

Previous exam questions on medical diagnosis and overdiagnosis

A recent study of different measures of left ventricular function gave the following results for mean values in the (healthy) study group: Mean ejection fraction 59%, mean systolic shortening of the left ventricle: 17 mm, mean systolic tissue velocity: 9.1 cm/s and mean systolic strain rate (shortening velocity per length unit) – 1.1 s⁻¹. To compare the usefulness of the different measures, we look at the reproducibility of the results, doing repeated measurements. Which measure of repeatability would be best here for comparison of the accuracy of the different measures?

- A. Limits of agreement
- B. Correlation
- C. Mean error
- D. kappa coefficient

Ejection fraction measured by ultrasound has a high variability. Limits of agreement are $\pm 10\%$ points. A patient who has had an infarct had an ejection fraction of 45% (slightly reduced) when he left the hospital. Two months later, he was admitted because of dyspnea, but had no new infarction. During the second stay, a new echocardiography was made, and they now found an EF of 38%. The diagnosis of worsening heart failure (HF) was made, and medical treatment for heart failure was started. The conclusion that the patient had worsening heart failure was:

- a) Correct because the EF was below 40%
- b) Correct because the EF had dropped by 7% points
- c) Incorrect because a change of 7% points is not a significant change with this method
- d) Incorrect because an EF of 38% is not compatible with heart failure.

Ejection fraction measured by ultrasound has a high variability. Limits of agreement are $\pm 10\%$ points. A patient who has had an infarct had an ejection fraction of 48% (slightly reduced) when he left the hospital. Two months later, he was admitted because of dyspnea, but had no new infarction. During the second stay, a new echocardiography was made, and they now found an EF of 35%. The diagnosis of worsening heart failure (HF) was made, and medical treatment for heart failure was started. The conclusion that the patient had worsening heart failure was:

- A) Correct because the EF was below 40%
- B) Correct because the EF had dropped by 13% points
- C) Incorrect because a change of 13% points is not a significant change with this method
- D) Incorrect because an EF of 35% is not compatible with heart failure.

What is the sensitivity of a certain diagnostic test?

- A The probability of having a positive test if you are ill
- B The probability of having a negative test if you are healthy
- C The probability of being ill if you have a positive test
- D The probability of being healthy if you have a negative test

What do we mean by the positive predictive value of a test:

- A The probability of being ill if the test shows the patient to be ill
- B The probability of being healthy, if the test shows the patient to be healthy
- C The probability of being ill, despite the test showing the patient to be healthy
- D The probability of being healthy, despite the test showing the patient to be ill

The usefulness of a diagnostic test is dependent on its ability to separate a healthy population from the population with the disease that the test is specific for. Validity is the bias of the test measure from the "truth". Reliability or precision is the statistical spread of repeated measurements. Which factor is most important for the diagnostic usefulness?

- A The validity of the test
- B The reliability of the test
- C Both are equally important
- D None of them

A 50 year old man with no symptoms want a checkup to be sure he doesn't have coronary heart disease, for safety's sake. He went to a private specialist centre in order to have a test, and have an exercise ECG. The risk of having significant coronary disease in men without symptoms at that age may be assumed to be 4%. Exercise ECG has a sensitivity of about 65%, and a specificity of 75%. This may be expressed in the following 4x4 table:

Probability of:	Positive test	Negative test	Total
healthy	24%	72%	96%
ill	2.6%	1.4%	4%
Total	26.6%	73.4%	100%

If the test shows a positive result (meaning it indicates coronary disease), what is the positive predictive value of this test (the probability that he really is ill)?

- a) 2,6%
- b) 10%
- c) 24%
- d) 26.6%

If the test shows a negative result (meaning it doesn't indicate coronary heart disease), what is the negative predictive value of the test (the probability that he really is healthy)?

- a) 1.4%
- b) 72%
- c) 73.4%
- d) 98%

A 50 year old man with no symptoms want a checkup to be sure he doesn't have coronary heart disease, for safety's sake. He went to a private specialist centre in order to have a test, and had a stress echocardiography. The risk of having significant coronary disease in men without symptoms at that age may be assumed to be 4%. Stress echo has a sensitivity of about 80%, and a specificity of 90%. This may be expressed in the following 4x4 table:

<i>Probability of</i>	<i>Positive test</i>	<i>Negative test</i>	<i>Total</i>
<i>healthy</i>	9.6%	86.4%	96%
<i>ill</i>	3.2%	0.8%	4%
<i>Total</i>	12.8%	87.2%	100%

If the test shows a positive result (meaning it indicates coronary disease), what is the positive predictive value of this test (the probability that he really is ill)?

- A) 3.3%
- B) 10.0 %
- C) 25.0 %
- D) 31.3%

If the test shows a negative result (meaning it doesn't indicate coronary heart disease), what is the negative predictive value of the test (the probability that he really is healthy)?

- A) 99.0%
- B) 90.1%
- C) 90.0%
- D) 86.4%

A man in his mid fifties go to a doctor to check if he has prostate cancer, and asks for a blood test (PSA) that marks the presence of this. He has a blood level of PSA of 10 ng/l, which means that it is elevated. With a cut of of 4 ng/l, the test has a sensitivity of 20% and a specificity of 90%. The prevalence of prostate cancer among men between 50 and 60 years can be taken as about 1 %. What is the probability that he has cancer (positive predictive value of the test)?

- A 2%
- B 20%
- C 50%
- D 90%

What is “downstream damage” from a diagnostic screening program?

- A Long term radiation damage from excessive X-ray use
- B As screening do not result in reduction of total mortality, the downstream damage is the statistical effect that the patients will have to die from something else
- C The fact that screening programs use resources that could have been used better for treatment of other patients, those patients suffer from downstream damage of lack of optimal treatment
- D The fact that screening results in over diagnosis – patients with harmless disease, but who will receive treatment, which will in some cases have side effects.

The cut off limits between anemia and normal hemoglobin are usually set at 12 g/dl for women and 13 g/dl for men. The measurement of hemoglobin is variable, and has limits of agreement around 1.2 g/dl. A man sees his doctor for a routine checkup, and the doctor finds a hemoglobin of 13.5. He checks his records and at the last check up visit, the hemoglobin was 14.2. The doctor prescribes iron medication due to the drop in the value, in order to prevent development of anemia (dropping below 13g/dl). This decision is:

- A Correct, as the drop in hemoglobin obviously shows iron deficiency or iron loss
- B Correct anyway, everybody should take iron supplements, as the normal food in any case contains too little iron.
- C Incorrect, because the value does not fulfill the definition of anemia
- D Incorrect, because the drop in hemoglobin is not significant.