

SOUNDING BOARD

SCREENING FOR HIV: CAN WE AFFORD THE FALSE POSITIVE RATE?

WE are a testing culture: we test our urine for drugs; we test our sweat for lies. It is not surprising that we should also test our blood for the acquired immunodeficiency syndrome (AIDS). But before we screen low-risk groups for antibody to the human immunodeficiency virus (HIV), we should consider what the results will mean. Tests for HIV antibody appear to be characterized by extraordinarily low false positive rates. Even so, positive initial and confirmatory tests in someone at low risk of HIV infection are by no means synonymous with infection, because of the possibility of false positive results. Furthermore, any increase in the false positive rate could turn a screening program into a social catastrophe.

Whatever its scientific merits, widespread HIV-antibody testing is becoming a political reality. Blood banks screen potential donors; the armed forces test recruits and personnel on active duty; the State Department tests Foreign Service officers and their dependents; and the Peace Corps and Job Corps test their applicants. Soon, screening of immigrants, prisoners in federal penitentiaries, and perhaps veterans will begin. Pregnant women have been advised to undergo testing in both the first and third trimesters.¹ President Reagan has suggested that applicants for marriage licenses should also be screened.²

Plans to test low-risk populations for HIV antibody generally ignore the possibility of false positive results. When screening of blood donors began two years ago, decontaminating the blood supply was an urgent need; it justified the assumption that confirmatory testing could identify most, or at least enough, of the testing errors. But before we establish a public policy of widespread screening, we should consider whether testing that is justified in the blood bank is also justified in other settings. If the false positive rate is not virtually zero, screening a population in which the prevalence of HIV is low will unavoidably stigmatize and frighten many healthy people. How will these mistakes change the lives of the unfortunate persons who are incorrectly identified as infected? Will such screening affect the course of the AIDS epidemic? Does the benefit of identifying infected persons justify the personal and social burden of false positive tests?

CHARACTERISTICS OF THE TESTS

The central issue is the false positive rate of tests for HIV infection. Current screening programs use a sequence of tests, starting with an enzyme immunoassay. Serum samples yielding repeatedly positive results on enzyme immunoassay are subjected to more complicated and expensive confirmatory testing, typically with a Western blot. A positive confirmatory test is considered evidence of HIV infection.

The results of screening among blood donors allow us to deduce an upper limit for the false positive rate in testing conducted to date. In 1985 and 1986, 0.01 percent of female blood donors in Atlanta and of both male and female blood donors in the north-eastern Netherlands had antibody to HIV on both enzyme immunoassay and Western blot assay.^{3,4} In the worst case, if none of those blood donors were truly infected, then the highest possible false positive rate for the pair of tests would be 0.01 percent. Because some of those blood donors were truly infected, the false positive rate was almost certainly even lower. If we make the best-case assumption that the probability of a false positive Western blot is independent of the probability of a false positive enzyme immunoassay, or if we have data about the false positive rate on Western blot tests among patients with false positive enzyme immunoassays, the joint false positive rate of the two tests in sequence will equal the product of their false positive rates. One recent study found that the false positive rates of six commercial enzyme immunoassay kits used to test blood from donors ranged from zero to 0.42 percent.⁵ Another study noted variations in false positive rates of enzyme immunoassays, even among different batches of one manufacturer's kit.⁶ Other investigators have found that the false positive rate of enzyme immunoassays can be as high as 6.8 percent among hospitalized patients.⁷

Confirmatory tests are intended to distinguish false positive results of enzyme immunoassays from those that truly represent HIV infection. Here, variations in the false positive rate may be even more important. The Western blot, the most common confirmatory test for HIV antibody and a standard against which new techniques are evaluated, is complex and very labor intensive. Its techniques have not been standardized, and the magnitude and consequences of interlaboratory variations have not been measured. Its results require interpretation, and the criteria for this interpretation vary not only from laboratory to laboratory but also from month to month. When widespread Western blot confirmation of positive findings on enzyme immunoassays began in 1985, a band indicating the presence of antibody to a protein of 24,000 to 25,000 daltons was regarded as evidence of infection. Some laboratories report this as a 24-kd band, whereas others report it as a 25-kd band. Within a year, many investigators had concluded that apparent bands in this region could represent artifacts and that even a definite band there was not specific for HIV infection.⁸

By mid-1986, the U.S. Army had adopted criteria that required either a band at 41 kd or bands at both 24 and 55 kd. But when investigators from the Army HIV-testing program sent panels of 15 serum samples from healthy adults at low risk to five large commercial firms offering HIV Western blot testing, six different specimens were classified as positive. All samples had yielded repeatedly negative results at the Walter

Reed Army Institute of Research. Three laboratories considered 1 of 15 specimens positive; one considered 3 positive.⁹

Within several months of the report from Walter Reed, investigators in both Sweden and Paris reported what they considered false positive results on Western blot tests despite the presence of both 25- and 55-kd bands. Their conclusion was based on the absence of risk factors in the individual blood donors and of concordant findings on confirmatory tests in research laboratories.^{10,11} Reactivity to the cultured human cells in which the virus had grown served to explain two unexpectedly positive Western blots.^{12,13} To find that explanation, one patient's serum was examined in three research laboratories. Other investigators have reported instances in which one specimen from a patient yielded results on a Western blot that were interpreted as positive, whereas subsequent specimens from the same patient yielded negative results.^{14,15} Several abstracts presented at the recent Third International Conference on AIDS described extensive retesting and follow-up of "atypical positive" results that would clearly be considered negative according to the U.S. Army criteria published a year earlier.¹⁶⁻²⁰ Another study described very sensitive Western blot tests that even showed reactivity in the 41-kd region to serum from normal donors at low risk for HIV infection.²¹ Thus, the lack of standardization persists.

A recent Army study compared the interpretation of the first Western blot performed with the final classification of the specimens after more extensive investigation. Among specimens that were repeatedly positive on enzyme immunoassay, the false positive rate was 1.17 percent.²² If the false positive rate of enzyme immunoassays is about 0.4 percent, the joint false positive rate of the two tests performed sequentially should be about 0.005 percent. A pair of tests with a joint false positive rate of 1 per 20,000 is unusual in clinical medicine.

These reports reflect the difficulty, uncertainty, and even disagreement that characterize testing for antibody to HIV. They suggest that positive results from low-risk populations deserve thoughtful interpretation and perhaps further testing. Despite these technical difficulties, laboratories testing blood donors and military recruits have achieved a very high standard of performance. However, specimens collected in more widespread screening programs might not all be analyzed in reference laboratories or with the same techniques. Decentralized testing might further compromise standardization. Smaller laboratories could not offer the research methods that are sometimes used to verify positive Western blot findings in persons at low risk. Technicians processing the specimens might not be as skilled as those who have developed the technique, and laboratories performing a large number of tests might be less inclined to scrutinize positive results. Interlaboratory variation in test characteristics may increase as a new generation of tests

(under development by more than 25 companies) becomes available.²³ Some new tests have been proposed to be used as a one-stage procedure, thus eliminating the extra protection of an independent confirmatory test.^{24,25}

PREVALENCE OF INFECTION

What do we know about the prevalence of HIV infection? Perhaps 50 percent of homosexual men in San Francisco have serologic evidence of the infection. The prevalence of seropositivity among intravenous drug abusers and among patients with hemophilia who received factor VIII concentrate pooled before the advent of heat inactivation is similar.^{3,8} At somewhat lower risk are patients who received repeated transfusions of red cells, platelets, and plasma before routine HIV testing of donated blood began in 1985. Antibody testing of one group of patients with leukemia treated between 1978 and 1985 showed that about 5 percent became seropositive. The patients who became seropositive had received an average of 164 units of blood products.²⁶

Other segments of the population are at much lower risk. Screening of military recruits has shown 0.16 percent of the men and 0.06 percent of the women to be seropositive.²⁷ When antibody screening of donated blood began in 1985, 1 unit of blood in 2500 had HIV antibody.²⁸ At that rate, the chance of infection from 2 units of blood donated before antibody screening began would be about 0.08 percent. Among female blood donors, as noted, the reported prevalence of seropositivity is 0.01 percent. Some of these donors may have had sexual contact with members of known high-risk groups; among women without such contact, the prevalence of infection may be even lower than 0.01 percent.

MEANING OF POSITIVE TESTS

Test sensitivity is not the issue here, and to emphasize our concern with the false positive rate, our analysis makes the best-case assumption that the combination of enzyme immunoassay and Western blot testing for HIV is 100 percent sensitive, identifying all persons who are infected. The meaning of positive tests will depend on the joint false positive rate. Because we lack a gold standard, we do not know what that rate is now. We cannot know what it will be in a large-scale screening program. However, we can be fairly sure that without careful quality control, it will rise.

Bayes' rule allows us to calculate the probability that a person with positive tests is infected.²⁹ Imagine testing 100,000 people, among whom the prevalence of disease is 0.01 percent. Of the 100,000, 10 are infected; 99,990 are not. A combination of tests that is 100 percent sensitive will correctly identify all 10 who are infected. If the joint false positive rate is 0.005 percent, the tests will yield false positive results in 5 of the 99,990 people who are not infected. Thus, of the 15 positive results, 10

will come from people who are infected and 5 from people who are not infected, and the probability that infection is present in a patient with positive tests will be 67 percent.

Figure 1 shows the consequences of screening in four populations. The implications of positive test results depend on the joint false positive rate. The horizontal axis shows a range of joint false positive rates from 0 to 0.5 percent. If the prevalence of infection is 5 percent or higher, more than 90 percent of persons with positive tests will truly be infected, whether the joint false positive rate is 0 or 0.5 percent. Unfortunately, this is not true in populations at lower risk. The probability that infection is present in a male army recruit with positive tests is 97 percent if the joint false positive rate is 0.005 percent, and 94 percent if the joint rate is 0.01 percent, but it will be only 62 percent if the joint rate rises to 0.1 percent. The probability that infection is present in a female blood donor with positive tests is about 67 percent if the joint false positive rate is 0.005 percent, and about 50 percent if the joint rate is 0.01 percent, but it will be only 9 percent if the joint rate rises to 0.1 percent. In other words, at this higher joint false positive rate, 10 women without HIV infection will be falsely identified as infected for each truly infected blood donor found. If the joint false positive rate increases to 0.5 percent, as might occur in a single-stage testing program, then 50 women without HIV infection will be stigmatized for every truly infected person identified.

The joint false positive rate may rise if single-stage testing is introduced into physicians' offices; a false positive rate of 0.6 percent was recently reported for such a test.^{24,25} The joint rate will rise if tests are performed and interpreted less carefully when the amount of testing increases substantially. Finally, it

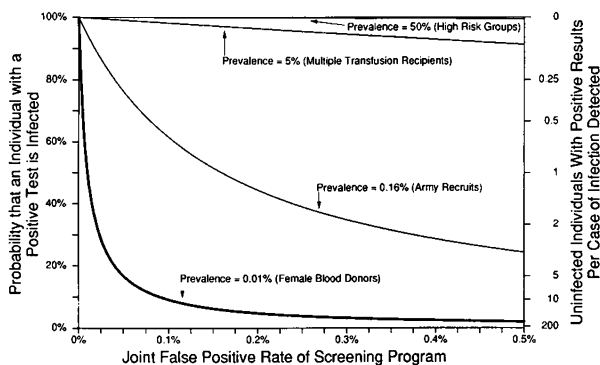


Figure 1. Meaning of Positive Screening Tests for HIV.

The horizontal axis shows the joint false positive rate of the tests. The left vertical scale shows the probability that HIV infection is present in a person with positive tests. The right vertical scale shows the number of uninfected persons falsely classified as infected for every infected person correctly identified. Sensitivity is assumed to be 100 percent. The four lines correspond to four populations that might be screened, each of which has a different prevalence of HIV infection. The boldface line represents low-prevalence populations such as those in which screening has recently been proposed.

will rise if criteria for defining a positive Western blot test are less stringent than those observed by the military and the Red Cross.

CONSEQUENCES OF WIDESPREAD SCREENING

How many cases of infection can we hope to prevent by screening groups at low risk? It is not clear how many of the few infected persons identified would have transmitted the virus to their sexual partners and children, or that testing will substantially reduce the transmission rate.^{8,30-34} Screening blood donors prevents transmission because we do not transfuse the blood. But how much does screening change behavior? By no means all seropositive persons are persuaded to practice "safer sex."³⁵⁻³⁷ Apparently only a minority abstain from childbearing.³⁸ What can we expect to happen when we screen other populations? We do not know what changes it would make in public health and our society.

Before we test, we should think again about the ethics of screening and about the social consequences of positive tests for HIV antibody. The first proposals to screen blood donors elicited widespread discussion of the potential threat to individual privacy. Special procedures were devised to ensure that this sensitive information remained private. The statutory requirement of HIV testing would in all likelihood eliminate such protection. The Secretary of Education has suggested that positive test results should be reported not only to public health authorities but also to the sexual partners of the person tested.³⁹

Despite educational efforts, public understanding of the epidemic is limited. As we contemplate recommendations and regulations, we should remember that most people consider a "positive AIDS test" to be a sentence to ghastly suffering and death. Patients with such results will take little comfort in Bayes' rule and will be offered little reassurance by their insurers, employers, and acquaintances.

A TIME FOR CAUTION

The AIDS epidemic frightens us all. But we should not allow our fear to cloud our judgment. Hasty and indiscriminate screening for antibody to HIV is imprudent and potentially dangerous, whether we suggest the tests to young women, require them of engaged couples, or impose them on our veterans. Although screening of blood donors and military recruits appears to have generated few false positive results, we do not know whether this performance can continue if the testing programs are expanded. Standardization and quality control should come first. These will take time and money; monitoring laboratory performance will require continuing effort, expenditure, and regulation.

Nor will our problems be purely technical. HIV screening poses questions that are at once scientific, political, legal, and philosophical. If laws are to link our fates to test results, should not due process be brought to the benches where those tests are performed? We will need guarantees not only of the confi-

dentiality of test results but also of the quality of the testing procedure. Should everyone be subjected to tests of uniform sensitivity and specificity, or should performance characteristics be tailored to the clinical situation? Should screening programs in the general population sacrifice specificity by adopting the highly sensitive tests designed to protect the blood supply? In the past, inexplicably positive results in persons at no apparent risk of HIV infection prompted extensive investigation of the specimens in research laboratories. Wider screening will inevitably yield more unanticipated positive results — perhaps far more than researchers can review. How will we decide whose positive results we scrutinize? Who will weigh the scientific evidence against the skepticism of the person who does not believe his positive test results? Will we recognize the results of tests performed in other countries? How often will we retest and reclassify on the basis of technical advances or because of the passage of time?

If we want to test each other, we should make a deliberate choice of the threshold probability of infection above which we will screen. We should make explicit the trade-offs implicit in any testing program. How many engagements should end to prevent one infection? How many jobs should be lost? How many insurance policies should be canceled or denied? How many fetuses should be aborted and how many couples should remain childless to avert the birth of one child with AIDS?

New England Medical Center
Boston, MA 02111

KLEMENS B. MEYER, M.D.
STEPHEN G. PAUKER, M.D.

Supported in part by a training grant (7044) and a research grant (4493) from the National Library of Medicine, Bethesda, Md.

REFERENCES

- Gruson L. AIDS toll in children is called 'deadly crisis.' *New York Times*. April 9, 1987:B8.
- Boffey PM. Reagan urges wide AIDS testing but does not call for compulsion. *New York Times*. June 1, 1987:A1.
- Francis DP, Chin J. The prevention of acquired immunodeficiency syndrome in the United States: an objective strategy for medicine, public health, business, and the community. *JAMA* 1987; 257:1357-66.
- Sibinga CTS, de Vries B, McShine RL, Das PC, Schooley RT. HIV antibody screening in a low-risk population (blood donors). *Transfusion* 1987; 27:215-6.
- Reesink HW, Lelie PN, Huisman JG, et al. Evaluation of six enzyme immunoassays for antibody against human immunodeficiency virus. *Lancet* 1986; 2:483-6.
- Füst G, Ujhelyi E, Héjjas M, Hollán SR. Traps of HIV serology: independent changes in sensitivity and specificity of ELISA kits. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:49.
- Cockerill FR, Edson RS, Chase RC, Katzmann JA, Taswell HF. "False-positive" antibodies to human immunodeficiency virus (HIV) detected by an enzyme-linked immunosorbent assay (ELISA) in patients at low risk for acquired immunodeficiency syndrome (AIDS). Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:34.
- Institute of Medicine, National Academy of Sciences. *Confronting AIDS: directions for public health, health care and research*. Washington, D.C.: National Academy Press, 1986:51,69,306.
- Burke DS, Redfield RR. False-positive Western blot tests for antibodies to HTLV-III. *JAMA* 1986; 256:347.
- Couroucé A-M, Muller J-Y, Richard D. False-positive western blot reactions to human immunodeficiency virus in blood donors. *Lancet* 1986; 2:921-2.
- Biberfeld G, Bredberg-Rådén U, Böttiger B, et al. Blood donor sera with false-positive western blot reactions to human immunodeficiency virus. *Lancet* 1986; 2:289-90.
- Saag MS, Britz J. Asymptomatic blood donor with a false positive HTLV-III Western blot. *N Engl J Med* 1986; 314:118.
- Roy S, Portnoy J, Wainberg MA. Need for caution in interpretation of Western blot tests for HIV. *JAMA* 1987; 257:1047.
- Sandstrom EG, Schooley RT, Ho DD, et al. Detection of human anti-HTLV-III antibodies by indirect immunofluorescence using fixed cells. *Transfusion* 1985; 25:308-12.
- Stoneburner RL, Chiasson MA, Solomon K, Rosenthal S. Risk factors in military recruits positive for HIV antibody. *N Engl J Med* 1986; 315:1355.
- Shriver K, Klanieki J, Houghton R, Masinovsky R, McClure J, Watson AJ. Analysis of sera exhibiting atypical reactions with HIV. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:16.
- McGrath K, Mijch A, Maskell W, et al. Follow-up of western blot positive blood donors — Victoria, Australia. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:103.
- Thomas D, Mundon FK, Zimmerman D, Larson D, Gowan L, Wilhelm S. Analysis of discrepant anti-HIV ELISA reactivities. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:103.
- Busch M, Shiota J, Nason M, Samson S, Vyas G, Perkins H. Serologic and culture follow-up study of anti-HIV reactive blood donors. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:108.
- Dock NL, Lamberson HV, O'Brien TA, Petteway SR, Alexander S, Poiesz BJ. Evaluation of HIV antibody reactivity in blood donors with atypical western blot patterns. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:148.
- Tan PL, Kay JWD, Munjal D. Studies of normal donor specimens causing various reactivity patterns in sensitive western blot assays. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:151.
- Burke DS, Brandt BL, Redfield RR, et al. Diagnosis of human immunodeficiency virus infection by immunoassay using a molecularly cloned and expressed virus envelope polypeptide: comparison to Western blot on 2707 consecutive serum samples. *Ann Intern Med* 1987; 106:671-6.
- Blakeslee S. Pressure for wider AIDS testing fuels search for better methods. *New York Times*. June 9, 1987:C14.
- Foreman J. Faster, cheaper AIDS test reported. *Boston Globe*. June 5, 1987:1.
- Quinn TC, Francis H, Kline R, et al. Evaluation of a latex agglutination assay using recombinant envelope polypeptides for detection of antibody to HIV. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:166.
- Human immunodeficiency virus infection in transfusion recipients and their family members. *MMWR* 1987; 36:137-40.
- Human T-lymphotropic virus type III/lymphadenopathy-associated virus antibody prevalence in U.S. military recruit applicants. *MMWR* 1986; 35:421-4.
- Schorr JB, Berkowitz A, Cumming PD, Katz AJ, Sandler SG. Prevalence of HTLV-III antibody in American blood donors. *N Engl J Med* 1985; 313:384-5.
- Pauker SG, Kassirer JP. Decision analysis. *N Engl J Med* 1987; 316:250-8.
- Selwyn PA, Schoenbaum EE, Feingold AR, et al. Perinatal transmission of HIV in intravenous drug abusers (IVDAs). Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:157.
- Willoughby A, Mendez H, Minkoff H, et al. Human immune deficiency virus in pregnant women and their offspring. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:158.
- Blanche S, Rouzioux C, Veber F, Le Deist F, Mayaux MJ, Griscelli C. Prospective study on newborns of HIV seropositive women. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:158.
- Braddick M, Kreiss JR, Quinn T, et al. Congenital transmission of HIV in Nairobi, Kenya. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:158.
- Nzilambi N, Ryder RW, Behets F, et al. Perinatal HIV transmission in two African hospitals. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:158.
- Willoughby B, Schechter MT, Boyko WJ, et al. Sexual practices and condom use in a cohort of homosexual men: evidence of differential modification between seropositive and seronegative men. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:5.
- Pollak M, Schiltz MA, Lejeune B. Safer sex and acceptance of testing: results of the nationwide annual survey among French gay men. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:5.
- Pesce A, Negre M, Cassuto JP. Knowledge of HIV contamination modalities and its consequence on seropositive patients behaviour. Presented at the 3rd International Conference on AIDS, Washington, D.C., June 1-5, 1987:60.
- Wofsy CB. Human immunodeficiency virus infection in women. *JAMA* 1987; 257:2074-6.
- Werner LM. Education chief presses AIDS tests. *New York Times*. May 1, 1987:A18.