Polymerase Chain Reaction for the Diagnosis of HIV Infection in Adults

A Meta-Analysis with Recommendations for Clinical Practice and Study Design

Douglas K. Owens, MD, MSc; Mark Holodniy, MD; Alan M. Garber, MD, PhD; John Scott, BA; Seema Sonnad, MS; Lincoln Moses, PhD; Bruce Kinosian, MD; and J. Sanford Schwartz, MD

Purpose: To do a meta-analysis of studies that have evaluated the sensitivity and specificity of polymerase chain reaction (PCR) assay for the diagnosis of human immunodeficiency virus (HIV) infection in adults. Evaluating the performance of PCR is difficult because in certain clinical situations, the sensitivity or specificity of PCR may exceed those of the current reference standard tests (enzyme immunoassay followed by confirmatory Western blot analysis). Therefore, an additional goal was to develop recommendations for 1) the design of future evaluative studies of PCR and 2) the use of PCR in persons with suspected HIV infection.

Data Sources: Studies published between 1988 and 1994 that were identified in a search of 17 computer databases, including MEDLINE, and abstracts identified from conference proceedings.

Study Selection: Studies were included if DNA amplification by PCR was done on peripheral blood mononuclear cells from adults. Ninety-six studies met the inclusion criteria.

Data Extraction: Data were extracted independently by two reviewers. Study design was assessed independently by two investigators blinded to study results.

Results: Reported sensitivities for PCR range from 10% to 100%, and specificities range from 40% to 100%. A summary receiver-operating characteristic curve based on all 96 studies has a maximum joint sensitivity and specificity (upper left point on the curve, where sensitivity equals specificity) of 97.0% to 98.1%. If the threshold value that defines a positive PCR result is chosen so that sensitivity is higher than 98.1%, specificity will decrease to less than 98.1%. Conversely, if the threshold value that defines a positive PCR result is chosen so that specificity is greater than 98.1%, sensitivity will decrease to less than 98.1%. If sensitivity and specificity are chosen to be equal, the corresponding false-positive rate is 1.9% to 3.0%. At the maximum joint sensitivity and specificity, the positive predictive value of PCR ranges from 34% to 85% as the prevalence of HIV increases from 1.0% to 10%. We identified seven areas in which study design could be modified to 1) reduce susceptibility to bias in estimates of the sensitivity and specificity of PCR and 2) to increase the generalizability of the study results. These modifications will also help to overcome methodologic problems created by the lack of a reference standard test.

Conclusions: The PCR assay is not sufficiently accurate to be used for the diagnosis of HIV infection without confirmation. Use of PCR for the diagnosis of HIV in adults should be limited to situations in which antibody tests are known to be insufficient. Future studies of PCR performance should be sufficiently large and should use adequate reference standard tests and standardized methods for the performance of PCR. Specimens should be evaluated by persons blinded to clinical status and to the results of other diagnostic tests for HIV infection.

Polymerase chain reaction (PCR) is a gene amplification technique that has found widespread use in medicine and molecular biology. The PCR assay was developed in 1985 (1, 2), and one of its earliest and most important clinical applications has been the diagnosis of human immunodeficiency virus (HIV) infection (3–9). The PCR assay received attention as a diagnostic test for HIV infection in part because numerous reports suggested that months to years might elapse between infection with HIV and the development of HIV antibodies that could be detected by enzyme immunoassay and Western blot analysis (10, 11). Because PCR directly amplifies proviral HIV DNA and does not depend on HIV antibody formation, it is a potentially attractive alternative to conventional antibody tests. However, the clinical role of PCR in the diagnosis of HIV infection remains uncertain because subsequent studies (12, 13) have not confirmed the occurrence of long "window" periods between infection and the development of antibodies.

Considerable controversy remains about the diagnostic accuracy of PCR. Some studies report that the test has perfect sensitivity and specificity, but others report high false-positive and false-negative rates. An understanding of the diagnostic performance of PCR for HIV infection is essential in determining the appropriate role of PCR in the clinical diagnosis of such infection. However, evaluation of the performance of PCR poses difficult methodologic challenges. To evaluate the sensitivity and specificity of PCR, investigators must ascertain whether study participants are infected with HIV. Typically, a new test is compared with a superior

reference (or gold standard) test, but PCR is an example of a class of diagnostic technologies (including, for example, genetic screening tests) that have the potential to outperform and displace existing tests. At least in certain clinical circumstances, PCR may be more sensitive or more specific than the current reference tests (enzyme immunoassay followed by confirmatory Western blot analysis). The lack of an appropriate reference test substantially complicates evaluation. A successful approach to the evaluation of such technologies would be broadly useful.

We sought to 1) assess the validity and reliability of the scientific evidence on the diagnostic accuracy of PCR; 2) characterize the sensitivity and specificity of PCR on the basis of a formal analysis of the available studies; 3) develop recommendations for the clinical use of PCR in persons with suspected HIV infection; and 4) develop recommendations for the design of future studies of the diagnostic accuracy of PCR. In pursuing our third objective, we paid particular attention to whether PCR technology has improved enough to play a broader clinical role in the diagnosis of HIV infection. We did not evaluate the use of PCR for the quantification of viral load (14) or for the prediction or assessment of response to antiviral therapy (15).

We postulated that more recent studies, because they would reflect advances in PCR technology, would report higher sensitivities and specificities. We also expected that the most methodologically rigorous studies would report lower sensitivities and specificities than other studies and that studies published as full articles would report higher sensitivities and specificities than studies published only as abstracts because of publication bias (the results of which would be that studies reporting high sensitivity and specificity would be published more frequently than studies reporting poor test performance).

Methods

We did a meta-analysis of the published English-language literature to examine the relation between study population, study characteristics, technical aspects of the assay, and measured test performance. We used statistical techniques to fit a summary receiver-operating characteristic (ROC) curve that characterizes the results of multiple studies (16). An ROC curve represents the tradeoff between sensitivity and specificity for a diagnostic test. It can be used to compare diagnostic tests by assessing the degree to which differences in test sensitivity and specificity result from the use of different cut-off points for abnormality rather than from actual dif-

ferences in test performance (17). Typically, an ROC curve is developed from a single study by varying the cut-off point for an abnormal test. In our study, we developed summary ROC curves on the basis of an analysis of multiple studies. Although the method for developing a summary ROC curve differs from the method for developing an ROC curve from a single study, the summary ROC curve also estimates the tradeoff between sensitivity and specificity for a diagnostic test.

Study Identification

An investigator and a professional librarian with extensive experience in medical literature searches independently developed search strategies to identify studies of PCR for the diagnosis of HIV infection that had been published through the middle of 1994 (Appendix). We also manually searched the bibliographies of retrieved articles and conference proceedings. We wrote to the authors of studies that were published only as abstracts and requested information about study design and updated data on PCR performance.

Study Selection

Two investigators independently examined all titles, abstracts, and full articles identified in the search. We included studies if 1) PCR was done on peripheral blood mononuclear cells; 2) DNA (as opposed to RNA) was amplified; 3) study participants were older than 16 years of age; 4) more than 10 participants were enrolled; and 5) primary data sufficient for the determination of both sensitivity and specificity were reported. We excluded studies with fewer than 10 participants because we believed such studies would provide unreliable estimates. We also excluded studies that determined only sensitivity or specificity, because calculation of each is needed to determine a point on the ROC curve. Disagreements were resolved by re-review and discussion.

Data Abstraction

Two investigators independently abstracted data from each study, including the characteristics and risk behaviors of the study sample; the technical details of the assay, including the use of heparin (18); the reference test used (for example, Western blot analysis or viral culture); the criteria used to interpret results of both PCR and the reference test; and the data needed to calculate the sensitivity, specificity, false-positive rate, and false-negative rate of PCR. Disagreements were resolved by re-review and discussion.

Calculation of Sensitivity and Specificity for Polymerase Chain Reaction

We abstracted primary data on the performance of PCR into a 3×3 table in which all participants (or test results) were classified as PCR-positive, PCR-negative, or PCR-indeterminate and as reference test–positive, reference test–negative, or reference test–indeterminate. We used the authors' criteria for PCR-positive, -negative, and -indeterminate test results (for example, the number of primers that had to be detected for a PCR test result to be positive) when they were stated. In the few instances in which these criteria were not stated, we defined a positive test result (in terms of the number of primer pairs detected) to maximize both sensitivity and specificity, if possible, or to maximize sensitivity if doing so did not substantially decrease specificity. We examined whether differences in these criteria affected test performance. We calculated both upper and lower estimates of PCR sensitivity and specificity. We calculated the upper estimate by excluding results that were PCR indeterminate (thereby overestimating sensitivity and specificity), and we calculated the lower estimate by considering reference-test–positive, PCR-indeterminate results to be false-negative results and by considering reference test–negative, PCR-indeterminate results to be false-positive results (thereby underestimating sensitivity and specificity). For the lower estimate, we also considered PCR test results to be false-positive if, after repeated PCR and antibody tests, the results remained PCR-positive and antibody test–negative throughout the follow-up period (8, 19–23). Excluding these few discordant samples did not produce a statistically significant change in our lower-bound estimate.

When possible, PCR performance was evaluated on the basis of the number of study participants rather than the number of tests conducted (some participants were tested more than once). This was done because repeated samples in the same individual person are not independent, and the use of multiple test results from an individual person may therefore spuriously inflate or deflate estimated sensitivity and specificity. Approximately 2% of the samples included in our analysis were repeated samples from individual persons that we could not exclude. Because we calculated sensitivity and specificity by using prospectively defined criteria for the patient's true disease state, the sensitivity and specificity we report for a study sometimes differ from those reported by the original authors. We calculated 95% CIs for individual study estimates of sensitivity and specificity (Figure 1) by using normal or Poisson approximations to the binomial distribution (24), as appropriate (25).

Assessment of Study Design

To assess the reliability of the evidence for the diagnostic accuracy of PCR for HIV infection, two investigators independently assessed the design of the studies by using prospectively developed criteria (Table 1). To develop these criteria, we modified a previously developed assessment framework for diagnostic tests (26–28). Investigators were blinded to the study title, study results, study authors, the name of the journal in which the study results were published, and the name of the institution where the study was done. We assessed the appropriateness of the study design for the evaluation of the diagnostic performance of PCR on a four-point scale (1, 2, 3, or 4). A rating of 1 indicated that the design made the study susceptible to significant bias; a rating of 4 indicated that the study design satisfied all criteria for the evaluation of diagnostic tests (Table 1). Some studies had primary research questions that were not about the accuracy of PCR, but our assessments of the potential sources of bias in the study apply only to the evaluation of the diagnostic performance of PCR. We also identified studies for which the evaluation of PCR performance was either the sole objective or a major goal. We analyzed these studies separately to evaluate whether their design and methods differed from those of studies in which the evaluation of PCR performance was not a primary objective. We accepted positive results on conventional antibody tests (if they included a confirmatory Western blot analysis or similar test) or viral cultures as high-quality evidence of infection. The absence of infection is more difficult to establish. Only studies that used serial testing or follow-up to establish the absence of HIV infection received the highest ratings for study design.

Development of Summary Receiver-Operating Characteristic Curves

The summary ROC curve characterizes the performance of a test as measured in multiple studies (16). Our statistical approach (16, 29–31) for developing summary ROC curves is described in detail in the Appendix. We characterize the summary ROC curve by the point that we call the "maximum joint sensitivity and specificity." This point is defined by the intersection of the ROC curve with a diagonal line that runs from the top left to the bottom right corner of the diagram, along which sensitivity and specificity are equal (specificity is equal to 1 minus the false-positive rate). This point is the maximum attainable common value for sensitivity and specificity for this test; a perfect test would have a joint sensitivity and specificity of 1.0. The maximum joint sensitivity and specificity provides a convenient
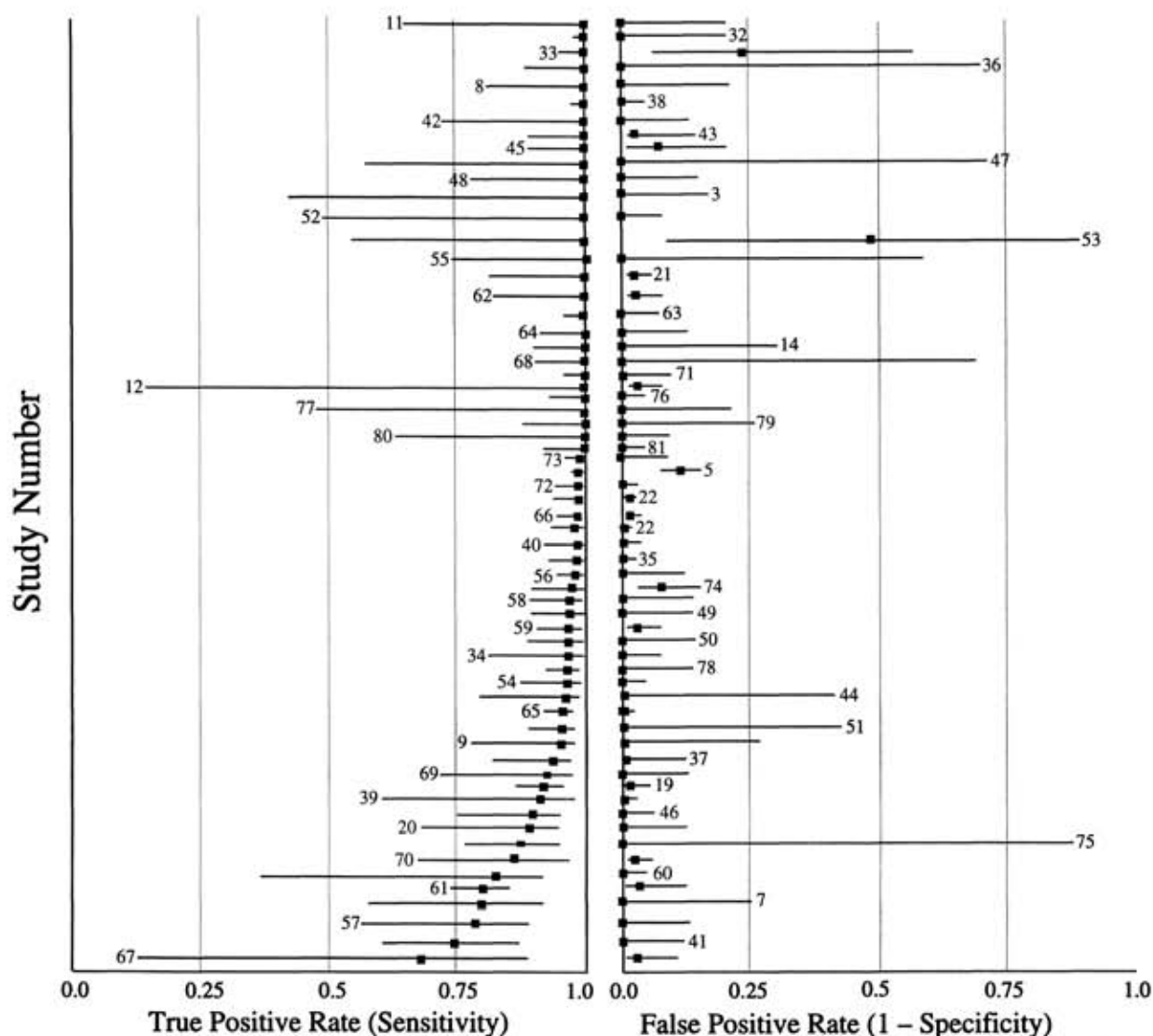
Figure 1. Calculated sensitivity and false-positive rate (1.0 − specificity) for included studies published before 1992. Black squares indicate the sensitivity or false-positive rate; horizontal bars indicate the 95% CIs. A perfect test would have a sensitivity of 1.0 and a false-positive rate of 0.0 (specificity of 1.0). The reference number for each study is shown.

point with which to compare two ROC curves (much like the area under the ROC curve). This point does not indicate the only, or even necessarily the best, combination of sensitivity and specificity for a particular clinical application. Rather, the ROC curve shows the tradeoff between sensitivity and specificity as the threshold for an abnormal PCR test result is changed. The developers of a test can choose a threshold for an abnormal result so that they balance test sensitivity and specificity and appropriately for particular clinical applications. For example, if the developers deem a false-negative result to be more harmful than a false-positive result (as they might for blood-bank screening), they could increase test sensitivity and thereby decrease the number of false-negative results. For each analysis, we report a summary ROC curve based on our upper estimates of sensitivity and specificity (inde-

terminate PCR results excluded) and a summary ROC curve based on our lower-bound estimates of PCR sensitivity and specificity (indeterminate PCR results counted as false-positive or as false-negative results).

Results

Studies Identified

Our literature search identified 5698 titles of potentially relevant articles. After independent review by two readers, 1735 titles were judged to be potentially relevant. We reviewed the associated abstracts and then selected 379 studies published as full articles for further review. Of these 379 articles, 96 met the inclusion criteria and were analyzed (1

Table 1. Criteria for Assessment of Study Design*

Criteria	Minimum Required Score for Overall Study Rating			
PCR test quality†	2	2	1	0
Reference test quality‡	2	2	1	0
Application of reference test§	2	1	1	0
Blinding‖	2	2	1	0
Clinical description¶	1	1	0	–
Cohort assembly**	2	1	1	0
Sample size††	2	1	1	0
Overall study rating	4	3	2	1

* The design of each study was rated according to the individual criteria listed. For each criterion, a study received a score of 2 (to indicate complete satisfaction of the criterion), 1 (to indicate partial satisfaction of the criterion), or 0 (to indicate that the criterion was not satisfied). Columns 2 through 5 show the minimum score required on each criterion for the overall study rating shown in the bottom row. For example, for a study to receive an overall rating of 4, it had to receive a score of at least 1 on clinical description and a score of 2 on all other criteria. PCR = polymerase chain reaction.

† A score of 2 indicates that performance of the PCR assay was described in sufficient detail to enable the method to be reproduced and that the assay included positive and negative controls. A score of 1 indicates incomplete description of the PCR protocol or no mention of controls. A score of 0 indicates that the PCR protocol was not described.

‡ Studies were rated according to the quality of the reference test in both the diseased and the nondiseased participants. For the diseased group, a score of 2 signifies any of the following: enzyme immunoassay with confirmatory Western blot analysis or positive viral culture (in independent samples) or positive antigen testing (only if serial antigen testing was done and concomitant evidence of seroconversion was present). A score of 1 indicates positive viral culture on one sample only. A score of 0 signifies no description of a reference test or positive antigen test result on a single sample only. For the nondiseased population, a score of 2 signifies participants who were negative by enzyme immunoassay with serial testing. A score of 1 indicates a negative result by enzyme immunoassay on one sample only, low-risk group (studies that used blood donors with a single negative result on enzyme immunoassay were rated as 1.5 because of the impracticality of serial testing). A score of 0 signifies no description, a single negative viral culture, or a single negative result on antibody enzyme immunoassay in a high-risk group.

§ A score of 2 indicates that the appropriate reference test was applied consistently within the diseased and nondiseased populations. A score of 1 indicates that all study participants received a reference test but did not consistently receive the same test. A score of 0 indicates that the reference test was not used for all participants.

‖ A score of 2 denotes that the PCR assay and the reference test were done with the investigator blinded to all other test and clinical information. A score of 1 indicates that either PCR or the reference test, but not both, was done with the investigator blinded. A score of 0 signifies no blinding or that blinding was not described.

¶ A score of 2 indicates description sufficient to enable a reader to determine whether the study patients resemble the reader's clinical population, including age, sex, risk factors, and clinical disease. A score of 1 signifies incomplete description. A score of 0 indicates no description.

** A score of 2 indicates that the study population had an adequate spectrum of participants and that assembly of the cohort was described in enough detail that a similar cohort could be assembled by another investigator. A score of 1 signifies an inadequate spectrum of participants or that assembly methods were incompletely described. A score of 0 indicates that the assembly methods were not described or that the results of PCR were used to determine which participants received the reference test (work-up bias).

†† A score of 2 indicates that both the diseased and the nondiseased population had more than 30 participants. A score of 1 signifies that either the diseased or the nondiseased population had more than 30 participants. A score of 0 indicates that both the diseased and the nondiseased population had fewer than 30 participants.

article reported two independent studies that were analyzed individually) (3–5, 7–9, 11, 12, 14, 19–22, 32–113). These studies included 5739 HIV-infected persons and 8929 uninfected persons. We excluded 26 of the 379 studies because they supplied data on either sensitivity or specificity but not both (references available from the authors). Other reasons for exclusion are noted in Table 2. Forty-five studies published only as abstracts met the inclusion criteria and were analyzed separately (references available from the authors).

Assessment of Study Design

The degree to which the 96 included studies satisfied each criterion for the design of an evaluation of a diagnostic test is shown in Figure 2. The information provided in studies whose results were published only in abstract form was insufficient for an assessment of study design. Because our criteria were rigorous, few published studies satisfied all of them. Identifiable aspects of the study design left many studies susceptible to potential bias (for example, lack of blinding during test interpretation) or produced imprecise estimates of the sensitivity and specificity of PCR (for example, small sample size). The numbers of studies receiving a rating of 1, 2, 3, or 4 for overall study design (*see* Methods and Table 1) were 73, 12, 6, and 5, respectively. Studies that focused solely or largely on the evaluation of PCR performance did not receive more favorable ratings than other studies. The criteria that were satisfied least often were adequacy of blinding during the interpretation of test results, adequacy of the reference test in uninfected study participants, and adequacy of sample size. In 57% of studies, there were fewer than 30 reference test–positive or reference test–negative participants; this resulted in wide 95% CIs on the estimates of sensitivity and specificity (Figure 1).

Most of the studies (74%) used acceptable reference tests in the HIV-infected participants. The clinical population of greatest interest for PCR testing, however, is that of persons at high risk for infection who have negative results on conventional antibody tests. Twenty-two of the 96 studies (23%)
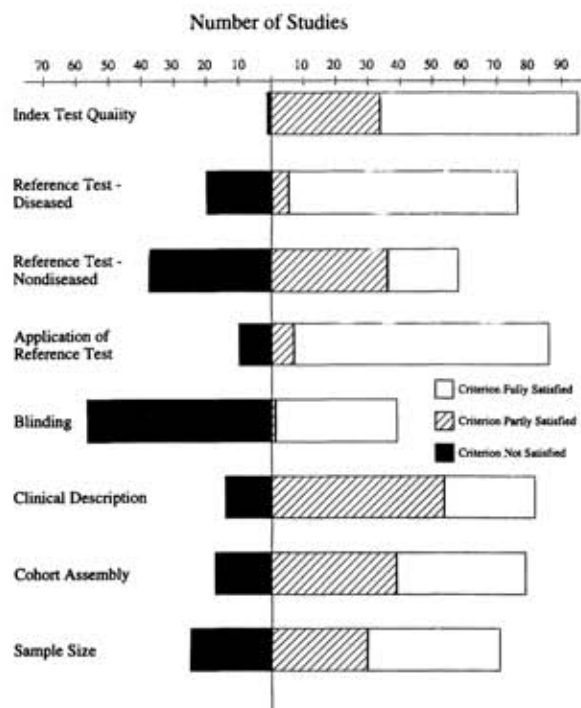

Figure 2. Results of quality ratings for individual quality criteria. The number of studies that satisfied, partly satisfied, or failed to satisfy each criterion is shown. For an explanation of the scoring system for each criterion, *see* Table 1.

Table 2. Results of Literature Search

Classification	Studies, n
Potentially eligible studies	379
Excluded studies (total)	283
Inadequate data to calculate sensitivity and specificity	72
PCR not done in PBMC*	30
Description of technical aspect of assay methods only	36
Pediatric sample	47
Total sample size <10 participants	22
Report not written in English	15
Other	61
Studies analyzed	96

* PBMC = peripheral blood mononuclear cells; PCR = polymerase chain reaction.

fully met the criterion for an adequate reference test in these persons. Thirty-six studies (38%) partially fulfilled this criterion, and 38 studies (40%) did not satisfy this criterion. Twenty-two studies (23%) fully met the reference test criteria for participants with and those without disease.

Sensitivity and Specificity of Polymerase Chain Reaction

Measured performance was extremely variable. When indeterminate PCR results were excluded, sensitivity ranged from 10% to 100% and specificity ranged from 40% to 100% (data available from the authors). In studies in which the design was rated as either 3 or 4, sensitivity ranged from 83% to 100%, and specificity ranged from 95% to 100%.

Summary Receiver-Operating Characteristic Curves

On the basis of all 96 studies, the upper estimate (indeterminate PCR results excluded) of the maximum joint sensitivity and specificity was 98.1%, and the lower estimate (indeterminate PCR results counted as false-positive or false-negative results) was 97.0% (Figure 3, Table 3). The corresponding log odds ratios (± SE) are 7.93 ± 0.330 and 6.96 ± 0.195, respectively (Table 3). The exclusion of four studies that reported sensitivity and specificity on the basis of the number of samples rather than the number of study participants did not significantly affect our results ($P > 0.2$). The exclusion of studies that used heparin to preserve blood samples provided slightly but not statistically significantly higher estimates of joint sensitivity and specificity (upper estimate, 98.5 [$P > 0.2$]; lower estimate, 97.1 [$P > 0.2$]). Figure 3 shows the tradeoff between sensitivity and specificity. For example, if a cut-off point for an abnormal PCR result is chosen so that the specificity of PCR is 99.0% (false-positive rate, 1.0%), the sensitivity decreases to approximately 91.0% to 96.0%.

Subgroup Analysis

Our subgroup analyses (Table 3) indicated that studies published only as abstracts reported lower values for sensitivity and specificity than did studies published as articles. The upper estimate of maximum joint sensitivity and specificity based on studies that received scores of 2, 3, or 4 did not differ significantly from the estimated performance based on studies that received a score of 1. However, the lower-bound estimate of joint sensitivity and specificity was significantly lower in studies with better study design scores (96.2 compared with 97.7 [$P = 0.02$]; *see* Table 3). In addition, rather than finding that the reported accuracy of PCR was greater in more recent studies, we found that studies published during or after 1991 gave lower estimates of the accuracy than did studies published before 1991. We analyzed studies reported in and after 1991 because we believed that PCR technology had matured by 1991. Finally, the upper estimate of sensitivity and specificity based on studies in which the primary purpose was to evaluate the accuracy of PCR did not differ from estimates based on other studies. Subgroups defined by reference test criteria and by study objective showed significant differences only as judged by lower-bound estimates of joint sensitivity and specificity (Table 3).

The criteria for determining when PCR gave a positive result varied among the studies. Two of the 23 studies in which study design was rated as 2, 3, or 4 considered a PCR test result to be positive if reactivity with any one primer pair was seen. Thirteen studies required reactivity with two primer pairs, 7 did not specify explicit criteria, and 1 used variable criteria depending on the PCR assay used.
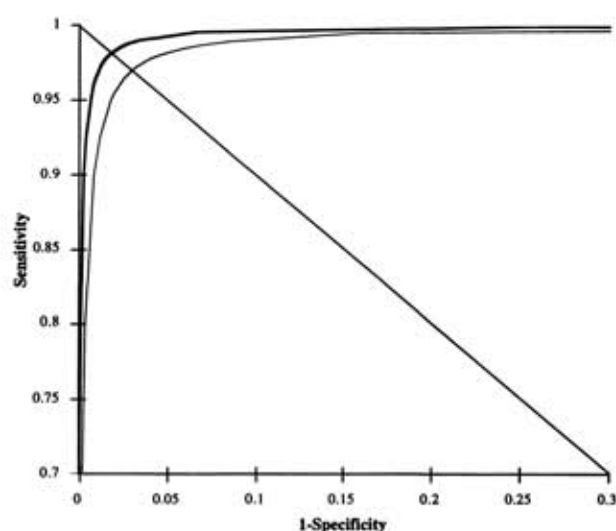

Figure 3. Summary receiver-operating characteristic curve for polymerase chain reaction (PCR). The upper left corner of the summary receiver-operating characteristic (ROC) curve is shown. The summary ROC curve is based on all 96 included studies. The lower estimate (*thin line*) was calculated by including indeterminate PCR test results to determine a conservative estimate for sensitivity and specificity. The upper estimate (*thick line*) was calculated by excluding indeterminate PCR test results. The intersection of the diagonal line with each curve represents the maximum joint sensitivity and specificity for that ROC curve, where sensitivity equals specificity.

Table 3. Subgroup Comparisons: Estimated Maximum Joint Sensitivity and Specificity*

Subgroup	Studies, n	Mantel–Haenszel Log Odds Ratio ± SE†		Two-Sided P Value		Maximum Joint Sensitivity and Specificity, %‡	
		Lower Estimate	Upper Estimate	Lower Estimate	Lower Estimate	Lower Estimate	Upper Estimate
All studies	96	6.96 ± 0.195	7.93 ± 0.330			97.0	98.1
Publication date							
Before 1991	37	7.83 ± 0.396	9.24 ± 0.528	0.02	0.02	98.0	99.0
During or after 1991	59	6.73 ± 0.216	7.66 ± 0.356			96.7	97.9
Sexual preference§							
Homosexual men	24	6.49 ± 0.249	8.04 ± 0.409	>0.2	>0.2	96.2	98.2
Other	35	7.01 ± 0.347	7.55 ± 0.525			97.1	97.8
Sample size							
≤140 persons	65	7.02 ± 0.454	7.28 ± 0.541	>0.2	0.08	97.1	97.4
>140 persons	31	6.93 ± 0.207	8.41 ± 0.321			97.0	98.5
Study design rating‖							
1	73	7.51 ± 0.358	7.82 ± 0.439	0.02	>0.2	97.7	98.0
2, 3, or 4	23	6.45 ± 0.229	8.15 ± 0.422			96.2	98.3
Reference test and PCR criteria¶							
Partly satisfied	78	7.71 ± 0.352	8.03 ± 0.445	0.001	>0.2	97.9	98.2
Fully satisfied	18	6.01 ± 0.243	7.70 ± 0.421			95.3	97.9
Study objective							
To evaluate test performance**	51	7.35 ± 0.280	7.88 ± 0.417	0.02	>0.2	97.5	98.1
Other	45	6.43 ± 0.273	8.06 ± 0.484			96.1	98.2
Publication status††							
Abstracts	30	5.60 ± 0.298	5.60 ± 0.298	0.001	<0.001	94.3	94.3
Published articles	63	6.86 ± 0.219	8.43 ± 0.357			96.9	98.5

* PCR = polymerase chain reaction.
† The log odds ratio measures the discriminatory power of a test. A higher ratio corresponds to higher sensitivity and specificity.
‡ The maximum joint sensitivity and specificity represents the upper left corner of the receiver-operating characteristic curve, where sensitivity equals specificity.
§ Fifty-eight studies gave sufficient information to determine risk group.
‖ Overall rating of the study design. A study with a rating of 1 was subject to major biases from design flaws. A study that satisfied all criteria received a rating of 4 (see Table 1).
¶ Refers to the criteria for the performance of PCR and the use of reference test in the diseased and nondiseased participants (see Table 1). "Satisfied" indicates that the studies fully satisfied criteria for the use of PCR and use of the reference test. "Partly satisfied" indicates that studies received a 0 or 1 rating for use of either PCR or the reference tests.
** Studies whose primary or major objective was to assess the sensitivity and specificity of PCR for the diagnosis of human immunodeficiency virus infection.
†† Compares only studies and abstracts published before 1992.

A summary ROC curve based on the 13 studies that required reactivity with two primer pairs yielded an upper estimate of the joint combined sensitivity and specificity of 98.0%.

Post-Test Probability

Post-test probabilities depend on the sensitivity and specificity of PCR. Figure 4 shows the post-test probability of disease after positive and negative PCR test results as calculated using Bayes theorem (17) if the threshold for an abnormal test result has been chosen so that the test has maximum joint sensitivity and specificity. For example, if the pretest probability of HIV infection is 10%, the post-test probability of disease after a positive PCR test result (positive predictive value) increases to between 78% (thin curve, Figure 4) and 85% (thick curve, Figure 4). At a pretest probability of 1.0% and a sensitivity and specificity of 98.1% (the upper estimate), the post-test probability of HIV infection after a positive PCR test result is only 34%.

Discussion

We sought to critically and systematically examine the many published studies that have reported on the use of PCR for the diagnosis of HIV infec-

tion in adults. If it is sufficiently accurate and inexpensive, PCR could supplant standard antibody tests for diagnosis and screening. Our investigation produced two main findings. First, the false-positive
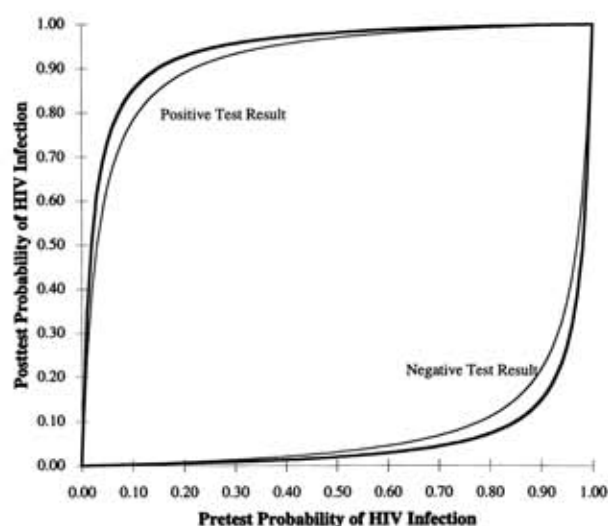
Figure 4. Post-test probability of human immunodeficiency virus (HIV) infection. Upper curves show the post-test probability of HIV infection after a positive t polymerase chain reaction (PCR test result). Lower curves show the post-test probability of HIV infection after a negative PCR test result. It is assumed that PCR has a joint sensitivity and specificity between 97.0 (thin curves) and 98.1 (thick curves), consistent with the upper- and lower-estimate summary receiver-operating characteristic curve based on all 96 included studies.

and false-negative rates of PCR that we determined are too high to warrant a broader role for PCR in either routine screening or in the confirmation of diagnosis of HIV infection. This conclusion is true even for the results reported from more recent, high-quality studies that used commercially available, standardized PCR assays. We did not address the emerging potential uses of PCR for use in quantification of viral load (14) or in the prediction or assessment of response to antiviral therapy (15), areas in which PCR may prove to have an important clinical role. Second, our evaluation of study design suggests several modifications of design that would substantially reduce susceptibility to bias.

We estimated the maximum joint sensitivity and specificity of PCR to range from 97.0% to 98.1%, with corresponding false-positive and false-negative rates between 1.9% and 3.0%. Our analysis of the post-test probability of disease (Figure 4) indicates that if we use the joint maximum sensitivity and specificity for PCR, the proportion of false-positive tests would be unacceptably high for screening or other common clinical applications. The post-test probability of disease will vary depending on the sensitivity and specificity, which in turn depend on the cut-off point used to define an abnormal test result. The summary ROC curve indicates how specificity will decrease (or increase) as sensitivity increases (or decreases).

To put the diagnostic performance of PCR in context, the conventional antibody test sequence of an enzyme immunoassay followed by confirmatory Western blot analysis has a sensitivity that exceeds 99% and a specificity greater than 99.5% (corresponding to a false-positive rate $\leq$ 0.5%) in high-quality screening programs (114–116). Although the meta-analytic techniques that we used have not been applied to HIV antibody tests, a study including 1400 participating laboratories, done by the Centers for Disease Control and Prevention (CDC), found the sensitivity and specificity of the enzyme immunoassay to be 99.68% and 98.46%, respectively, in 1988 (6566 infected samples, 3051 negative samples) and 99.3% and 99.7%, respectively, in 1989 (115). False-positive rates as low as 6 per million have been reported in blood-bank screening programs (116), although such low rates may not be attainable in all programs. The log odds ratio associated with the 1988 findings of the CDC study is 9.90 ± 0.26, which substantially exceeds the log odds ratio we found for PCR (7.93 ± 0.33); the sensitivity and specificity of the enzyme immunoassay in 1989 were even higher. The studies included in our analysis suggested that the sensitivity of the p24 antigen assay (in contrast to that of antibody tests) is inferior to that of other tests. For example, p24 antigen was detected in only 14% of HIV-infected hemo-philiacs (3) and in only 8% to 32% of participants with PCR-positive, antibody test–positive test results (50, 63). Although these studies suggest that PCR is superior to the p24 antigen assay, we cannot directly compare the sensitivities and specificities produced by the two assays, because the p24 antigen assays has not been evaluated formally with summary ROC curves.

Our subgroup analyses show that studies published only as abstracts provided lower estimates of the sensitivity and specificity of PCR. This may indicate publication bias (the preference for publishing favorable rather than unfavorable studies). Although publication bias is a concern in meta-analyses, few examples of it have been documented. Studies with more rigorous designs provided similar upper estimates of joint sensitivity and specificity but decreased lower estimates of joint sensitivity and specificity relative to other studies. Rigorous study design (for example, blinding) may prevent the inadvertent overestimation of test performance.

We did not find evidence that the performance of PCR improved over time. The problem of false-positive and false-negative PCR results for the diagnosis of HIV infection has led to efforts to develop quality assurance programs for the performance of PCR (90, 95, 100). For example, laboratory personnel now take extensive precautions to prevent carryover contamination, which was an important cause of false-positive test results in early studies. A particularly rigorous program of quality assurance was instituted recently by the AIDS (acquired immunodeficiency syndrome) Clinical Trials Group investigators. Because training of laboratory personnel is probably an important component of laboratory performance, the study done by these investigators (95) used only experienced laboratories that met strict performance criteria. The sensitivity and specificity of PCR were found to be 97.4% and 94.8%, respectively, in an ongoing quality assurance program that used the latest generation of commercially available PCR kits (95) and standardized protocols for the performance of PCR. These results are consistent with the results of our analysis and, along with the findings of another multicenter quality assurance study (90), indicate that the problem of false-positive and false-negative results persists in currently available test programs, including those that use commercially available standardized PCR tests rather than assays developed in-house.

Recommendations for the Clinical Role of Polymerase Chain Reaction

Our analysis confirms that at present, PCR is not sufficiently accurate to be a reference or gold standard test. The frequency of false-positive and false-negative results, even in more recent studies, pre-

cludes this. Clearly, the performance of PCR is not adequate to justify its use as a clinical screening test. The PCR assay will be most useful in settings in which conventional antibody tests are indeterminate or are likely to be inaccurate. Depending on the criteria used, 13% to 48% of Western blot analyses in low-risk persons who have repeated reactive enzyme immunoassay results may be indeterminate (48). In these situations, PCR is a useful alternative test.

The PCR assay may also be useful in persons who have recently had a known or suspected exposure to HIV whose infection status must be determined urgently (for example, health care workers who have sustained a percutaneous exposure to HIV-infected blood). Although the PCR assay provides interim information that may be useful in selected cases, clinicians and health care workers should be aware that the false-positive rate of PCR probably exceeds that of conventional antibody tests. Therefore, the benefit of early detection should be weighed against the increased risk for a false-positive result. Conventional antibody tests and clinical follow-up can minimize the effect of false-negative or false-positive PCR test results. We conclude that for the diagnosis of HIV infection in adults, the role of PCR should continue to be limited to circumstances in which antibody tests are known to be insufficient or indeterminate.

Recommendations for Study Design

Our analysis highlights the importance of a crucial aspect of study design: the choice, use, and description of the index test (PCR) and reference tests. Whenever possible, studies of the performance of a diagnostic test should use reference tests that unequivocally establish the true state of disease or health. Because PCR can detect HIV infection before antibodies have developed, a positive PCR test result in a person with negative results on an HIV enzyme immunoassay could represent either a false-positive PCR result or a false-negative enzyme immunoassay result. Evaluation of PCR is challenging because no single diagnostic test can resolve this dilemma with certainty. For current studies of HIV infection, the discrepancy can be resolved by serially testing seronegative persons with enzyme immunoassay and Western blot analysis and doing clinical follow-up for a period long enough to exclude acute infection. If a person is truly infected with HIV, then eventually peripheral blood mononuclear cell culture or plasma culture should become positive, the enzyme immunoassay and Western blot analysis should become reactive, or clinical illness should ensue. Although some reports indicate that the period between infection and antibody production may last as long as 4 years, more than 95% of HIV-infected persons seroconvert within 9 to 12 months

(117). Studies of other diagnostic tests for HIV have successfully used serial testing and clinical follow-up to determine true infection status (118). In high-risk populations, however, the value of long-term serial testing may be attenuated by incident infections. In many of the studies that we reviewed, longer follow-up would have enabled the investigators to convincingly establish the disease status of antibody test–negative participants.

Once the procedure for determining the infection status has been chosen, it should be applied consistently to all the study participants, regardless of their PCR test results. A particular PCR test result should not be used to decide which persons are given the reference test, because such a selection procedure can create "referral bias." Referral bias spuriously reduces the number of true-negative and false-negative PCR test results in the study population and thereby overestimates sensitivity and underestimates specificity (17). Investigators can further avoid potential bias in the interpretation of test results by doing the PCR assays and the reference tests while blinded to the results of other tests for HIV and to all clinical information. Investigators were blinded to previous test results in only 40% of the 96 studies that we evaluated. In addition, interpretation of studies can be enhanced if both the PCR assays and the reference tests are described in sufficient detail to allow another investigator to reproduce the test procedures. Descriptions should address how the tests were done and how the results were interpreted.

Many of the studies we analyzed had design limitations that are commonly found in studies of other types of diagnostic tests: incomplete representation of the spectrum of patients in the study population, insufficient sample size, and incomplete reporting of test results. To increase the generalizability of study results, the study sample should reflect the entire spectrum of disease encountered in the clinical population of interest (119). For example, the nondiseased population should include persons who are at risk for HIV infection and would be candidates for testing rather than healthy controls. The usefulness of the study will be enhanced if the study sample is described in enough detail to 1) enable readers to determine whether the sample is sufficiently similar to their clinical setting to permit application of the study findings and 2) allow another investigator to assemble a cohort similar to the sample to confirm the study findings (120). Investigators can reduce uncertainty to acceptable levels in the estimates of sensitivity and specificity by increasing the sample size. As shown in Figure 1, the 95% CIs for sensitivity and specificity are broad if the sample size is small. Recommendations for determining appropriate sample sizes have been published (121).

Finally, studies of test performance can be improved if investigators report the sensitivity and specificity of a test for various definitions of test reactivity (114). Because both sensitivity and specificity are determined by the choice of the threshold for an abnormal test result, there is an inherent tradeoff between them. The threshold for a reactive test can be chosen so that PCR is 100% sensitive or 100% specific, but usually not both (unless the test is perfect and the diseased and nondiseased populations have no overlap for the attribute being measured). Thus, a study that evaluates only the sensitivity of PCR (that is, that includes only diseased persons) or the specificity of PCR (that is, that includes only nondiseased persons) provides insufficient information for an evaluation of test performance. Investigators can develop an ROC curve by calculating sensitivity and specificity for varying definitions of test reactivity (122). The ROC curve represents the performance of a test much more thoroughly than do single values of sensitivity and specificity, in which differences in test performance may merely indicate that different criteria for test positivity were used. Such reporting also facilitates the development of summary ROC curves, such as those used in our meta-analysis and used by others in the analysis of other diagnostic tests (123, 124).

Technical advances will probably improve the performance of the PCR assay. As the sensitivity and specificity of PCR for the diagnosis of HIV improve, the clinical role of PCR may change. Such a change should occur only after a rigorous evaluation of test performance that incorporates the recommendations for study design discussed above. Currently, interpretation of PCR test results for the diagnosis of HIV infection should be combined with careful consideration of the clinical circumstances and with the use of confirmatory tests and clinical follow-up whenever possible.

Appendix

In this Appendix, we describe the methods we used to search the literature and develop summary ROC curves.

Literature Search

Two literature searches were done by a professional research librarian to identify pertinent published data. For articles published in or before 1991, 17 databases were searched: MEDLINE, AIDSline, Cancerlit, Embase, Federal Research in Progress, Compendex, Scisearch, Inspec, Conference Papers, Diogenes, Chemical Abstracts, Biosis, Life Sciences Collection, Biobusiness, Pharmaceutical News Index, National Technical Information Service, and International Pharmaceutical Abstracts. For articles published in 1992 through 1994, we limited our computer-based search to MEDLINE because we found other databases to be redundant. In the initial search, we used the following strategy.

1. S1 Acquired (W) Immunodeficiency OR Acquired (W) Immune (W) Deficiency OR AIDS
2. S2 HIV OR HIV1 OR HIV2 OR HIV-1 OR HIV-2
3. S3 Human (W) (Immunodeficiency OR Immune (W) Deficiency) (W) (Virus OR Viruses)
4. S4 HTLV3 OR HTLVIII OR HTLV (5W) (3 OR III)
5. S5 Human (W) T (W) Cell (W) (Leukaemia OR Leukemia) (W) (Virus OR Viruses) (5W) (3 OR III)
6. S6 LAV OR Lymphadenopathy (W) Associated (W) (Virus OR Viruses)
7. S7 ARC
8. S8 PCR OR Polymerase (W) Chain (W) Reaction
9. S9 PCR OR Polymerase (W) Chain
10. S10 Amplif? (3N) (Gene OR Genes OR Genetic OR DNA OR Deoxyribonucleic)
11. S11 Sequence (W) Tagged (W) Site?
12. S12 (S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7) AND (S9 OR S10 OR S11)
13. S13 Remove Duplicates S12

This search was updated with a slightly different strategy.
1. S1 Acquired (W) Immunodeficien? OR Acquired (W) Immune (W) Deficien? OR AIDS
2. S2 HIV OR Human (W) Immunodeficien? (W) Virus? OR Human (W) Immune (W) Deficien? (w) Virus? OR HIV-1 OR HIV-2
3. S3 DC = D24.611.216.327.570.470.?
4. S4 ARC
5. S5 Polymerase (W) Chain OR PCR
6. S6 (Gene OR Genetic OR DNA OR Sequence? OR Deoxyribonucleic OR Nucleic OR Nucleotide? OR Genome?) (5N) Amplif?
7. S7 Amplicon OR Amplicons
8. S8 Sequence (W) Tagged (W) Site?
9. S9 (S1 OR S2 OR S3 OR S4) AND (S5 OR S6 OR S7 OR S8)
10. Limit S9 to Updates since the earlier search
11. Eliminate Duplicates

When the Chemical Abstracts database was searched, the following strategy was used.
1. L1 Acquired (W) Immunodeficien? OR Acquired (W) Immune (W) Deficien?
2. L2 AIDS OR HIV OR Human (W) Immunodeficien? (W) Virus?
3. L3 Human (W) Immune (W) Deficien? (W) Virus?
4. L4 HIV-1 OR HIV-2 OR ARC
5. L5 Polymerase (W) Chain OR PCR
6. L6 (Gene OR Genetic OR DNA OR Sequence? OR Deoxyribonucleic) (3A) Amplif?
7. L7 S Amplicon OR Amplicons
8. L8 Amplicon OR Amplicons OR Sequence (W) Tagged (W) Site?
9. L9 (L1 OR L2 OR L3 OR L4) AND (L5 OR L6 OR L8)

Summary Receiver-Operating Characteristic Curves

We used two approaches for estimating summary ROC curves. The first method, described previously (16), uses a logistic transformation of sensitivity and specificity so that a summary ROC curve can be fitted with linear regression. To do the logistic transformation, we added a correction factor of 0.5 when the data for a study included zero values (which occurred when either the number of false-positive tests or the number of false-negative tests was zero). The ROC curve was then determined by back transformation of the fitted linear regression line. The method also provides a statistical test to evaluate whether the ROC curve is symmetrical. If the summary ROC curve is symmetrical, a common log odds ratio uniquely determines the entire ROC curve. The test of symmetry is to determine whether the slope of the fitted regression line differs significantly from zero. Regression lines with a slope near zero can be represented by a common log odds ratio; if the slope differs from zero, the odds ratio changes for different points on the ROC curve. Our analysis indicated that the slope for both our upper estimate (slope [$\pm$ SE] = -0.156 ± 0.118 [95% CI, -0.369 to 0.078]; $P = 0.10$)

and our lower estimate (slope = -0.174 ± 0.114 [CI, -0.40 to 0.05]; $P = 0.13$) of the summary ROC curves did not differ significantly from zero. We therefore felt justified in estimating a common odds ratio, and we used the Mantel–Haenszel estimator (29).

We also chose the Mantel–Haenszel method because the alternative method uses a logistic transformation that requires a correction factor for zero values. The correction factor can introduce bias in the estimation of summary ROC curves for highly accurate tests such as PCR. We calculated the SE of the estimated log odds ratio using both the method of Robins and coworkers (31) and the jackknife and bootstrap methods (30). Reported comparison statistics are based on the SE as calculated by using the method of Robins and coworkers because this method produced the most conservative estimates of statistical significance (that is, the largest SEs). To determine whether the sensitivity and specificity of PCR differed among certain subgroups, we compared the Mantel–Haenszel estimated common log odds ratio for each group in terms of their SEs. We compared both the upper and lower estimate of sensitivity and specificity in the subgroups.

Requests for Reprints: Douglas K. Owens, MD, MSc, Section of General Internal Medicine (111A), Veterans Affairs Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304.

Current Author Addresses: Drs. Owens and Garber: Veterans Affairs Palo Alto Health Care System, 3801 Miranda Avenue (111A), Palo Alto, CA 94304.
Dr. Holodniy: Veterans Affairs Palo Alto Health Care System, 3801 Miranda Avenue (111-ID), Palo Alto, CA 94304.
Mr. Scott: 1 Cloister Court, Apartment 205, Bethesda, MD 20814-1460.
Ms. Sonnad: University of Michigan, SPHII, Department of Health Management and Policy, 109 Observatory Road, Ann Arbor, MI 48109-2029.
Dr. Moses: Department of Health Research and Policy, Stanford University, Redwood Building, Room T-160, Stanford, CA 94305-5092.
Dr. Kinosian: Veterans Affairs Medical Center, Hospital Based Home Care Program (111F), University and Woodland Avenues, Philadelphia, PA 19104.
Dr. Schwartz: Leonard Davis Institute of Health Economics, 3641 Locust Walk, Room 209, Philadelphia, PA 19104.

References

1. **Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H.** Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol. 1986;51:263-73.
2. **Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al.** Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 1985;230:1350-4.
3. **Jackson JB, Sannerud KJ, Hopsicker JS, Kwok SY, Edson JR, Balfour HH Jr.** Hemophiliacs with HIV antibody are actively infected. JAMA. 1988; 260:2236-9.
4. **Farzadegan H, Vlahov D, Solomon L, Munoz A, Astemborski J, Taylor E, et al.** Detection of human immunodeficiency virus type 1 infection by polymerase chain reaction in a cohort of seronegative intravenous drug users. J Infect Dis. 1993;168:327-31.
5. **Sheppard HW, Ascher MS, Busch MP, Sohmer PR, Stanley M, Luce MC, et al.** A multicenter proficiency trial of gene amplification (PCR) for the detection of HIV-1. J Acquir Immune Defic Syndr. 1991;4:277-83.
6. **Sheppard HW, Dondero D, Arnon J, Winkelstein W Jr.** An evaluation of the polymerase chain reaction in HIV-1 seronegative men. J Acquir Immune Defic Syndr. 1991;4:819-23.
7. **Ou CY, Kwok S, Mitchell SW, Mack DH, Sninsky JJ, Krebs JW, et al.** DNA amplification for direct detection of HIV-1 in DNA of peripheral blood mononuclear cells. Science. 1988;239:295-7.
8. **Loche M, Mach B.** Identification of HIV-infected seronegative individuals by a direct diagnostic test based on hybridisation to amplified viral DNA. Lancet. 1988;2:418-21.
9. **Hart C, Schochetman G, Spira T, Lifson A, Moore J, Galphin J, et al.** Direct detection of HIV RNA expression in seropositive subjects. Lancet. 1988;2:596-9.
10. **Imagawa DT, Lee MH, Wolinsky SM, Sano K, Morales F, Kwok S, et al.** Human immunodeficiency virus type 1 infection in homosexual men who remain seronegative for prolonged periods. N Engl J Med. 1989;320:1458-62.
11. **Wolinsky SM, Rinaldo CR, Kwok S, Sninsky JJ, Gupta P, Imagawa D, et al.** Human immunodeficiency virus type 1 (HIV-1) infection a median of 18 months before a diagnostic western blot. Evidence from a cohort of homosexual men. Ann Intern Med. 1989;111:961-72.
12. **Lee TH, el-Amad Z, Reis M, Adams M, Donegan EA, O'Brien TR, et al.** Absence of HIV-1 DNA in high-risk seronegative individuals using high-input polymerase chain reaction. AIDS. 1991;5:1201-7.
13. **Busch MP, Eble BE, Khayam-Bashi H, Heilbron D, Murphy EL, Kwok S, et al.** Evaluation of screened blood donations for human immunodeficiency virus type 1 infection by culture and DNA amplification of pooled cells. N Engl J Med. 1991;325:1-5.
14. **Lee TH, Sunzeri FJ, Tobler LH, Williams BG, Busch MP.** Quantitative assessment of HIV-1 DNA load by coamplification of HIV-1 gag and HLA-DQ-alpha genes. AIDS. 1991;5:683-91.
15. **Holodniy M, Katzenstein D, Winters M, Montoya J, Shafer R, Kozal M, et al.** Measurement of HIV virus load and genotypic resistance by gene amplification in asymptomatic subjects treated with combination therapy. J Acquir Immune Defic Syndr. 1993;6:366-9.
16. **Moses LE, Shapiro D, Littenberg B.** Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med. 1993;12:1293-316.
17. **Owens D, Sox HJ.** Medical decision making: probabilistic medical reasoning. In: Shortliffe EH, Perreault L, Fagan LE, Wiederhold G, Fagan LM. Medical Informatics: Computer Applications in Health Care. Reading, MA: Addison-Wesley; 1990:70-116.
18. **Holodniy M, Kim S, Katzenstein D, Konrad M, Groves E, Merigan T.** Inhibition of human immunodeficiency virus gene amplification by heparin. J Clin Microbiol. 1991;29:676-9.
19. **Horsburgh CR Jr, Ou CY, Jason J, Holmberg SD, Lifson AR, Moore JL, et al.** Concordance of polymerase chain reaction with human immunodeficiency virus antibody detection. J Infect Dis. 1990;162:542-5.
20. **Ensoli F, Fiorelli V, Mezzaroma I, D'Offizi GP, Aiuti F.** Proviral sequences detection of human immunodeficiency virus by polymerase chain reaction in seronegative subjects by polymerase chain reaction. Mol Cell Probes. 1990;4:153-61.
21. **Hewlett IK, Laurian Y, Epstein J, Hawthorne CA, Ruta M, Allain JP.** Assessment by gene amplification and serological markers of transmission of HIV-1 from hemophiliacs to their sexual partners and secondarily to their children. J Acquir Immune Defic Syndr. 1990;3:714-20.
22. **Young KK, Peter JB, Winters RE.** Detection of HIV DNA in peripheral blood by the polymerase chain reaction: a study of clinical applicability and performance. AIDS. 1990;4:389-91.
23. **Ensoli F, Fiorelli V, Mezzaroma I, et al.** Plasma viremia in prolonged seronegative human immunodeficiency virus type-1 infection [Abstract]. In: Istituto Superiore Di Sanita: Science Challenging AIDS. Seventh International Conference on AIDS, Florence 16–21 June 1991. Rome: The International Conference on AIDS; 1991:302.
24. **Fleiss J.** Statistical Methods for Rates and Proportions. 2d ed. New York: J Wiley; 1981.
25. **Moses L.** Think and Explain with Statistics. Reading, MA: Addison-Wesley; 1986.
26. **Hoffman RM, Kent DL, Deyo RA.** Diagnostic accuracy and clinical utility of thermography for lumbar radiculopathy. A meta-analysis. Spine. 1991;16:623-8.
27. **Kent DL, Haynor DR, Larson EB, Deyo RA.** Diagnosis of lumbar spinal stenosis in adults: a metaanalysis of the accuracy of CT, MR, and myelography. AJR Am J Roentgenol. 1992;158:1135-44.
28. **Kent DL, Larson EB.** Disease, level of impact, and quality of research methods. Three dimensions of clinical efficacy assessment applied to magnetic resonance imaging. Invest Radiol. 1992;27:245-54.
29. **Kelsey JL, Thompson WD, Evans AS.** Methods in Observational Epidemiology. New York: Oxford Univ Pr; 1986.
30. **Efron B.** The Jackknife, the Bootstrap and Other Resampling Plans. 5th ed. Philadelphia: Society for Industrial and Applied Mathematics; 1982.

31. Robins J, Breslow N, Greenland S. Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. Biometrics. 1986;42:311-23.

32. Wages JM Jr, Hamdallah M, Calabro MA, Fowler AK, Oster CN, Redfield RR, et al. Clinical performance of a polymerase chain reaction testing algorithm for diagnosis of HIV-1 infection in peripheral blood mononuclear cells. J Med Virol. 1991;33:58-63.

33. Sonnerborg A, Abens J, Johansson B, Strannegard O. Detection of human immunodeficiency virus-1 by polymerase chain reaction and virus cultivation. J Med Virol. 1990;31:234-40.

34. Simmonds P, Balfe P, Peutherer JF, Ludlam CA, Bishop JO, Brown AJ. Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. J Virol. 1990;64:864-72.

35. Mariotti M, Lefrere JJ, Noel B, Ferrer-Le-Coeur F, Vittecoq D, Girot R, et al. DNA amplification of HIV-1 in seropositive individuals and in seronegative at-risk individuals. AIDS. 1990;4:633-7.

36. Lucotte G, Reveilleau J. Identification of HIV-1 infected seropositive subjects by a direct diagnostic test involving hybridization of amplified viral DNA. Mol Cell Probes. 1989;3:299-306.

37. Li JJ, Friedman-Kien AE, Huang YQ, Mirabile M, Cao YZ. HIV-1 DNA proviral sequences in fresh urine pellets from HIV-1 seropositive persons [Letter]. Lancet. 1990;335:1590-1.

38. Lefrere JJ, Mariotti M, Courouce AM, Rouger P, Salmon C, Vittecoq D. Polymerase chain reaction testing of HIV-1 seronegative at-risk individuals. Lancet. 1990;335:1400-1.

39. Lefrere JJ, Mariotti M, Ferrer-Le-Coeur F, Rouger P, Noel B, Bosser C. PCR testing in HIV-1 seronegative haemophilia [Letter]. Lancet. 1990;336: 1386.

40. Lefrere JJ, de Montalembert M, Mariotti M, Girot R, Salmon C, Rouger P, et al. Absence of HIV DNA sequences in seronegative polytransfused thalassemic patients. Vox Sang. 1990;59:218-21.

41. Kwok S, Mack DH, Sninsky JJ, Ehrlich GD, Poiesz BJ, Dock NL, et al. Diagnosis of human immunodeficiency virus in seropositive individuals: enzymatic amplification of HIV viral sequences in peripheral blood mononuclear cells. In: Luciw PA, Steimer KS, eds. HIV Detection by Genetic Engineering Methods. New York: Marcel Dekker; 1989:243-55.

42. Wormser GP, Joline C, Bittker S, Forseter G, Kwok S, Sninsky JJ. Polymerase chain reaction for seronegative health care workers with parenteral exposure to HIV-infected patients [Letter]. N Engl J Med. 1989;321: 1681-2.

43. Gibbons J, Cory JM, Hewlett IK, Epstein JS, Eyster ME. Silent infections with human immunodeficiency virus type 1 are highly unlikely in multitransfused seronegative hemophiliacs. Blood. 1990;76:1924-6.

44. Genesca J, Wang RY, Alter HJ, Shih JW. Clinical correlation and genetic polymorphism of the human immunodeficiency virus proviral DNA obtained after polymerase chain reaction amplification. J Infect Dis. 1990;162:1025-30.

45. Perrin LH, Yerly S, Adami N, Bachmann P, Butler-Brunner E, Burckhardt J, et al. Human immunodeficiency virus DNA amplification and serology in blood donors. Blood. 1990;76:641-5.

46. Ou CY, McDonough SH, Cabanas D, Ryder TB, Harper M, Moore J, et al. Rapid and quantitative detection of enzymatically amplified HIV-1 DNA using chemiluminescent oligonucleotide probes. AIDS Res Hum Retroviruses. 1990;6:1323-30.

47. Kemp DJ, Churchill MJ, Smith DB, Biggs BA, Foote SJ, Peterson MG, et al. Simplified colorimetric analysis of polymerase chain reactions: detection of HIV sequences in AIDS patients. Gene. 1990;94:223-8.

48. Jackson JB, MacDonald KL, Cadwell J, Sullivan C, Kline WE, Hanson M, et al. Absence of HIV infection in blood donors with indeterminate western blot tests for antibody to HIV-1. N Engl J Med. 1990;322:217-22.

49. Jackson JB, Kwok SY, Sninsky JJ, Hopsicker JS, Sannerud KJ, Rhame FS, et al. Human immunodeficiency virus type 1 detected in all seropositive symptomatic and asymptomatic individuals. J Clin Microbiol. 1990;28:16-9.

50. Conway B, Adler KE, Bechtel LJ, Kaplan JC, Hirsch MS. Detection of HIV-1 DNA in crude cell lysates of peripheral blood mononuclear cells by the polymerase chain reaction and nonradioactive oligonucleotide probes. J Acquir Immune Defic Syndr. 1990;3:1059-64.

51. Ayehunie S, Sonnerborg A, Johansson B, Fehniger TE, Zewdie DW, Yeamane-Berhan T, et al. Differences in PCR reactivity between HIV proviruses from individuals in Ethiopia and Sweden. J Acquir Immune Defic Syndr. 1990;3:975-80.

52. Alter HJ, Epstein JS, Swenson SG, VanRaden MJ, Ward JW, Kaslow RA, et al. Prevalence of human immunodeficiency virus type 1 p24 antigen in United States blood donors—an assessment of the efficacy of testing in donor screening. The HIV-Antigen Study Group. N Engl J Med. 1990;323: 1312-7.

53. Kumar R, Goedert JJ, Hughes SH. A method for the rapid screening of human blood samples for the presence of HIV-1 sequences: the probe-shift assay. AIDS Res Hum Retroviruses. 1989;5:345-54.

54. Jehuda-Cohen T, Slade BA, Powell JD, Villinger F, De B, Folks TM, et al. Polyclonal B-cell activation reveals antibodies against human immunodeficiency virus type 1 (HIV-1) in HIV-1-seronegative individuals. Proc Natl Acad Sci U S A. 1990;87:3972-6.

55. Chadwick EG, Yogev R, Kwok S, Sninsky JJ, Kellogg DE, Wolinsky SM. Enzymatic amplification of the human immunodeficiency virus in peripheral blood mononuclear cells from pediatric patients. J Infect Dis. 1989; 160:954-9.

56. Albert J, Fenyo EM. Simple, sensitive, and specific detection of human immunodeficiency virus type 1 in clinical specimens by polymerase chain

57. Abbott MA, Poiesz BJ, Byrne BC, Kwok S, Sninsky JJ, Ehrlich GD. Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in-vitro. J Infect Dis. 1988;158:1158-69.

58. Coutlee F, Yang BZ, Bobo L, Mayur K, Yolken R, Viscidi R. Enzyme immunoassay for detection of hybrids between PCR-amplified HIV-1 DNA and a RNA probe: PCR-EIA. AIDS Res Hum Retroviruses. 1990;6:775-84.

59. Lifson AR, Stanley M, Pane J, O'Malley PM, Wilber JC, Stanley A, et al. Detection of human immunodeficiency virus DNA using the polymerase chain reaction in a well-characterized group of homosexual and bisexual men. J Infect Dis. 1990;161:436-9.

60. Dock NL, Kleinman SH, Rayfield MA, Schable CA, Williams AE, Dodd RY. Human immunodeficiency virus infection and indeterminate western blot patterns. Prospective studies in a low prevalence population. Arch Intern Med. 1991;151:525-30.

61. Schechter MT, Neumann PW, Weaver MS, Montaner JS, Cassol SA, Le TN, et al. Low HIV-1 proviral DNA burden detected by negative polymerase chain reaction in seropositive individuals correlates with slower disease progression. AIDS. 1991;5:373-9.

62. Yagi MJ, Joesten ME, Wallace J, Roboz JP, Bekesi JG. Human immunodeficiency virus type 1 (HIV-1) genomic sequences and distinct changes in CD8+ lymphocytes precede detectable levels of HIV-1 antibodies in high-risk homosexuals. J Infect Dis. 1991;164:183-8.

63. Schmidt BL. A rapid chemiluminescence detection method for PCR-amplified HIV-1 DNA. J Virol Methods. 1991;32:233-44.

64. Bagnarelli P, Menzo S, Manzin A, Varaldo PE, Montroni M, Giacca M, et al. Detection of human immunodeficiency virus type 1 transcripts in peripheral blood lymphocytes by the polymerase chain reaction. J Virol Methods. 1991;32:31-9.

65. Cassol S, Salas T, Lapointe N, Arella M, Rudnik J, O'Shaughnessy M. Improved detection of HIV-1 envelope sequences using optimized PCR and inosine-substituted primers. Mol Cell Probes. 1991;5:157-60.

66. Cassol S, Salas T, Arella M, Neumann P, Schechter MT, O'Shaughnessy M. Use of dried blood spot specimens in the detection of human immunodeficiency virus type 1 by the polymerase chain reaction. J Clin Microbiol. 1991;29: 667-71.

67. Celum CL, Coombs RW, Lafferty W, Inui TS, Louie PH, Gates CA, et al. Indeterminate human immunodeficiency virus type 1 western blots: seroconversion risk, specificity of supplemental tests, and an algorithm for evaluation. J Infect Dis. 1991;164:656-64.

68. Dahlen PO, Iitia AJ, Skagius G, Frostell A, Nunn MF, Kwiatkowski M. Detection of human immunodeficiency virus type 1 by using the polymerase chain reaction and a time-resolved fluorescence-based hybridization assay. J Clin Microbiol. 1991;29:798-804.

69. de la Salle C, Baas MJ, Laustriat D, Guy B, Bordes E, Wiesel ML, et al. DNA amplification of HIV genome in hemophiliacs and in newborns from seropositive mothers. Ann Hematol. 1991;62:165-8.

70. Ferrer-Le-Coeur F, Mariotti M, Hivert P, Satre EP, Bouchardeau F, Courouce AM, et al. No evidence of HIV-1 infection in seronegative hemophiliacs and in seronegative partners of seropositive hemophiliacs through polymerase chain reaction (PCR) and anti-NEF serology. Thromb Haemost. 1991; 65:478-82.

71. Gibson KM, McLean KA, Clewley JP. A simple and rapid method for detecting human immunodeficiency virus by PCR. J Virol Methods. 1991;32: 277-86.

72. Lefrere JJ, Mariotti M, Vittecoq D, Noel B, Courouce AM, Lambin P, et al. No evidence of frequent human immunodeficiency virus type 1 infection in seronegative at-risk individuals. Transfusion. 1991;31:205-11.

73. Lefrere JJ, Mariotti M, Salpetrier J, Demange F, Wattel E, Rouger P, et al. Polymerase chain reaction (PCR) in various stages of HIV infection. Relationship to disease progression. Nouv Rev Fr Hematol. 1991;33:245-9.

74. Pan LZ, Sheppard HW, Winkelstein W, Levy JA. Lack of detection of human immunodeficiency virus in persistently seronegative homosexual men with high or medium risks for infection. J Infect Dis. 1991;164:962-4.

75. Scarlatti G, Lombardi V, Plebani A, Principi N, Vegni C, Ferraris G, et al. Polymerase chain reaction, virus isolation and antigen assay in HIV-1-antibody-positive mothers and their children. AIDS. 1991;5:1173-8.

76. Yerly S, Chamot E, Deglon JJ, Hirschel B, Perrin LH. Absence of chronic human immunodeficiency virus infection without seroconversion in intravenous drug users: a prospective and prospective study. J Infect Dis. 1991; 164:965-8.

77. Zachar V, Mayer V, Aboagye-Mathiesen G, Norskov-Lauritsen N, Ebbesen P. Enhanced chemiluminescence-based hybridization analysis for PCR-mediated HIV-1 DNA detection offers an alternative to ^{32}P-labelled probes. J Virol Meth. 1991;33:391-5.

78. Coutlee F, Saint-Antoine P, Olivier C, Voyer H, Kessous-Elbaz A, Berrada F, et al. Evaluation of infection with human immunodeficiency virus type 1 by using nonisotopic solution hybridization for detection of polymerase chain reaction-amplified proviral DNA. J Clin Microbiol. 1991;29: 2461-7.

79. Bruisten SM, Koppelman MH, van der Poel CL, Huisman JG. Enhanced detection of HIV-1 sequences using polymerase chain reaction and a liquid hybridization technique. Application for individuals with questionable HIV-1 infection. Vox Sang. 1991;61:24-9.

80. Nielsen C, Teglbjaerg LS, Pedersen C, Lundgren JD, Nielsen CM, Vestergaard BF. Prevalence of HIV infection in seronegative high-risk individuals examined by virus isolation and PCR. J Acquir Immune Defic Syndr. 1991;4:1107-11.

81. Ensoli F, Fiorelli V, Mezzaroma I, D'Offizi G, Rainaldi L, Luzi G, et al. Plasma viraemia in seronegative HIV-1-infected individuals. AIDS. 1991;5: 1195-9.
82. Aiuti F, Ensoli F, Fiorelli V, Mezzaroma I, Pinter E, Guerra E, et al. Silent HIV infection. Vaccine. 1993;11:538-41.
83. Bailly E, Kleim JP, Schneweis KE, van Loo B, Hammerstein U, Brackmann HH. Absence of human immunodeficiency virus (HIV) proviral sequences in seronegative hemophilic men and sexual partners of HIV-seropositive hemophiliacs. Transfusion. 1992;32:104-8.
84. Bournique B, Akar A, Broly H, Ajana F, Counis R, Scholler R. Detection of HIV-1 infections by PCR: evaluation in a seropositive subject population. Mol Cell Probes. 1992;6:443-50.
85. Brettler DB, Somasundaran M, Forsberg AF, Krause E, Sullivan JL. Silent human immunodeficiency virus type 1 infection: a rare occurrence in a high-risk heterosexual population. Blood. 1992;80:2396-400.
86. Bruisten SM, Koppelman MH, Dekker JT, Bakker M, de Goede RE, Roos MT, et al. Concordance of human immunodeficiency virus detection by polymerase chain reaction and by serologic assays in a Dutch cohort of seronegative homosexual men. J Infect Dis. 1992;166:620-2.
87. Coutlee F, Olivier C, Cassol S, Voyer H, Kessous-Elbaz A, Saint-Antoine B, et al. Absence of prolonged immunosilent infection with human immunodeficiency virus in individuals with high-risk behaviors. Am J Med. 1994;96:42-8.
88. Dannatt AH, Goodwin SJ, Dasani H, Bowen DJ, Peake IR, Bloom AL. The relationship of HIV-1 viral sequences detected by the polymerase chain reaction in haemophilic patients to clinical and other markers of infection. Clin Lab Haematol. 1992;14:1-7.
89. Dawood MR, Allan R, Fowke K, Embree J, Hammond GW. Development of oligonucleotide primers and probes against structural and regulatory genes of human immunodeficiency virus type 1 (HIV-1) and their use for amplification of HIV-1 provirus by using polymerase chain reaction. J Clin Microbiol. 1992;30:2279-83.
90. Defer C, Agut H, Garbarg-Chenon A, Moncany M, Morinet F, Vignon D, et al. Multicentre quality control of polymerase chain reaction for detection of HIV DNA. AIDS. 1992;6:659-63.
91. Eble BE, Busch MP, Khayam-Bashi H, Nason MA, Samson S, Vyas GN. Resolution of infection status of human immunodeficiency virus (HIV)-seroindeterminate donors and high-risk seronegative individuals with polymerase chain reaction and virus culture: absence of persistent silent HIV type 1 infection in a high-prevalence area. Transfusion. 1992;32:503-8.
92. Gupta P, Kingsley L, Anderson H, Ho M, Enrico A, Ding M, et al. Low prevalence of HIV in high-risk seronegative homosexual men evidenced by virus culture and polymerase chain reaction. AIDS. 1992;6:143-9.
93. He Y, Coutlee F, Saint-Antoine P, Olivier C, Voyer H, Kessous-Elbaz A. Detection of polymerase chain reaction-amplified human immunodeficiency virus type 1 proviral DNA with a digoxigenin-labeled RNA probe and an enzyme-linked immunoassay. J Clin Microbiol. 1993;31:1040-7.
94. Henrard DR, Laurian Y, Mehaffey WF, Allain JP. Lack of detection of human immunodeficiency virus type 1 DNA by polymerase chain reaction in the plasma and lymphocytes of seronegative exposed hemophiliacs. Transfusion. 1993;33:405-8.
95. Jackson JB, Drew J, Lin HJ, Otto P, Bremer JW, Hollinger FB, et al. Establishment of a quality assurance program for human immunodeficiency virus type 1 DNA polymerase chain reaction assays by the AIDS Clinical Trials Group. ACTG PCR Working Group, and the ACTG PCR Virology Laboratories. J Clin Microbiol. 1993;31:3123-8.
96. Jehuda-Cohen T, Vonsover A, Miltchen R, Bentwich Z. 'Silent' HIV infection among wives of seropositive HIV carriers in the Ethiopian community in Israel. Scand J Immunol Suppl. 1992;11:81-3.
97. Kelen GD, Chanmugam A, Meyer WA 3d, Farzadegan H, Stone D, Quinn TC. Detection of HIV-1 by polymerase chain reaction and culture in seronegative intravenous drug users in an inner-city emergency department. Ann Emerg Med. 1993;22:769-75.
98. Kunakorn M, Wichukchinda N, Raksakait K, Petchclai B, Jutavijittum P, Mandee Y, et al. Screening of HIV-1 proviral DNA using simple protocol polymerase chain reaction in Thai blood donors [Letter]. AIDS. 1993;7: 1681-2.
99. Luque F, Leal M, Pineda JA, Torres Y, Aguado I, Olivera M, et al. Failure to detect silent HIV infection by polymerase chain reaction in subjects at risk for heterosexually transmitted HIV type 1 infection. Eur J Clin Microbiol Infect Dis. 1993;12:663-7.
100. Lynch CE, Madej R, Louie PH, Rodgers G. Detection of HIV-1 DNA by PCR: evaluation of primer pair concordance and sensitivity of a single primer pair. J Acquir Immune Defic Syndr. 1992;5:433-40.

101. Mariotti M, Wattel E, Lefrere F, Demange F, Lefrere JJ. Polymerase chain reaction procedure for rapid diagnosis of HIV infection [Letter]. AIDS. 1993;7:1680-1.
102. Moodley D, Reddy K, Smuts H, Govender T, Coovadia HM. Heterogeneity of HIV-1 in South Africa detected by polymerase chain reaction [Letter]. AIDS. 1993;7:1538-9.
103. Nandi J, Banerjee K, Thakar M, Bhavalkar V, Rodrigues J. Human immunodeficiency virus-1 infection in spouses of seropositive individuals. Natl Med J India. 1993;6:156-9.
104. Palenicek J, Fox R, Margolick J, Farzadegan H, Hoover D, Odaka N, et al. Longitudinal study of homosexual couples discordant for HIV-1 antibodies in the Baltimore MACS Study. J Acquir Immune Defic Syndr. 1992; 5:1204-11.
105. Quiros E, Garcia F, Gonzalez I, Cabezas T, Bernal MC, Maroto MC. Diagnosis of HIV-1 infection by PCR with two primer pairs. Eur J Epidemiol. 1993;9:426-9.
106. Rapier JM, Villamarzo Y, Schochetman G, Ou CY, Brakel CL, Donegan J, et al. Nonradioactive, colorimetric microplate hybridization assay for detecting amplified human immunodeficiency virus DNA. Clin Chem. 1993;39:244-7.
107. Sauvaigo S, Barlet V, Guettari N, Innocenti P, Parmentier F, Bastard C, et al. Standardized nested polymerase chain reaction-based assay for detection of human immunodeficiency virus type 1 DNA in whole blood lysates. J Clin Microbiol. 1993;31:1066-74.
108. Sevall JS, Prince H, Garratty G, O'Brien WA, Zack JA. Rapid enzymatic analysis for human immunodeficiency virus type 1 DNA in clinical specimens. Clin Chem. 1993;39:433-9.
109. Whetsell AJ, Drew JB, Milman G, Hoff R, Dragon EA, Adler K, et al. Comparison of three nonradioisotopic polymerase chain reaction-based methods for detection of human immunodeficiency virus type 1. J Clin Microbiol. 1992;30:845-53.
110. Willerford DM, Bwayo JJ, Hensel M, Emonyi W, Plummer FA, Ngugi EN, et al. Human immunodeficiency virus infection among high-risk seronegative prostitutes in Nairobi. J Infect Dis. 1993;167:1414-7.
111. Yourno J, Conroy J. A novel polymerase chain reaction method for detection of human immunodeficiency virus in dried blood spots on filter paper. J Clin Microbiol. 1992;30:2887-92.
112. Zazzi M, Romano L, Brasini A, Valensin PE. Nested polymerase chain reaction for detection of human immunodeficiency virus type 1 DNA in clinical specimens. J Med Virol. 1992;38:172-4.
113. Zazzi M, Romano L, Brasini A, Valensin PE. Simultaneous amplification of multiple HIV-1 DNA sequences from clinical specimens by using nested-primer polymerase chain reaction. AIDS Res Hum Retroviruses. 1993;9:315-20.
114. Schwartz JS, Dans PE, Kinosian BP. Human immunodeficiency virus test evaluation, performance, and use. Proposals to make good tests better. JAMA. 1988;259:2574-9.
115. Update: serologic testing for HIV-1 antibody—United States, 1988 and 1989. MMWR Morb Mortal Wkly Rep. 1990;39:380-3.
116. MacDonald KL, Jackson JB, Bowman RJ, Polesky HF, Rhame FS, Balfour HH Jr, et al. Performance characteristics of serologic tests for human immunodeficiency virus type-1 (HIV-1) antibody among Minnesota blood donors. Ann Intern Med. 1989;110:617-21.
117. Horsburgh CR Jr, Ou CY, Jason J, Holmberg SD, Longini IM Jr, Schable C, et al. Duration of human immunodeficiency virus infection before detection of antibody. Lancet. 1989;2:637-40.
118. Landesman S, Weiblen B, Mendez H, Willoughby A, Goedert JJ, Rubinstein A, et al. Clinical utility of HIV-IgA immunoblot assay in the early diagnosis of perinatal HIV infection. JAMA. 1991;266:3443-6.
119. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med. 1978;299:926-30.
120. Philbrick JT, Horwitz RI, Feinstein AR. Methodologic problem in exercise testing for coronary artery disease: groups, analysis and bias. Am J Cardiol. 1980;46:807-12.
121. Arkin CF, Wachtel MS. How many patients are necessary to assess test performance? JAMA. 1990;263:275-8.
122. Mushlin AI, Detsky AS, Phelps CE, O'Connor PW, Kido DK, Kucharczyk W, et al. The accuracy of magnetic resonance imaging in patients with suspected multiple sclerosis. The Rochester-Toronto Magnetic Resonance Imaging Study Group. JAMA. 1993;269:3146-51.
123. Hurlbut TA 3d, Littenberg B. The diagnostic accuracy of rapid dipstick tests to predict urinary tract infection. Am J Clin Pathol. 1991;96:582-8.
124. Littenberg B, Mushlin AI. Technetium bone scanning in the diagnosis of osteomyelitis: a meta-analysis of test performance. Diagnostic Technology Assessment Consortium. J Gen Intern Med. 1992;7:158-63.