Yuning Hu
Antonio Olivas
SPA DRP final reflection
Autumn 2023

Biomarkers for Ovarian Cancer

With the expanding variety of food choices, increasing pace of life and other factors, people nowadays are easily trapped in a bad lifestyle, and thus have a high risk of disease. However, diseases such as diabetes and cancer do not have clear symptoms that allow people to diagnose disease themselves. Then, those seemingly healthy populations would need a screening test to identify or prevent disease. While imaging is expensive and biopsy is both expensive and invasive, blood tests are relatively expensive and less invasive, and therefore a good choice for screening tests.

Biomarkers are objective indications of medical states, and they can be measured accurately and reproducibly. Although diabetes and some cancers have a unified biomarker, HbA1c, there is controversy for the selection of ovarian cancer. Ovarian cancer is a common disease (1 out of 75 people who have ovaries), and it is serious that many ovarian cancers are undetected until the late stage. We know that it is necessary to find a biomarker for screening tests of ovarian cancer.

Since a person can only have the disease or don't have the disease, the screening test is binary. When operating screening tests for a biomarker, there will be 4 possible results: the person have the disease, and the test confirms that the person have the disease (true positive); the person have the disease, but the test fails show the person have the disease (false negative); the person do not have the disease, and the test confirms that the person do not have the disease (true negative); the person do not have the disease, but the test instead shows the person have the disease (false positive). Define sensitivity to be the ability to detect the disease in the disease population, i.e. true positive / (true positive + false negative). Define specificity to be the ability to claim non-disease in the non-disease population, i.e. true negative / (true negative + false positive). Since screening tests are for populations that are healthy from appearance, in the context of statistics, the null hypothesis is that the person does not have the disease. Then, 1 - sensitivity represents Type 2 Error and 1 - specificity represents Type 1 Error.

During this quarter, I have learned to visualize the performance of biomarkers by the ROC (receiver operating characteristic) curves. By fixing a cutoff and operating a binary test based on the cutoff, the ROC curve records the performance of a biomarker at varying threshold values. The x-axis and y-axis represent sensitivity (true positive rate, from 0 to 1) and 1- specificity (false positive rate, from 0 to 1, or specificity, from 1 to 0) respectively. Accordingly, AUC, the area under the curve, for the ROC curve is the integration over the ROC curve. Statistically, it

means the expected performance of the biomarker. Higher AUC is higher performance on average. When sensitivity is the same as specificity, the biomarker has no ability to distinguish the disease population from the total population. In the ROC curve, this situation corresponds to the diagonal, and its AUC is 0.5. Therefore, if the ROC curve for a biomarker is under 0.5, it is considered useless.

When already having a best biomarker, scientists would have to find an optimal threshold differentiating the disease and non-disease population. The standard approach to find this cutoff is to penalize Type 2 Error and Type 1 Error, which is, minimizing (1 - Sensitivity) +(1 - Specificity). However, it is worth noting that Type 2 Error and Type 1 Error may have different weights, instead of the same weight that the standard approach suggested. If I'm asked whether I would prefer "I have the disease, but the doctor says I don't" or "I don't have the disease, but the doctor says I have", I would say the first is worse, because I could prevent my disease to deteriorate if I knew I have the disease. Therefore, in my modified calculation for the cutoff, I would give 1.5 weight to Type 2 Error, which is, minimizing 1.5* (1 - Sensitivity) +(1 - Specificity).

Applying the studies onto the case of ovarian cancer, I used the dataset which comes from 349 patients with tumors. There are 6 biomarkers: CA19-9: most commonly used for ovarian cancer; CA72-4: most commonly used for gastric cancer; AFP: related to cancer of the liver, ovaries or testicle; CA125: related to ovarian cancer; HE4: significant for epithelial ovarian cancer; CEA: certain types of cancer can increase CEA levels, but one can have high CEA without having cancer. Comparing by eye-baling and the DeLong method, the result is that HE4 has the best performance. However, considering that HE4 is expensive and not available for every institution, I chose CA125 as an alternative, as it is the second best. The original cutoff for HE4is 70.73, and the modified is 71.44, corresponding to the findings of researchers from different countries, which is 70. The original and modified cutoff for CA125 are both 98.445, which is a huge difference with the cutoff from test manufacturers. There are two possible explanations for this difference. One is a leap in the dataset, another is that manufacturers' cutoff is determined from distribution in healthy individuals to include 99% of the normal population, while the dataset contains unhealthy individuals.

We can conclude for the case of ovarian cancer that HE4 is the best biomarker, but for more affordable and accessible options, institutions can use CA125. However, it is worth pointing out that although ROC is frequently used in the medical fields, there may exist better ways to find the biomarker or the cutoff.