

A New Approach for Extreme Proportion Calculation: Median Estimate and Its Application to Evaluate Performance of HIV Assays With 100% Test Sensitivity or Specificity

Jin Liu (✉ liujin@jsph.org.cn)

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

<https://orcid.org/0000-0002-0557-026X>

Fang Shao

Nanjing Medical University

Jiaying Yang

Southeast University

Jinxi Liu

AstraZeneca R&D Shanghai, China

Technical advance

Keywords: Extreme proportion, diagnostic accuracy, positive predictive value, HIV, evaluation of assays, median estimate

Posted Date: March 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-358014/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A new approach for extreme proportion calculation: median estimate and its application to evaluate performance of HIV assays with 100% test sensitivity or specificity

Jin Liu^{1*}, Fang Shao², Jiaying Yang³, Jinxi Liu⁴

¹Clinical Research Institute, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China.

²Department of Biostatistics; School of Public Health; Nanjing Medical University; Nanjing, China

³Department of Epidemiology and Biostatistics, School of Public Health, Southeast University, Nanjing, China

⁴R&D China, AstraZeneca, Shanghai, China

*Corresponding author: Jin Liu

324, Guangzhou Road

Nanjing, 210029, China

Tel: +86-025-68306156

Email: liujin@jsph.org.cn

22 **Abstract**

23 **Background:** In the evaluation of performance of HIV assays, extreme sample
24 proportions often occur, with test sensitivity and/or specificity of 100%, which
25 making it challenging to assess the assays accuracies. To overcome these challenges,
26 we propose using median estimate as an evaluation indicator for such testing.

27 **Methods:** Based on the principles of binomial distribution and confidence interval,
28 median estimate was defined as $p = 0.5^{\frac{1}{n}}$, which means that, when the sample size n is
29 equal to the event number x , namely the sample proportion (e.g., test sensitivity) is
30 100%, the 50th percentile (median) of p (the estimate of population proportion) is
31 $0.5^{\frac{1}{n}}$. After demonstrating the mathematical proof of the median estimate, the key
32 programming commands of commercial software SAS and free software R were given.
33 Subsequently, we developed an Excel-based calculation tool that allows users to fill in
34 data in an Excel sheet without writing any program. Six cases of HIV screening and
35 diagnostic tests and HIV infections incidence data were selected from related articles
36 and World Health Organization reports published between 2009 and 2020.

37 **Results:** The median estimates, which were proved in the range from $\frac{n-1}{n}$ to 1 and
38 within the confidence interval range, showed statistical plausibility. Six HIV testing
39 cases were presented to illustrate its application and elaborate on the relationship
40 between the median estimate and the conventional simple estimate. These cases
41 demonstrate that, when extreme proportions occurred (i.e., false positive and/or false
42 negative in testing were zero), the conventional simple estimates of sensitivity,
43 specificity, positive predictive value, and negative predictive value were 100%
44 regardless of the sample size and prevalence. In contrast, the corresponding median
45 estimates varied depending on the sample size and prevalence.

46 **Conclusions:** As evaluation indicators of HIV assays with extreme proportions,
47 median estimates were more effective than conventional simple estimates. However,
48 simple estimates objectively expressed the results of HIV testing. Because the

49 correlation between median estimates and simple estimates was seamless, the two types
50 of indicators were complementary in the evaluation of testing with extreme proportions.
51 Hence, using both types of estimates could help evaluate HIV assays with extreme
52 proportions more comprehensively.

53 **Keywords:** Extreme proportion, diagnostic accuracy, positive predictive value, HIV,
54 evaluation of assays, median estimate

Background

The performance of HIV assays is typically described in terms of sensitivity and specificity which are defined as the probabilities of correctly identifying a patient and a non-patient, respectively [1]. One significant feature of HIV screening and diagnostic tests is the high accuracy of these assays; for example, sensitivity and/or specificity can be as high as 100% compared to reference assays. According to the WHO/UNAIDS report [2], 62.5% (20/32) of enzyme-linked immunosorbent assays (ELISA) and rapid diagnostic tests (RDTs) demonstrated 100% sensitivity. In the China CDC report, 91.7% (22/24) of ELISA and RDTs achieved 100% sensitivity [3]. A study in Tanzania used five types of assays for large-scale screening of HIV antibodies in four populations (hospital patients, pregnant women, voluntary participants and testing attendees, and blood donors) showed that 100% sensitivity was achieved in 85% (16/20) of the tests, whereas 100% specificity was achieved in 60% of the tests (12/20) [4]; similar results have been reported in other countries [5-7].

The occurrence of extreme proportion (i.e., the point estimate of HIV test sensitivity and/or specificity reaches 100%) presents challenges for the assessment of test accuracy. Because inherent problems of false negative and false positive are inevitable in HIV testing [8, 9], these challenges are presented in two aspects: one is that the point estimate calculated by the current method is fixed for any value of sample size when extreme proportion occur [5, 10]. Another is that some comprehensive indicators fail to reasonably assess the operational performance of HIV assays when the sample proportion is 100% [10, 11]. The study by Manak et al.

77 shows that it is difficult to interpret the results of HIV screening tests using the
78 conventional simple estimates when a combination of two RDTs has 100% specificity.
79 Therefore, the positive predictive value and the positive likelihood ratio, two important
80 comprehensive indicators, are shown as "NaN" and "infinity", respectively, in their
81 article [11]. Similar challenges have also emerged in the recent evaluation of 10
82 commercially available COVID-19 Enzyme immunoassays [12]. Therrien et al.
83 showed that the four assays with 100% specificity (at two sites) showed 100% positive
84 predictive value at every specified prevalence level. Same positive predictive value
85 (100%) in a prevalence ranging between 1% and 15% may preclude results
86 interpretation [12].

87 To date, either frequentist or Bayesian approaches [10, 13-15] have been proposed
88 for the evaluation of diagnostic tests with 100% sensitivity and/or specificity; however,
89 challenges remain in this regard. To overcome these challenges, we propose a new
90 indicator, called the median estimate. This work elaborates on the importance of the
91 median estimate and the detailed calculation of its point estimates and interval
92 estimates. Moreover, six HIV screening cases [2, 4, 6-7, 9] from the WHO, China,
93 Tanzania, and other countries were used to illustrate the application of median
94 estimates and elaborate on the relationship between the median estimate and
95 conventional estimates. Key program scripts developed in the commercial software
96 SAS (SAS Institute Inc., Cary, NC, USA) and the free software R (the R Project for
97 Statistical Computing) are provided herein to facilitate the application of this new
98 approach. Furthermore, we developed an Excel-based calculation tool that allows users

to fill data in an Excel sheet without writing any program script.

Methods

Evaluation indicators for diagnostic tests

The results of diagnostic test are classified into four categories: true positive (TP), false positive (FP), true negative (TN) and false negative (FN). Sensitivity (Se) is defined as $Se = TP / (TP + FN)$. Specificity (Sp) is defined as $Sp = TN / (TN + FP)$. Let p be the estimate of the target population proportion (θ), the positive predictive value (PPV) and negative predictive value (NPV) are defined as follows: [16],

$$PPV = \frac{p \times Se}{p \times Se + (1 - p) \times (1 - Sp)}$$

$$NPV = \frac{(1 - p) \times Sp}{(1 - p) \times Sp + p \times (1 - Se)}$$

Thus, positive and negative predictive values are not intrinsic to an assay but are markedly affected by prevalence [16]. The development of testing strategies for HIV diagnosis is guided by the predictive values of test results that best indicate the proportion of results representing the actual status [1].

In the evaluation of the performance of HIV assays, positive likelihood ratio ($PLR = Se / (1 - Sp)$) and negative likelihood ratio ($NLR = (1 - Se) / Sp$) are two common comprehensive indicators. The larger the PLR, the better the ability of the assay to confirm HIV infection. The smaller the NLR, the better the ability of the assay to exclude HIV infection.

Point estimation of performance of HIV assays

When the sensitivity and/or specificity of a HIV assay are not equal to 100%, its performance can be easily calculated by the expression of $\frac{x}{n}$, where x is the number of events and n is the total number of tests; in the present study, this expression is referred to as the simple estimate because of the straightforward calculation involved. However, when the sensitivity and/or specificity are 100%, the performance of an HIV assay cannot be reasonably evaluated using the conventional simple estimate. To overcome this dilemma, we introduced a novel indicator called the median estimate, which is defined as follows:

Let X denote a binomial random variable for n trials with population proportion θ , and p denote the estimate of θ ; when $x=n$, the 50th percentile of p is defined as the median estimate,

$$p_{50} = 0.5^{\frac{1}{n}} \quad (1)$$

Further, it was proved that

$$\frac{n-1}{n} < p_{50} < 1 \quad (2)$$

For example, when $n=100$, $p_{50} = 0.5^{\frac{1}{100}} = 99.31\%$, which means that when $x=n$, that is, the sample proportion is 100%, the 50th percentile (median) of p (the estimate of the population proportion of interest) is 99.31%. The median estimate is guaranteed to be in the range from $\frac{n-1}{n} = \frac{100-1}{100} = 99\%$ to 100%. Mathematical proofs of formulae (1) and (2) are provided in Additional file 1.

Confidence intervals of performance characteristics of HIV assays

To evaluate HIV assays, the WHO uses the Exact (Clopper-Pearson) confidence interval [2]. The F -distribution method [17] is adopted in many software packages

available from both commercial and open-source statistical packages (such as SAS and R). The lower and upper confidence limits of $1 - \alpha$ confidence level, respectively, are defined as follows:

$$\theta_L = \left(1 + \frac{n-x+1}{x F(\frac{\alpha}{2}, 2x, 2(n-x+1))}\right)^{-1} \quad (3)$$

$$\theta_U = \left(1 + \frac{n-x}{(x+1) F(1-\frac{\alpha}{2}, 2(x+1), 2(n-x))}\right)^{-1} \quad (4)$$

Where $F(\frac{\alpha}{2}, b, c)$ is the $(\frac{\alpha}{2}) \times 100$ th percentile of the F distribution with b and c representing the degrees of freedom.

Further, it was proved (see Additional file 1) that when $x = n$, the lower limit of the confidence interval $1 - \alpha$ is reduced to

$$\theta_L = \alpha^{\frac{1}{n}} \quad (5).$$

When $x = n$, the lower limit of the confidence interval calculated by formulae (3) and (5) are equal, and the upper confidence limit is defined as 100% [17]. When $x < n$, formulae (3) and (4) are used to calculate the confidence interval.

Data for case analysis

To illustrate the application of the median estimate, Six cases of HIV screening and diagnostic tests and HIV infections incidence data were selected from related articles and World Health Organization reports, which are previously published between 2009 and 2020 [2- 4, 6-7, 9].

Available software and calculating tool

In the commercial software SAS and the free software R, the key commands to calculate the median estimate and lower limit of the 95% confidence interval, respectively, were:

$p_{50} = 0.5 ** (1/n); p_{0.025} = 0.025 ** (1/n).$

In Excel, the corresponding commands were:

$p_{50} = POWER(0.5, 1/n); p_{0.025} = POWER(0.025, 1/n).$

Since Excel is the most popular data processing software for personal computers, we developed an Excel calculation tool (see Additional file 2) to facilitate the practical application. The results of the Excel calculation tool were verified using the SAS commercial software.

Results

Proposed median estimate

To illustrate the reasons for the proposed median estimate, the theoretical percentages when the sample proportion was 100% for various population proportions and sample sizes were calculated using formula (1) in Additional file 1.

Table 1 Theoretical percentages at sample proportion of 100% for various population proportions (θ) and sample sizes (n)

n	θ	98.00	99.00	99.90	99.99
10		81.71	90.44	99.00	99.90
50		36.42	60.50	95.12	99.50
100		13.26	36.60	90.48	99.00
200		1.76	13.40	81.87	98.02
460†		0.01	0.98	63.11	95.50
600		0.00	0.24	54.87	94.18
1000		0.00	0.00	36.77	90.48

†The number of specimens of WHO HIV reference panel to evaluate sensitivity for rapid HIV assays [2].

Table 1 shows that the greater the population proportion or the smaller the sample size, the greater the chance of a 100% sample proportion. For example, when the population proportion θ was 99.9% (common for sensitivity and specificity in ELISA and RDTs to detect HIV) and the sample size was 460 (the number of specimen from the WHO/UNAIDS reference panel to evaluate the sensitivity of HIV rapid assays [2]), the probability of 100% sample proportion was 63.11%. When θ was 99.99%, the probability was as high as 95.50%. Since no assay among the current HIV tests has a population sensitivity or specificity of 100% [10, 18], it is unreasonable to consider a sample proportion of 100% as the estimate of the population proportion. In fact, for extreme accuracy tests, the smaller the sample size, the more likely it is to obtain a seemingly perfect test result with 100% sensitivity or specificity. These concepts are extremely important to understand the rationale and practical significance of the median estimate proposed in this study.

Case analysis

The sensitivity, specificity, PPV, and NPV calculated based on the respective median estimates and simple estimates, the corresponding 95% confidence interval, and raw data from six cases that used for evaluation of the performance characteristics of HIV assays [2-4, 6-7, 9] are summarised in Table 2.

As shown in Table 2, extreme test results with a false positive (FP) or false negative (FN) of 0 resulted in 100% sensitivity (Se) and specificity (Sp) of simple estimates, regardless of sample size. Another notable result was that these simple estimates of 100% were not within the corresponding 95% confidence interval, but equal to the upper limit

of the confidence interval. However, the corresponding median estimates (SeM and SpM) were not 100% and within the corresponding confidence intervals. For example, in assay 1, when the sample size was 460, the median estimate of sensitivity was 99.8%, which was contained in the confidence interval (99.2-100.0). Moreover, the median estimate depended on the sample size.

Table 3 demonstrates that the PPV calculated using simple estimate was always equal to 100%, regardless of the prevalence, when the specificity was 100%. On the contrary, the PPVM varied with the prevalence. For example, the PPVM was 50.39% for assay 1 in African region, whereas the PPVM in Western Pacific region was 5.38%. When the specificity was not equal to 100%, the values of PPV and PPVM were close.

Table 3 shows that NPV was always 100%, irrespective of prevalence, when sensitivity was 100%. However, NPVM overcame this limitation, reflecting the effect of prevalence on them. Notably, the occurrence of NPVM=100% in Table 3 results from rounding of the calculation. According to its calculation formula, as the prevalence decreases and the sensitivity increases, NPVM gets closer and closer to 100%, but is not exactly equal to 100%.

Table 4 shows that when the specificity was 100%, PLR became infinite (∞) since the denominator was 0; however, PLRM did not become infinite. When sensitivity was equal to 100%, NLR was 0, irrespective of sample size; however, NLRM overcame this limitation.

223

224 **Table 2.** Comparisons of sensitivities, specificities of HIV assays, calculated using simple estimation and median estimation

Assay	Source	Results of test				Sensitivity (%)			Specificity (%)		
		TP	FP	FN	TN	Se	SeM	95% CI	Sp	SpM	95% CI
1 [2]	WHO	460	0	0	658	100	99.8	99.2-100.0	100	99.9	99.4-100.0
2 [6]	Cameroon	235	0	5	240	97.9	-	95.2-99.1	100	99.7	98.5-100.0
3 [7]	Ghana	140	24	0	116	100	99.5	97.4-100.0	82.9	-	75.6-88.1
4 [3]	China	164	0	0	286	100	99.6	97.8-100.0	100	99.8	98.7-100.0
5 [4]	Tanzania	390	4	0	1039	100	99.8	99.1-100.0	99.6	-	99.0-99.8
6 [9]	sub-Sahara Africa	222	0	0	224	100	99.7	98.4-100.0	100	99.7	98.4-100.0

225 TP: true positive; FP: false positive; FN: false negative; TN: true negative; CI: confidence interval; Se, SeM, Sp, SpM, PPV, PPVM, NPV and
226 NPVM: sensitivity, specificity, positive predictive value and negative predictive value calculated by simple estimation and median estimation,
227 respectively; “-”: the median estimate did not be calculated.

228

229 **Table 3.** Comparisons of predictive values in different regions, calculated using simple estimation and median estimation

Assays	African Region		Western Pacific Region		African Region		Western Pacific Region	
	p=0.107% [19]		p=0.006% [19]		p=0.107% [19]		p=0.006% [19]	
	PPV	PPVM	PPV	PPVM	NPV	NPVM	NPV	NPVM
1 [2]	100.000	50.393	100.000	5.384	100.000	100.000	100.000	100.000
2 [6]	100.000	26.669	100.000	1.997	99.998	99.998	100.000	100.000
3 [7]	0.621	0.618	0.035	0.035	100.000	99.999	100.000	100.000
4 [3]	100.000	30.586	100.000	2.409	100.000	100.000	100.000	100.000
5 [4]	21.832	21.802	1.540	1.538	100.000	100.000	100.000	100.000
6 [9]	100.000	25.684	100.000	1.899	100.000	100.000	100.000	100.000

230 p: HIV infections incidence; PPV, PPVM, NPV and NPVM: positive predictive value and negative predictive value calculated by simple

231 estimation and median estimation, respectively.

232

Table 4. Comparisons of likelihood ratios of HIV assays, calculated using simple estimation and median estimation

Assays	PLR	PLRM	NLR	NLRM
1 [2]	infinity	948.363	0.000	0.002
2 [6]	infinity	339.523	0.021	0.021
3 [7]	5.833	5.805	0.000	0.006
4 [3]	infinity.	411.369	0.000	0.004
5 [4]	260.75	260.287	0.000	0.002
6 [9]	infinity	322.655	0.000	0.003

PLR, PLRM, NLR and NLRM: positive likelihood ratio and negative likelihood ratio calculated by simple estimation and median estimation, respectively.

Discussion

Extreme results of HIV testing assays (i.e., zero false positive and/or false negative) make it challenging to interpret the calculated sample proportion using conventional simple estimate. As shown in Tables 2 to 4, the simple estimate was always 100%, regardless of sample size and prevalence. Although some strategies have been proposed to overcome this challenge, such as substituting 100% directly [13] and adding “pseudo-counts” of successes and failures [14], these strategies do not consider the effect of sample size. Other researchers and we have tried to use Bayesian methods to overcome this dilemma, and the results obtained are similar to those obtained using median estimate [10, 20]. However, Bayesian statistical calculations involve

computationally intensive techniques and require specialised software [10], being difficult for researchers unfamiliar with such tools. Recently, Kravitz et al. reported the worst- and best- case performance estimates calculated using the lower and upper 95% bounds for each test, respectively, and noted that algorithms using data from the lower end of the confidence interval had lower the diagnostic accuracy, whereas those using data from the upper end of the confidence interval often outperformed anticipations, even in comparison to the manufacturers' performance claims stated in the instructions for use [15].

The limitations of simple estimate and the study by Kravitz et al. [15] inspired us to propose a new indicator. Ideally, this new indicator should have the following characteristics: reasonable methodology, simple calculation, and easy to understand for clinical researchers; the median estimate proposed in this study satisfies these requirements. First, the calculation of median estimate was based on binomial distribution and interval estimation (see Proof of Formula (1) in Additional file 1), and the median estimate was proved to be in the range from $\frac{n-1}{n}$ to 1, which showed its statistical plausibility (see Proof of Formula (2) in Additional file 1). Second, the median estimate (Formula (1)) can be easily understood by non-statisticians because of its simple definition and intuitive accompanying examples. Finally, its calculation was straightforward using an Excel sheet that allowed users to directly fill in the data without writing program scripts (see Additional file 2).

Practical cases examined in this study demonstrated that median estimates provided more reasonable evaluation of HIV testing assays with extreme performance than

simple estimates. The rationales for using median estimate are presented in terms of both clinical practice and statistical methods. First, in clinical practice, HIV assays have false positives or false negatives in HIV screening and diagnostic testing [10]. When evaluating high accuracy assays, the common reason for 100% sample proportion is that the sample size that can be detected is too small relative to the high accuracy population proportion of interest to detect rare false negatives and false positives, resulting in 100% sample sensitivity and specificity. However, this dilemma can be overcome using the median estimate. Second, based on statistical theory, the point estimate of a proportion should be within the range of the confidence interval, but when x is equal to n , the simple estimate is at the upper limit of the confidence interval. Clearly, this result contradicts the intuitive perception and statistical definition of confidence intervals [21]. In contrast, the median estimate is within the confidence interval.

The median estimate has remarkable advantages in the evaluation of comprehensive indicators such as predictive values and likelihood ratios. PPV is the most critical indicator to the WHO because it reflects the real state of the epidemic [1]. Careful scrutiny of Table 3 revealed that when extreme data occurred, the PPVs calculated using simple estimates were always equal to 100%, regardless of prevalence. Apparently, this is unreasonable in both theory and practice. In contrast, the PPVs calculated using the median estimates varied with prevalence. These results were consistent with the fact that the higher the prevalence of HIV infection in the population, the greater the probability that individuals who tested positive were truly infected (i.e.,

the greater the PPV) [16]. In addition, the PLRs of simple estimates in Table 4 have implausible results of infinity (∞), which is consistent with the findings of Manak et al. [11]. However, the PLRs of the median estimates presented reasonable results.

Finally, two viewpoints need to be emphasised. First, the simple estimates objectively express the outcomes of HIV tests, regardless of whether x is equal to n or less than n . Herein, we suggest that when using median estimates to evaluate the performance of HIV assays with extreme data, this objective result should be reported in an appropriate form, such as assay 1 shown in Table 2, whose sensitivity could be reported as “SeM=99.8% (460/460)” or described as “the median sensitivity is 99.8% with both positive number and sample size of 460”. The interval estimates of the two methods are exactly the same when $x=n$ (Table 2). Second, the relation between the median estimate and the simple estimate is seamless owing to $p_{50} > \frac{n-1}{n}$. This suggests that median estimates should be used only when $x=n$ (that is, the sample sensitivity and/or specificity equal 100%). When $x<n$, the conventional simple estimate should be used. Therefore, these two types of indicators are complementary in the evaluation of HIV assays. Because the median estimate could compensate for the deficiencies of the simple estimate, it is proposed that the two estimates should be combined for more comprehensive evaluation of HIV assays with extreme test results.

Conclusions

In the evaluation of HIV test assays with extreme performance characteristics, such as substantial sensitivity and/or specificity of 100%, it is challenging to interpret the test

results. In this study, we proposed a novel evaluation indicator for such extreme proportion, the median estimate, and proposed a simplified calculation method for its confidence interval. The results demonstrated that the median estimates provided more reasonable results compared to simple estimates for extreme proportion; however, simple estimates objectively expressed the results of HIV testing. The relationship between the median estimate and simple estimate is seamless. Thus, the two types of estimates should be combined for a more comprehensive evaluation of HIV assays with extreme test results.

Additional File

Additional file 1: Proof of Formulae (1) and (2)

Formulae (1) and (2) were proved based on the principle of binomial distribution and interval estimation.

Additional file 2: Excel calculating tool

The Excel calculation tool was developed to evaluate the performance of the diagnostic test when the sensitivity and/or specificity are 100%. By inputting the raw data, the results of Tables 2, 3, and 4 were obtained. The calculation process is straightforward using an Excel sheet, and hence no specialised statistical software or programming knowledge is required.

Abbreviations

TP: True positive; FP: False positive; FN: False negative; TN: True negative; Se: Sensitivity; SeM: Median sensitivity; Sp: Specificity; SpM: Median specificity; PPV:

Positive predictive value; PPVM: Median positive predictive value; NPV: Negative predictive value; NPVM: Median negative predictive value; PLR: Positive likelihood ratio; PLRM: Median positive likelihood ratio; NLR: Negative likelihood ratio; NLRM: Median negative likelihood ratio; CI: Confidence interval; ELISA: Enzyme-linked immunosorbent assays; RDTs: Rapid diagnostic tests

Acknowledgements

The authors would like to thank Professor Pei Liu for helpful discussion regarding the statistical method.

Authors' contributions

JL contributed to the conception and design of this research. JL, FS, and JXL performed the research and analysed the data. JL wrote the manuscript with consultation from FS and JYY. All authors have read and approved the final manuscript.

Funding

This work received funding from the Youth Foundation of the National Natural Science Foundation of China (81602938).

Availability of data and materials

All data generated or used during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Authors details

¹Clinical Research Institute, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China. ²Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China. ³Department of Epidemiology and Biostatistics, School of Public Health, Southeast University, Nanjing, China. ⁴R&D China, AstraZeneca, Shanghai, China.

References

1. World Health Organization. ANNEX 7. Diagnostics for HIV diagnosis in: WHO Consolidated guidelines on HIV testing services: 5Cs: consent, confidentiality, counselling, correct results, and connection. Geneva, Switzerland: World Health Organization; 2015.
https://apps.who.int/iris/bitstream/handle/10665/180221/WHO_HIV_2015.26_eng.pdf;jsessionid=3E5FB256DBE1AFC452B4AB1E4AC7D2D3?sequence=1. Accessed 6 Mar 2021.
2. World Health Organization. HIV assays: laboratory performance and other operational characteristics: rapid diagnostic tests (combined detection of HIV-1/2 antibodies and discriminatory detection of HIV-1 and HIV-2 antibodies). Report 18. Geneva, Switzerland: World Health Organization; 2015.
https://www.who.int/diagnostics_laboratory/publications/15032_hiv_assay_report_18.pdf?ua=1,&ua=1. Accessed 6 Mar 2021.
3. Chinese Center for Disease Control and Prevention. Report on clinical quality evaluation of national HIV antibody diagnostic test in 2019. Beijing, China: Chinese Center for Disease Control and Prevention; 2019. In Chinese.

<http://www.chinacdc.cn/zxdt/201912/W020191218527929912835.pdf> . Accessed
6 Mar 2021.

4. Lyamuya EF, Aboud S, Urassa WK, Sufi J, Mbwana J, Ndugulile F, et al. Evaluation of simple rapid HIV assays and development of national rapid HIV test algorithms in Dar es Salaam, Tanzania. BMC Infect Dis. 2009;9:19.
<http://link.springer.com/content/pdf/10.1186/1471-2334-9-19>
5. Dagnra AY, Dossim S, Salou M, Nyasenu T, Ali-Edje K, Ouro-Médéli A, et al. Evaluation of 9 rapid diagnostic tests for screening HIV infection, in Lomé, Togo. Med Mal Infect. 2014;44(11-12):525-9.
<https://linkinghub.elsevier.com/retrieve/pii/S0399077X1400294>
6. Njouom R, Ngono L, Mekinda-Gometi DD, Ndé CK, Sadeuh-Mba SA, Vernet MA, et al. Evaluation of the performances of twelve rapid diagnostic tests for diagnosis of HIV infection in Yaounde, Cameroon. J Virol Methods. 2017;243:158-63.
<https://linkinghub.elsevier.com/retrieve/pii/S0166093416300507>
7. Boadu R, Darko G, Nortey P, Akweongo P, Sarfo B. Assessing the sensitivity and specificity of first Response HIV-1-2 test kit with whole blood and serum samples: a cross-sectional study. AIDS Res Ther. 2016;13(9):9.
<http://link.springer.com/content/pdf/10.1186/s12981-016-0092-0>
8. Klarkowski D, Glass K, O'Brien D, Lokuge K, Piriou E, Shanks L. Variation in specificity of HIV rapid diagnostic tests over place and time: an analysis of discordancy data using a Bayesian approach. PLOS ONE. 2013;8(11):e81656.
<https://dx.plos.org/10.1371/journal.pone.0081656>
9. Kosack CS, Shanks L, Beelaert G, Benson T, Savane A, Ng'ang'a A, et al. HIV misdiagnosis in sub-Saharan Africa: performance of diagnostic algorithms at six testing sites. J Int AIDS Soc. 2017;20(1):21419.
<http://doi.wiley.com/10.7448/IAS.20.1.21419>
10. Liu P, Yang HT, Qiang LY, Jin H, Xiao S, Shi ZX. Evaluation of 30 commercial assays for the detection of antibodies to HIV in China using classical and Bayesian statistics. J Virol Methods. 2010;170(1-2):73-9.
<https://linkinghub.elsevier.com/retrieve/pii/S0166093410003149>

11. Manak MM, Njoku OS, Shutt A, Malia J, Jagodzinski LL, Milazzo M, et al. Evaluation of performance of two rapid tests for detection of HIV-1 and -2 in high- and low-prevalence populations in Nigeria. *J Clin Microbiol.* 2015;53(11):3501-6. <https://jcm.asm.org/content/53/11/3501>
12. Therrien C, Serhir B, Bélanger-Collard M, Skrzypeczak J, Shank DK, Renaud C, et al. Multicenter evaluation of the clinical performance and the neutralizing antibody activity prediction properties of 10 high-throughput serological assays used in clinical laboratories. *J CLIN MICROBIOL.* 2021; 59(3): e02511-20. <https://jcm.asm.org/content/59/3/e02511-20>
13. Abramson JH. WINPEPI (PEPI-for-Windows): computer programs for epidemiologists. *Epidemiol Perspect Innov.* 2004;1(1):6. https://www.researchgate.net/publication/8122960_WINPEPI_PEPI-for-Windows_computer_programs_for_epidemiologists
14. Zou GY. From diagnostic accuracy to accurate diagnosis: interpreting a test result with confidence. *Med Decis Mak.* 2004;24(3):313-8. <https://journals.sagepub.com/doi/abs/10.1177/0272989x04265483>
15. Kravitz Del Solar AS, Parekh B, Douglas MO, Edgil D, Kuritsky J, Nkengasong J. A Commitment to HIV Diagnostic Accuracy - a comment on “Towards more accurate HIV testing in sub-Saharan Africa: a multi-site evaluation of HIV RDTs and risk factors for false positives ‘and’ HIV misdiagnosis in sub-Saharan Africa: a performance of diagnostic algorithms at six testing sites”. *J Int AIDS Soc.* 2018;21(8):e25177. <http://doi.wiley.com/10.1002/jia2.25177>
16. Zhou X, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. 2th ed. Hoboken, NJ: John Wiley & Sons, Inc; 2011. <https://www.wiley.com/en-us/Statistical+Methods+in+Diagnostic+Medicine-p-9780470317921>
17. Newcombe RG. Two-sided confidence intervals for the single proportion:

comparison of seven methods. Stat Med. 1998;17(8):857-72.

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=9595616&query_hl=1

18. Noble L, Scott L, Stewart-Isherwood L, Molifi SJ, Sanne I, Da Silva P, et al. Continuous quality monitoring in the field: an evaluation of the performance of the Fio Deki Reader™ for rapid HIV testing in South Africa. BMC Infect Dis. 2020;20(1):320.

<https://bmcinfectdis.biomedcentral.com/track/pdf/10.1186/s12879-020-4932-0.pdf>

19. World Health Organization. ANNEX 2. Tables of health statistics by country, WHO region and globally in: World health statistics 2020: monitoring health for the SDGs, sustainable development goals. Geneva, Switzerland: World Health Organization; 2020.

<https://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf>

. Accessed 6 Mar 2021.

20. Liu J, Chen F, Zhu FC, Bai JL, Li JX, Yu H, et al. Evaluation of vaccine seroresponse rates and adverse event rates through Bayesian and frequentist methods. Hum Vaccin Immunother. 2015;11(6):1557-63.

<https://www.tandfonline.com/doi/pdf/10.1080/21645515.2015.1008932>

21. Pradhan V, Evans JC, Banerjee T. Binomial confidence intervals for testing non-inferiority or superiority: a practitioner's dilemma. Stat Methods Med Res. 2016;25(4):1707-17.

<http://journals.sagepub.com/doi/10.1177/0962280213498324>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1.docx](#)
- [AdditionalFile2ExcelcalculatingtoolSuppInfo.xls.xls](#)