EDITORIAL

Section Editor: Diane D. Davey

False-Negative Rate of Cervical Cytology:
Sense and Sensitivity

Raimond W.M. Giard, M.D., Ph.D.*

In medicine, there has been a gradual shift from individual to public health and from curative to preventive care, with focus on the secondary prevention of cancer by early diagnosis. Mass screening for cervical and breast cancer has been implemented in most Western countries, but not without problems. Cancer is repeatedly diagnosed after a "normal" outcome of the test (either a cervical smear or a mammography), and this delay in diagnosis caused by a false-negative test has become a frequent cause for litigation.[1] False-negative outcomes unmistakably are a problem and deserve further attention and, when feasible, prevention. We will focus in this commentary on cervical cytology, but many of these problems are shared by mammographic screening.

Suitability of a Cancer Screening Test

What makes a test suitable for cancer screening? It should be simple, nontraumatic, inexpensive; precise estimates of its diagnostic properties must be known; and above all, it should have been properly put to the test. Knowing the correct test characteristics of cervical cytology is of consequence, because these are fundamental to recommendations for optimal frequency of screening cycles, the management of minor abnormalities, and the rational choice between different operation methods for basically the same cytology test. Ideally, not only the test as a whole but also the properties of each outcome class should be characterized separately.[2] Estimation of the false-negative rate of cancer screening tests such as smears, however, is much more complicated then in symptomatic individuals. Factors pertinent to test assessment are: the demographic and social characteristics of the population, the disease spectrum, the choice of cutoff point on the ordinal Papanicolaou or other (e.g., Bethesda) scale, how long and how exhaustive follow-up has been, and the choice of the proper "gold standard."[3]

Among a glut of publications on cervical cytology, very few articles meet with the methodological criteria for proper test evaluation. Most studies are heavily biased, very few explore the final clinical outcome of all test results through meticulous follow-up, and the clinical and/or histologic reference standard itself is often flawed. In their excellent and still very timely article, Ransohoff and Feinstein pointed out the problems in evaluating the efficacy of a diagnostic test.[4] A recent systematic literature survey on the accuracy of cervical cytology identified only 12 articles with appropriate data and the least biased estimates of test properties.[5] Sensitivity ranged from 30–87%, and when limited to four publications, the highest quality score was between 30–58%. A few years ago, a meta-analysis of Papanicolaou test accuracy also expressed concerns about the poor methodological quality of studies and produced a weighed mean sensitivity of 58%.[6] A report on 846 women, who were ultimately diagnosed with cervical (pre)neoplasia, where conventional light microscopy was compared with interactive neural network-assisted screening with 7 yr of follow-up, found a false-negative rate of 45% for both diagnostic modalities.[7] An additional problem is that cervical cytology is not reported as "normal" or "abnormal," but in different categories using an ordinal scale. Each outcome class has different diagnostic properties, which are only demonstrable when using the likelihood ratio concept.[8] These performance figures, heavily dependent on the prevalence and spectrum of the disease and the choice of the cutoff point, are relevant not only to professionals but to the public as well: every woman should know about the certainties and limitations of the test.[9]

Several studies indicate that sensitivity figures may be much lower than is generally assumed. The Court of Appeal in England recently ruled that, since sensitivity in screening is paramount, the false-negative rate should be reduced, preferably to or below 5%.[10] If the actual false-negative rate

Department of Clinical Pathology, St. Clara Hospital, Rotterdam, The Netherlands

*Correspondence to: Raimond W.M. Giard, M.D., Ph.D., Department of Clinical Pathology, St. Clara Hospital, PO Box 9119, 3007 AC Rotterdam, The Netherlands. E-mail: raimondgiard@hotmail.com

Received 6 March 2001; Accepted 25 June 2001

of cervical cytology seems to be so high, can we really do better? This calls for further research and exploration of opportunities for improvement, but how?

More Research and Evaluation

Potential paths for further investigation are: correct testing of the test to know exactly where we stand, introduction of innovative technology, and last but not least rethinking of diagnostic and therapeutic strategies.

Correct Testing of the Test

Poor accuracy may limit the clinical value of a test, but on the other hand, a near-perfect test will not automatically ensure an improvement in patient outcome. Just as in any field of medicine, we have to struggle with uncertainty, but with appropriate test evaluation we are able to quantify it and subsequently we are able to deal with it. There is one particular point worth considering in screening: when using a single test in asymptomatic persons, we are confronted with the full catastrophe of the erroneous test outcome. No clinical information and no positive result from any other test may "correct" the false-negative test outcome. This probably is the principal reason why we are confronted with the sharp rise in litigation for missing cancer.[1] We are awaiting the long-term results from diagnostic multimodality trials such as the ASCUS-LSIL Triage Study, which may elucidate the best testing strategies.[11]

Introduction of Innovative Technology

Often, the solution of a diagnostic problem is sought in introducing ancillary technologies. During the last few years, we have seen many such examples in cervical cytology: morphometry, image analysis, computer-aided screening, interactive neural network-assisted screening, liquid-based technology, human papilloma virus-typing, and so on. A clear but discouraging pattern can be observed in all these examples: introduction of the new technique, excitement, uncritical acceptance, and finally disappointment.[4, 12] There is a very fundamental reason for this peculiarity. Neoplastic and preneoplastic lesions of the cervix are morphologically defined by their histologic appearance. The multistep process of carcinogenesis is complex, and many factors may be identified as contributing to the etiology and pathogenesis of cervical cancer. None of the biochemical, genetic, or viral factors can serve as a better alternative for the morphologic disease definition. So, as long as we are tied down to morphology for diagnosis, we are confronted with its intrinsic limitations. The resolving power of cervical cytology between normal/abnormal or reactive changes/(pre)neoplastic lesions is limited, as is human cognition. No magistrate can enforce a rule of law resulting in a low false-negative rate of this test, and no advanced technology can overrule fundamental definition problems underlying the diagnostic difficulty, unless it provides a novel and much sharper disease definition. There is a delicate balance between sensitivity and specificity: augmenting sensitivity lowers specificity, and for each type of cancer screening, this balance has to be established, weighing all relevant factors. The only approach whereby a zero false-negative rate is achieved is to state beforehand that all test outcomes are abnormal.

Rethinking Diagnostic and Therapeutic Strategies

If we must decide to go on or to think about changing diagnostic or therapeutic strategies, the choice needs to be rational, and not emotional; for proper judgment, we need sound data. Time trends may be deceptive to ascertain if a test result finally affects patient outcome, and hence randomized comparisons are required. The ultimate proof of whether or not, and if so how much, cervical screening works should come from randomized clinical trials, but this hardly seems feasible, given the already settled nature of cervical cytology services. With creativity, however, we might still be able to uncover some of the impact of Papanicolaou smears.[13] With the utilization of regional or national computerized databases, linking test outcomes with clinical findings, we should be able to estimate test characteristics.

Conclusions

The disappointing results of screening for cervical cancer have defied the apparent logic of screening for cervical cancer based on scientific speculations.[14] Cervical cancer is not always detectable, it is not always preventable, and it is not always curable. Not only is it time to reflect, it also is time to tell.[9] The recent report from the UK National Screening Committee urges a shift from overselling to demystification, by explaining the limitations and risks associated with screening, together with its benefits, but this will take considerable time.[15] In screening for cancer, the false-negative rate may be a serious but not insurmountable problem, which deserves more attention. The general opinion is that we want women to continue to participate in screening programs for cervical cancer, because this can prevent morbidity and mortality. Above all, we need to consider the sense and the sensitivities of screening.

References

1. Wilson RM. Screening for breast cancer and cervical cancer as a common cause for litigation. Br Med J [Clin Res] 2000;320:1352–1353.
2. Giard RWM, Hermans J. The evaluation and interpretation of cervical cytology: application of the likelihood ratio concept. Cytopathology 1993;4:131–137.
3. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926–930.
4. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926–930.
5. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, Matchar DB. Accuracy of the Papanicolaou test in screening for

and follow-up of cervical cytologic abnormalities: a systematic review. Ann Intern Med 2000;132:810–819.

6. Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. Am J Epidemiol 1995;141:680–689.
7. Doornewaard H, van der Schouw YT, van der Graaf Y, Bos AB, Habbema JD, van den Tweel JG. The diagnostic value of computer-assisted primary cervical smear screening: a longitudinal cohort study. Mod Pathol 1999;12:995–1000.
8. Giard RWM, Hermans J. The evaluation and interpretation of cervical cytology: application of the likelihood ratio concept. Cytopathology 1993;4:131–137.
9. Anderson CM, Nottingham J. Bridging the knowledge gap and communicating uncertainties for informed consent in cervical cytology screening; we need unbiased information and a culture change. Cytopathology 1999;10:221–228.
10. Perrey and Ors versus East Kent Health Authority. Court of Appeal (Civil Division) Case QBENF1999/0294/1, 16 November 1999.
11. Schiffman M, Adrianza ME. ASCUS-LSIL Triage Study. Design, methods and characteristics of trial participants. Acta Cytol 2000;44: 726–742.
12. Kurtycz DF, Hoerl HD. Thin-layer technology: tempered enthusiasm. Diagn Cytopathol 2000;23:1–5.
13. Bossuyt PMM, Lijmer JG, Mol BWJ. Randomised comparisons of medical tests: sometimes invalid, not always efficient. Lancet 2000; 356:1844–1847.
14. Raffle AE. How long will screening myths survive? Lancet 1999;354: 431.
15. Kmietowicz Z. Warn the public of screening limitations, staff told. Br Med J [Clin Res] 2000;321:914.