer screening, women are invited to

participate when visiting health services for their convenience. Screening may be recommended by a healthcare professional during a consultation or requested by the woman. This method of screening relies on the patient-practitioner relationship.

Organised screening is considered to be more efficient than opportunistic screening, given that it implies a better use of available resources and benefits a greater number of people more equitably (Diaz et al., 2018; Dubé, 2018; Schiller-Fruehwirth et al., 2017). Organised screening is focused on obtaining higher coverage and on the quality of the process, providing benefits that outweigh any harm caused. The target age group and screening frequency are usually decided at the national or regional level. Organised screening programmes minimise inequalities in access to screening by giving every eligible person access independently of their relationship with the healthcare system. Opportunistic screening incurs on access barriers more frequently, with low screening coverage in certain groups (high-risk groups, certain ages, low-resource groups) and overuse in others. This implies a decrease in the effectiveness of the programme and its economic performance.

Table 2. Summary of characteristics of the types of screening programmes

Organised screening	Opportunistic screening
Organised implementation of the diagnostic and early treatment activities in pre-defined groups of the population at risk	Offered only to patients who voluntarily attend health services, either to get screened or for other reasons
The targeted population is clearly defined (subjects to be screened are identifiable)	Lack of organisation negatively impacts fairness, since subjects not requiring consultation for other reasons do not get screened
Generally, uses a census registry to invite the target population with recall systems for non-attendees	Generates methodological confusion between screening and clinical practice, which is hardly efficient or efficacious
Only offers validated screening techniques and has its own referral, treatment and follow-up algorithms for detected cases	Tends to unnecessarily repeat the screening test. Also, sufficient coverage is difficult to achieve

Organised screening	Opportunistic screening
Programme evaluation and monitoring are defined and planned, so incidence and mortality rates can be calculated separately for participants and non-participants at the total targeted population level	Potentially more expensive than any population-based screening programme and the results of its implementation are difficult to assess
Establishes quality control of these epidemiological data	

ACTIVITY

Read the following statements and decide if they refer to organised or opportunistic screening.

1. It is the most efficient screening delivery system in countries or settings with a publicly-funded healthcare system. It ensures high and equitable coverage and high quality in the processes involved. Achieving this requires an information system that identifies all individuals at risk for the disease, in addition to a call-recall system to reach all members of the target age group.
2. It is generally accepted as more cost-effective since it makes better use of available resources and ensures that the greatest possible number of women benefit.
3. It tends to reach younger women at lower risk, such as those attending antenatal, child health and family planning services.
4. It tends to be inefficient, though when applied with full adherence to professional guidelines it can also achieve a high reduction in disease incidence and mortality.

The correct answers are:

1 Organized screening, 2 Organized screening, 3 Opportunistic screening, 4 Opportunistic screening.

SUMMARY

- Screening involves the systematic application of safe, easy-to-use and economically affordable tests to enable early detection of disease. Screening aims to reduce disease mortality by reducing its prevalence, shortening its duration, reducing the incidence of complications associated with the disease and increasing the quality of life of the patient.
- Organised screening is a population-based programme designed by a central public health structure (national or regional) to achieve the highest possible coverage. Opportunistic or spontaneous screening is conducted independently of an organised or population-based programme: women are invited to participate when visiting health services, for their convenience.
- Measures of diagnostic test accuracy provide data about the ability of a test to:

  1) discriminate health and disease status distinguishing subjects who are ill from those who are not (sensitivity and specificity)

  2) predict disease: estimation of the post-test probability of disease given a certain test result (predictive values)
- Several criteria must be met when organising a cancer screening programme, including: a defined target population; defined age intervals and choice of screening test(s); a healthcare system with the capacity to screen, follow-up those who screened positive and provide treatment as indicated; and quality control systems, among others.

BIBLIOGRAPHY

Altman, D. G., & Bland, J. M. (1994a). Diagnostic tests 2: Predictive values. BMJ (Clinical Research Ed.), 309(6947), 102. https://doi.org/10.1136/bmj.309.6947.102

Altman, D. G., & Bland, J. M. (1994b). Statistics Notes: Diagnostic tests 1: sensitivity and specificity. BMJ, 308(6943), 1552. https://doi.org/10.1136/bmj.308.6943.1552

Andermann, A. (2008). Revisiting Wilson and Jungner in the genomic age: A review of screening criteria over the past 40 years. Bulletin of the World Health Organization, 86(4), 317–319. https://doi.org/10.2471/BLT.07.050112

Brenner, H., & Gefeller, O. (1997). Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Statistics in Medicine, 16(9), 981–991. https://doi.org/10.1002/(sici)1097-0258(19970515)16:9<981::aid-sim510>3.0.co;2-n

Bunting, P. S. (2002). Screening for prostate cancer with prostate-specific antigen: Beware the biases. Clinica Chimica Acta; International Journal of Clinical Chemistry, 315(1–2), 71–97. https://doi.org/10.1016/s0009-8981(01)00717-3

Castle, P. E., Pierz, A. J., Adcock, R., Aslam, S., Basu, P. S., Belinson, J. L., Cuzick, J., El-Zein, M., Ferreccio, C., Firnhaber, C., Franco, E. L., Gravitt, P. E., Isidean, S. D., Lin, J., Mahmud, S. M., Monsonego, J., Muwonge, R., Ratnam, S., Safaeian, M., … Xi, L. F. (2020). A Pooled Analysis to Compare the Clinical Characteristics of Human Papillomavirus–positive and -Negative Cervical Precancers. Cancer Prevention Research, 13(10), 829–840. https://doi.org/10.1158/1940-6207.CAPR-20-0182

Clavel, C., Masure, M., Bory, J.-P., Putaud, I., Mangeonjean, C., Lorenzato, M., Nazeyrollas, P., Gabriel, R., Quereux, C., & Birembaut, P. (2001). Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: A study of 7932 women. British Journal of Cancer, 84(12), 1616–1623. https://doi.org/10.1054/bjoc.2001.1845

Cox, B., & Sneyd, M. J. (2013). Bias in breast cancer research in the screening era. Breast (Edinburgh, Scotland), 22(6), 1041–1045. https://doi.org/10.1016/j.breast.2013.07.046

Deeks, J. J., & Altman, D. G. (2004). Diagnostic tests 4: Likelihood ratios. BMJ : British Medical Journal, 329(7458), 168–169.

Diaz, M., Moriña, D., Rodríguez-Salés, V., Ibáñez, R., Espinás, J. A., & de Sanjosé, S. (2018). Moving towards an organized cervical cancer screening: Costs and impact. European Journal of Public Health, 28(6), 1132–1138. https://doi.org/10.1093/eurpub/cky061

Dubé, C. (2018). Organized Screening Is Better Than Opportunistic Screening at Decreasing the Burden of Colorectal Cancer in the United States. Gastroenterology, 155(5), 1302–1304. https://doi.org/10.1053/j.gastro.2018.10.010

Eeles, R. A., Tobias, J. S., & Berg, C. D. (Eds.). (2018). Cancer prevention and screening: Concepts, principles and controversies. Wiley.

Franco, E. L. (2000). Statistical issues in human papillomavirus testing and screening. Clinics in Laboratory Medicine, 20(2), 345–367.

Franco, E. L., Mahmud, S. M., Tota, J., Ferenczy, A., & Coutlée, F. (2009). The expected impact of HPV vaccination on the accuracy of cervical cancer screening: The need for a paradigm change. Archives of Medical Research, 40(6), 478–485. https://doi.org/10.1016/j.arcmed.2009.06.003

Gray, J. A. M. (2004). New concepts in screening. British Journal of General Practice, 7.

Hesselink, A. T., Heideman, D. a. M., Steenbergen, R. D. M., Gök, M., van Kemenade, F. J., Wilting, S. M., Berkhof, J., Meijer, C. J. L. M., & Snijders, P. J. F. (2014). Methylation marker analysis of self-sampled cervico-vaginal lavage specimens to triage high-risk HPV-positive women for colposcopy. International Journal of Cancer, 135(4), 880–886. https://doi.org/10.1002/ijc.28723

IARC. (2022). Cervical Cancer Screening (Vol. 18). IARC Press. http://www.iarc.fr/en/publications/pdfs-online/prev/handbook10/

McGee, S. (2002). Simplifying Likelihood Ratios. Journal of General Internal Medicine, 17(8), 647–650. https://doi.org/10.1046/j.1525-1497.2002.10750.x

Principles of Epidemiology | Lesson 1—Section 9. (2020, May 11). http://www.cdc.gov/csels/dsepd/ss1978/lesson1/section9.html

Schiller-Fruehwirth, I., Jahn, B., Einzinger, P., Zauner, G., Urach, C., & Siebert, U. (2017). The Long-Term Effectiveness and Cost Effectiveness of Organized versus Opportunistic Screening for Breast Cancer in Austria. Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research, 20(8), 1048–1057. https://doi.org/10.1016/j.jval.2017.04.009

Schüz, J., Espina, C., Villain, P., Herrero, R., Leon, M. E., Minozzi, S., Romieu, I., Segnan, N., Wardle, J., Wiseman, M., Belardelli, F., Bettcher, D., Cavalli, F., Galea, G., Lenoir, G., Martin-Moreno, J. M., Nicula, F. A., Olsen, J. H., Patnick, J., … Working Groups of Scientific Experts. (2015). European Code against Cancer 4th Edition: 12 ways to reduce your cancer risk. Cancer Epidemiology, 39 Suppl 1, S1-10. https://doi.org/10.1016/j.canep.2015.05.009

Thomsen, L. T., Kjær, S. K., Munk, C., Frederiksen, K., Ørnskov, D., & Waldstrøm, M. (2020). Clinical Performance of Human Papillomavirus (HPV) Testing versus Cytology for Cervical Cancer Screening: Results of a Large Danish Implementation Study. Clinical Epidemiology, 12, 203–213. https://doi.org/10.2147/CLEP.S243546

Wilson, J., & Jungner, G. (1968). Principles and practice of screening for disease. World Health Organization - WHO. http://apps.who.int/iris/bitstream/10665/208882/1/WHO_PA_66.7_eng.pdf

ADDITIONAL MATERIAL

- Bandos AI, Rockette HE. Use of likelihood ratios for comparisons of binary diagnostic tests: Underlying ROC curves. Med. Phys. 37 (11):5821-5830.
- Dobrow, Mark J., Victoria Hagens, Roger Chafe, Terrence Sullivan, and Linda Rabeneck. Consolidated Principles for Screening Based on a Systematic Review and Consensus Process. CMAJ 190, no. 14 (9 April 2018): E422–29. https://doi.org/10.1503/cmaj.171154.
- Hofvind S, ponti A, Patnick J, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes.J Med Screen. 2012;19 Suppl 1:57-66.
- IARC. (2022). Cervical Cancer Screening (Vol. 18). IARC Press. http://www.iarc.fr/en/publications/pdfs-online/prev/handbook10/
- Public Health England. Criteria for appraising the viability, effectiveness and appropriateness of a screening programme. Accessible at: https://www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes/criteria-for-appraising-the-viability-

effectiveness-and-appropriateness-of-a-screening-programme. Last access: 20/11/2020

- World Health Organization. WHO report on cancer: setting priorities, investing wisely and providing care for all. Geneva, Switzerland: World Health Organization; 2020.
- World Health Organization. Comprehensive cervical cancer control: a guide to essential practice (2nd edition). Geneva: WHO editions., 2014