Confidence limits for prevalence of disease adjusted for estimated sensitivity and specificity

Zsolt Lang, Jenő Reiczigel

Szent István University, Faculty of Veterinary Science, Budapest, Hungary

Aims

Prevalence of a disease is usually assessed by diagnostic tests that may produce false results. Rogan and Gladen (1978) described a method to estimate the true prevalence and its variance correcting for sensitivity and specificity of the diagnostic procedure. Reiczigel et al. (2010) provided exact confidence intervals for the true prevalence assuming sensitivity and specificity were known. When sensitivity and specificity are not known a priori but estimated from a sample, the uncertainty of their estimates will increase the variance of the prevalence estimate, and this must be taken into account in the confidence interval construction. In this paper we construct approximate confidence intervals for the true prevalence when sensitivity and specificity are estimated from independent samples.

Methods

The confidence interval using the Rogan-Gladen variance estimate with normal approximation performs as poorly as the Wald interval for a binomial proportion usually does. Since the exact computation seemed to be computationally infeasible we adapted the "add 2 successes and 2 failures" method proposed by Agresti and Coull (1998) and found that this made the method quite acceptable. We studied the coverage of this procedure by simulation for various values of sensitivity, specificity and prevalence.

Results

According to an extensive simulation study the new confidence intervals maintain the nominal level fairly well even for sample sizes as small as 30; minimum coverage is above 88%, 93%, and 98% at nominal 90%, 95%, and 99%, respectively. The accuracy of coverage was preserved even when the three sample sizes differed considerably. The loss of precision due to application of confidence interval procedures assuming known sensitivity and specificity in such cases when these are estimated becomes negligible only when sample sizes for sensitivity and specificity are five to ten times greater than that for estimating prevalence.

Conclusion

In practical research situations it is quite typical that sensitivity and specificity are determined from samples of a few hundreds or even less, and when the diagnostic procedure is applied to screening, the samples are much larger (typically a few thousands). Simulation results showed that in these cases treating sensitivity and specificity as known values leads to invalid interval estimates. Our proposed CI provides good coverage for a wide range of prevalence, sensitivity and specificity values, including prevalence close to 0 and 1 and sensitivity and specificity close to 1.

References

Agresti A, Coull BA (1998) Approximate is better than 'exact' for interval estimation of binomial proportions. Am Stat 52, 119–126.

Reiczigel J, Földi J, Ózsvári L (2010) Exact confidence limits for prevalence of a disease with an imperfect diagnostic test. Epidemiol Infect 138, 1674–1678.

Rogan WJ, Gladen, B (1978) Estimating prevalence from the results of a screening test. Am J Epidemiol 107, 71–76.

Research was supported by TÁMOP 4.2.2/B-10/1-2010-011 „Development of a complex educational assistance/support system for talented students and prospective researchers at the Szent István University".