Christina Martin
Kazi Russell
MED INF 406 – Decision Support Systems & Healthcare
October 26, 2014

Group Project PPV & NPV
(Session 5)

The positive and negative predictive values are important in studies pertaining to serological tests, disease prevalence, or when evaluating the feasibility or success of a screening or diagnostic test. A serological test analyzes the immunological properties and actions of serum. This analysis is important in being able to detect the presence of antibodies common in certain diseases or antibodies that fight against specific diseases. For example, celiac disease (NDDIC, 2014) can be detected by three antibodies common to the disease (i.e.: anti-tissue transglutaminase (tTg) antibodies, endomysial antibodies (EMA), or deamidated gliadin peptide (DGP) antibodies). Given this information, there needs to be a way to evaluate the quality or performance of such a test to size up its validity (*Note that the gold standard for testing for Celiac Disease is an intestinal biopsy showing villous lesions). This can be accomplished by examining the **positive and negative predictive values** of the diagnostic test itself. In other words the proportion of tests that will have a ***true positive*** or ***true negative*** result will be indicative of how reliable the test is for detecting celiac disease. These predictive values can also be used for ***screenings*** that determine whether or not a ***diagnostic*** test should be given (i.e.: A patient with or without a disease precursor who has a mammogram screening versus a biopsy to first determine any possible risk factors for breast cancer such as a tumor that cannot be felt). More specifically, in a screening that yields high sensitivity there will be more false positives (i.e.: positive results but disease not actually present) which in turn means that there will be fewer false negative results. A diagnostic test that yields high specificity will show more false negatives (results show disease is not present but it really does exists) and therefore there will be less false positive results. Sensitivity and specificity are therefore measures of a test's or screening's ability to correctly identify a patient as having or not having a disease. Henceforth, the positive predictive value is the probability that a patient who test positive for a disease such as celiac or breast cancer actually has the disease, and a patient who test negative for any one of these diseases really is actually free of disease. This can best be illustrated by a 2X2 table:

	Diseased	Not Diseased
Test Positive	132	983
Test Negative	45	63650
	177	64,633

The 2 x 2 table above shows the results of the evaluation of a screening test that was conducted in 64,810 subjects. The gold standard indicated that 177 of these actually had the disease, and the other 64,633 subjects were not diseased. If we focus on the rows, we find that 1,115 subjects had a positive screening disease, i.e., the test results were abnormal and suggested disease. However, only 132 of these were found to actually have disease, based on the gold standard test. Also note that 63,695 people had a negative screening test, suggesting that they did not have the disease, BUT, in fact 45 of these people were actually diseased.

The probability of the screening test correctly identifying diseased individuals or sensitivity = 132/177 = 74.6%. The probability of the screening test correctly identifying non-diseased individuals or specificity = 63650/64633 = 98.3%

Another measure that is important and works in accordance with both the negative and positive predictive value is prevalence. In epidemiology (i.e.: the context with which this essay refers to regarding disease presence), prevalence refers to the proportion of a population found to have a disease. Prevalence answers the question "how many people have this disease right now." Prevalence unequivocally affects the predictive value of any test. For example, depending on the clinical environment in which the test is applied, the predictive accuracy will be contingent upon the number of people in the population who have the disease. So if prevalence increases from 5% (i.e.: breast cancer among 35 year-olds) to 25% (i.e.: breast cancer among 50 year-olds), then the predictive positive value is expected to rise as well. As you can clearly see, the relevance of the predictive value is important to the interpretation of test results or studies where population statistics matter in many clinical trials (i.e.: falling prevalence lends itself to false positive results). Sensitivity, specificity, and prevalence are all measures that are needed to give relevance to both the Positive and Negative Predictive values. Two by two tables are a valuable representation of a population or sample of a population that have (test positive) or do not have a disease (test negative). The data in these tables can be used to approximate cells with the use of descriptive statistics (i.e.: learning the probability for having a disease if you belong to a particular culture or were exposed to an environmental agent). Risk ratios can be determined given certain measures examined as well (i.e.: use of Chi Square test to test if people in low income communities are more at risk for diabetes than people from middle class communities). In concluding, it is evident that use of the PPV and NPV along with the 2X2 table are critical agents that render valuable information in clinical studies.

References:

National Digestive Diseases Information Clearninghouse (NDDIC). (2014). Testing for Celiac Disease. Retrieved from: http://digestive.niddk.nih.gov/DDISEASES/pubs/celiactesting/index.aspx

Wikipedia. (2014). Prevalence. Retrieved from: http://en.wikipedia.org/wiki/Prevalence