Research Article

Machine Learning Classification of False-Positive Human Immunodeficiency Virus Screening Results

Mahmoud Elkhadrawi[1], Bryan A Stevens[2], Bradley J Wheeler[3], Murat Akcakaya[1], Sarah Wheeler[4,5]

[1]Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA, [2]Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA, [3]Department of Pathology, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, USA, [4]Department of Pathology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA, [5]Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Submitted: 20-Jan-2021 Revised: 29-Jun-2021 Accepted: 13-Jul-2021 Published: 20-Nov-2021

Abstract

Background: Human immunodeficiency virus (HIV) screening has improved significantly in the past decade as we have implemented tests that include antigen detection of p24. Incorporation of p24 detection narrows the window from 4 to 2 weeks between infection acquisition and ability to detect infection, reducing unintentional spread of HIV. The fourth- and fifth-generation HIV (HIV5G) screening tests in low prevalence populations have high numbers of false-positive screens and it is unclear if orthogonal testing improves diagnostic and public health outcomes. **Methods:** We used a cohort of 60,587 HIV5G screening tests with molecular and clinical correlates collected from 2016 to 2018 and applied machine learning to generate a classifier that could predict likely true and false positivity. **Results:** The best classification was achieved by using support vector machines and transformation of results with principle component analysis. The final classifier had an accuracy of 94% for correct classification of false-positive screens and an accuracy of 92% for classification of true-positive screens. **Conclusions:** Implementation of this classifier as a screening method for all HIV5G reactive screens allows for improved workflow with likely true positives reported immediately to reduce infection spread and initiate follow-up testing and treatment and likely false positives undergoing orthogonal testing utilizing the same specimen already drawn to reduce distress and follow-up visits. Application of machine learning to the clinical laboratory allows for workflow improvement and decision support to provide improved patient care and public health.

Keywords: Fifth-generation human immunodeficiency virus testing, human immunodeficiency virus, principal components analysis, serology, support vector machine

Introduction

Infection with the human immunodeficiency virus (HIV) continues to be a global concern affecting approximately 38 million people with approximately 1.7 million new cases annually.[1] In 2006, the CDC announced an initiative for voluntary, routine testing of all Americans aged 13–64 during health-care encounters. Despite the progress in the last decade since this announcement, approximately 55% of adults in the United States have never been tested. An estimated 15% of infected individuals are unaware of their status, which means that they can unwittingly infect others.[2] HIV is most infectious during initial infection stages when people are least likely to be aware of infection. Appropriate test algorithms and analytically sensitive tests that can identify infection early are needed to reduce HIV infection rates and improve outcomes.[3,4] The most analytically sensitive approved screening tests are fourth- and fifth-generation HIV screening tests which test for anti-HIV-1 antibodies (HIV-1 Ab), anti-HIV-2 antibodies (HIV-2 Ab), and HIV-1 antigen (HIV-1 Ag). Fourth-generation HIV (HIV4G) screens report one result without the ability to indicate which of the targets is positive, while fifth-generation screening (HIV5G) reports an index value (IDV) for each of the targets.[5,6] HIV5G testing does not yet have a CDC recommended testing algorithm.

Access this article online

Quick Response Code:

Website: www.jpathinformatics.org

DOI: 10.4103/jpi.jpi_7_21

Address for correspondence: Dr. Bradley J Wheeler, Department of Pathology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.
Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
Clinical Laboratory Building, 3477 Euler Way, Pittsburgh, PA 15213, USA.
E-mail: wheelerse3@upmc.edu

How to cite this article: Elkhadrawi M, Stevens BA, Wheeler BJ, Akcakaya M, Wheeler S. Machine learning classification of false-positive human immunodeficiency virus screening results. J Pathol Inform 2021;12:46.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2021/12/1/46/330787

The current CDC algorithm follows HIV4G-positive screening tests with an additional differentiation assay to determine if anti-HIV 1 or anti-HIV 2 is positive. That screening test is followed by nucleic acid testing if differentiation results are negative or indeterminate.[7] In theory, HIV5G testing eliminates the need for a differentiation assay. At present, there is only one FDA-approved differentiation assay and it generates complex results that have led the CDC to release a full-page table of possible results and recommended interpretative comments for this testing.[8] The complexity is difficult for both patients and physicians and it is unclear if a second screening assay is needed to provide orthogonal testing for all patient samples positive by HIV5G testing. In addition, there is debate if a differentiation assay is the appropriate 2nd step in testing for HIV with HIV4G when a quantitative viral load assay will provide clinically actionable information not provided by the differentiation assay.[9] At a practical level, HIV5G testing is ongoing in an increasing number of clinical laboratories, and data are limited on best practices and limitations of this testing.

To add complexity to the diagnostic process, positive screening results must be reported to local health authorities in many countries including the US. Despite the high sensitivity and specificity of HIV4G and HIV5G (>99%), false positives are still common in the low prevalence populations most clinical laboratories serve. In addition, false positives for serologic tests generally and HIV serologic tests specifically are increased among pregnant women and those with autoimmune diseases. Prenatal HIV screening is considered standard of care to allow for early intervention to protect the fetus, further complicating the screening assays. There are competing needs for accuracy compared to time to diagnosis. Accuracy can be improved by incorporating orthogonal testing or reflex testing to molecular methods for confirmation, but this delays diagnosis. At present, the CDC algorithm and clinical laboratories favor increased accuracy at the expense of delaying diagnoses. Most clinical laboratories require a separate specimen type or unopened specimen for molecular confirmation, which requires the patient to return for a second blood draw or requires the laboratory to hold a tube in reserve on all patients undergoing testing. Timely reporting of true positives is vital to reduce transmission and speed antiretroviral treatments as early treatment improves outcomes.[4] Alternately, reporting of a false-positive screen that is reported to public health authorities and requires additional testing follow-up to determine true negative status is psychologically and economically stressful, particularly for our prenatal screening patients. We wanted a rapid intermediate step that could allow us to treat likely positive screens differently than likely false-positive screens.

We hypothesized that machine learning could be used as an interim step after screening to determine the likelihood of a positive screen being true positive versus false positive. Using this classification, we could then follow separate workflows to balance accuracy and time to diagnosis [Figure 1]. A screen classified as likely true positive could be reported immediately to allow for the necessary notifications and clinical follow-up.

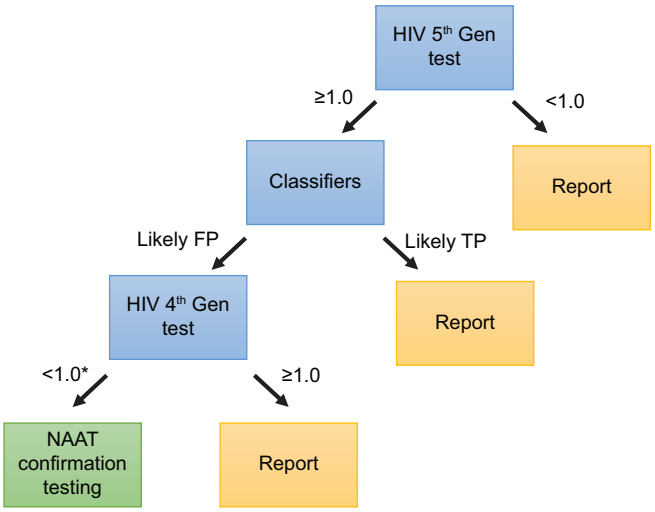

Figure 1: Proposed workflow with machine learning classifiers. The proposed workflow allows for differential handling of likely true-positive versus false-positive screens by HIV5G. *This testing can be resulted as "indeterminate" before human immunodeficiency virus nucleic acid amplification testing, clinician consultation can be initiated, or utilization of the same sample for nucleic acid amplification testing can be performed

A screen classified as likely false positive could be followed up with an orthogonal screening test, which allows for use of the same specimen already drawn. Positivity by the orthogonal test could be reported, while disagreement between the initial and orthogonal testing could be discussed with the physician and the patient brought back in for molecular confirmation.

Machine learning to date has had limited application in laboratory medicine and specifically in HIV diagnostic processes. Existing machine learning methods assess patient response to HIV infection and treatment using laboratory, demographic, and other electronic medical record (EMR) inputs.[10-18] There has been limited work using EMR demographic data to predict HIV status which could be used by health-care systems to proactively encourage screening.[19-21] One group has applied machine learning to laboratory testing, by using flow cytometry images of CD4+ cells to assess diagnosis and treatment of HIV,[18] providing valuable information, but not creating a decision support tool for the clinical laboratory.

Using support vector machines (SVM) and 60,587 patient screening results, we were able to create classifiers that appropriately classified 94% of false positives and 92% of true positives. The results from this work provide the first evidence of an effective HIV5G screening algorithm which will reduce unnecessary testing and interventions, decrease the time it takes patients to get an accurate diagnosis, and reduce HIV infection spread due to delayed test results.

Methods

Study design

A retrospective analysis of the laboratory information system (LIS) was performed by collecting all data generated

by the BioPlex 2200 HIV Ag-Ab assay platform (Bio-Rad Laboratories, Hercules, CA, USA; screening assay) between January 2, 2016, and December 31, 2018, at UPMC Clinical Laboratories (Pittsburgh, PA). This work was performed under the auspices of the University of Pittsburgh IRB study #20020103.

A total dataset of 60,587 individual HIV assays was collected. All nucleic acid amplification testing (NAAT) for HIV-1 run at UPMC was also collected from our LIS between January 2, 2016 and December 31, 2018. In addition, chart review of HIV screening assay positive samples was performed to assess for NAAT testing by reference laboratories, and screen assay positive cases without NAAT results were chart reviewed for clinical diagnosis of HIV. HIV NAAT-positive or negative testing was considered diagnostic if the patient was not on HIV treatment medications. NAAT and chart review combined to create a final disposition for each screen-positive patient of HIV positive, negative, or unknown. There were a total of 580 specimens that were assessed by both HIV5G and NAAT or screened positive by HIV5G. 475/580 had NAAT testing, 453/580 were HIV5G screen positive (representing all screen-positive specimens during this time period), 127/580 were HIV5G screen negative. 45/580 specimens were excluded as they either did not have all three HIV5G screen values or did not have polymerase chain reaction or a conclusive clinical diagnosis. To create these algorithms, we used random sampling of 25% of the total negative cases ($n = 59\,913$ total) to provide improved class balance due to the over representation of negative cases (i.e. low HIV prevalence).

BioPlex 2200 human immunodeficiency virus assay

The presence or absence of HIV viral antigens and antibodies were assessed using serum samples on the BioPlex 2200 HIV Ag-Ab assay. The BioPlex HIV Ag-Ab assay is a multiplex bead-based assay that allows for the detection of HIV-1 p24 antigen, anti-HIV-1 Ab, and anti-HIV-2 Ab. Index values are generated for each individual HIV analyte. If all analytes have indices <1.00, the specimen is interpreted as nonreactive with no additional testing. If at least one analyte is equal to or >1.00, repeat testing is performed in triplicate. If two of the three tests demonstrate no analytes with index values ≥1.00, then the specimen is interpreted as nonreactive. If repeat testing demonstrates two repeats having an index value equal to or >1.00 for the same analyte, then the specimen is interpreted as reactive for the specific analyte. We refer to the interpretation of the assay result as the "assay disposition." If a specimen is reactive for both HIV-1 Ab and HIV-2 Ab, and the indices have less than a 5-fold difference, the interpretation is reactive, undifferentiated for the specimen.

Nucleic acid amplification testing human immunodeficiency virus-1 RNA assay

For patients considered reactive on the BioPlex 2200 HIV Ag-Ab assay, the viral load of HIV RNA was assessed in-house using the Abbott RealTime HIV-1 assay on the m2000 system (Abbott Laboratories, Chicago, IL, USA) or sent to commercial reference laboratories. The linear range of the in-house assay is 40 copies/mL (1.6 log copies/mL) to 10 million copies/mL (7.0 copies/mL).

Retrospective chart analysis

A subset of the data, which included all individual HIV assays with at least one reactive analyte, was identified for the final HIV clinical diagnosis identification (we refer to this as clinical disposition). An HIV assay is considered a true positive in the presence of (1) any history of HIV viral load demonstrating detection of circulating HIV viral copies and/or (2) any history of treatment with HIV antiretroviral therapies. An HIV assay is considered a true negative in the presence of (1) a follow-up HIV viral load demonstrating no detection of circulating HIV viral copies in the absence of antiretroviral therapy or (2) in the absence of a follow-up viral load, a follow-up nonreactive fourth or fifth-generation HIV screening assay. In the context of the study, HIV preexposure prophylaxis identified by chart review is not considered antiretroviral therapy for final disposition categorization. If an individual sample did not meet the above criteria, it was categorized as unknown. Unknown samples were excluded from this analysis.

Classification for the identification of false assay positives

As mentioned above, in this study, we develop an ML-based method to distinguish between true and false assay positives. Our method uses the three assay results from HIV5G (HIV-1 Ab, HIV-1 Ag, and HIV-2 Ab) and assay and clinical dispositions (as defined above) in the following 3-step approach. Our method is implemented on MATLAB R2019b.

Step 1: In the 1st step, we develop two different SVM-based HIV-positive versus HIV-negative classifiers for feature extraction using (i) clinical and (ii) assay dispositions as the ground truth. We apply 10-fold cross validation to learn the parameters of the radial basis function (RBF) kernel, we utilize for SVM classification. Furthermore, during the cross validation, for each fold for each classification problem, we learn the SVM scores. The SVM scores from each classifier for each assay result are then concatenated to form a two-dimensional feature vector to be used in the next steps of our method. In addition, for the clinical classifier and the assay classifier, we also report the confusion matrix and present the accuracy, specificity, and sensitivity.

Step 2: In the 2nd step, we apply principal component analysis (PCA) using all the two-dimensional features extracted from Step 1 and project these samples on the principal components (PCs). This transforms the two-dimensional features from Step 1 and results in two orthogonal features PCs for each positive assay result.

Step 3: In the 3rd step, we use the PC transformed features for the positive assay results together with the clinical disposition in an SVM-based classifier to distinguish between true-positive and false-positive assay results. Here, clinical dispositions are used as the ground truth to develop an SVM-based classifier with RBF kernel through 10-fold cross validation. In addition,

in this step, we consider feature selection. Specifically, during 10-fold cross validation, we consider a cost function that is a weighted summation of the true and false-positive accuracies and learn the optimum weight for each PC transformed feature separately. Accordingly, we choose one of the PCs from Step 2 as the optimum feature to be used for false-positive vs. true-positive assay classification. In our cost function, we define:

True-positive accuracy = (correctly identified true positives)/(total number of true positives)

False-positive accuracy = (correctly identified false positives)/(total number of false positive*s)*

Results

A positive HIV screening assay should be followed by molecular confirmation of HIV infection; however, the initial screen result is psychologically stressful and is reported to public health agencies. Therefore, false-positive screen results carry consequences that are best mitigated where possible. We considered that it may be possible to create a classification algorithm that could appropriately determine false positive from true-positive screening results, allowing for alternate workflows on possible false versus true-positive screens.

Our current rules-based approach uses a cutoff of ≥1.00 for reactivity, using this cutoff to assess all cases with a clinical disposition ($n = 535$), we found a baseline accuracy of 73.5%, and a specificity of 47.2% [Figure 2a]. This subset of cases was chosen as all screen-positive cases and any additional cases that had both HIV5G and NAAT, as such the sensitivity is expected to be 100% [Figure 2a]. We began with a SVM classifier that was trained on negative cases and all cases with clinical disposition. This clinical classifier utilized the clinical disposition as the ground truth and found reasonable performance with 119 false positives correctly classified as false positive [Figure 2a and b] when using the classifier to assess cases with clinical disposition. We considered if there could be added value from assessing the assay disposition as ground truth and found poor detection of screen false-positive results with only 9 of 142 false positives detected [Figure 2a and c].

We graphed the SVM scores for assay-positive cases ($n = 408$) and the score distribution to assess for further information [Figure 3]. The clinical classifier demonstrated that most true positives had a score <0 and most correctly predicted false positives had a score >0 as demonstrated by the distribution probability [Figure 3b]. The false positives that were missed by the clinical classifier had a score <0 which overlapped with the true positives. The assay classifier demonstrated that most true-positive results had a score between −1 and −1.5 as shown by the distribution probability [Figure 3a], and all false positives that were caught by the assay classifier had a score of >0. These results indicated that both the assays and clinical classifiers appeared to have several areas of further separation possible to improve the false-positive classification with a data transformation [Figure 3].

We used PCA to transform the SVM scores from both the assay and clinical classifiers for assay-positive cases and found when PC1 versus PC2 were graphed against each other that there appeared to be reasonable separation by PC1 [Figure 4]. Assessment of each of the PCs using an accuracy evolution plot determined the optimal SVM classifier through the selection of hyperparameter C that balances between true and false-positive accuracies in the classification optimization cost function [Figure 4c]. That is, we used the hyperparameter C which weights classification of false-positive versus true-positive accuracy. When C is equal to 1, correct classification of true positive and correct classification of false positive are equally weighted. By shifting this parameter, we can optimize correct classification of false-positive cases. We created an accuracy evolution plot to determine the optimal value of C to maximize the correct identification of false positives and found the optimal value of C to be 1.5 for PC1 and 4 for PC2.

Using PC1 and PC2 each as an input for new classifiers with the clinical disposition of screen-positive cases as the ground truth, we were able to optimize accuracy. The PC1 classifier provided significantly more accurate results than the PC2 classifier [92.6% vs. 84.6%, respectively, Figure 5]. To ensure the accuracy, sensitivity, and specificity did not vary significantly between iterations of the algorithm, we assessed 10 repetitions of 10-fold cross validation and assessed the average and standard deviation for these. We found the standard deviations to be small, indicating high reproducibility and high generalizability [Figure 5].

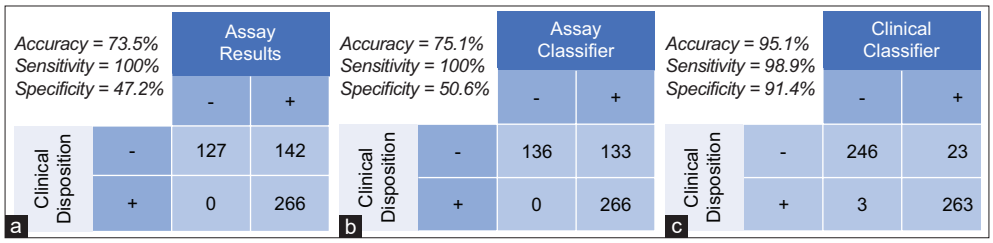

Figure 2: Confusion matrices for initial support vector machines. (a) Confusion matrix for all samples screened by HIV5G with available clinical disposition through nucleic acid amplification testing or chart review. (b) Confusion matrix for the performance of the assay classifier compared to clinical disposition. The assay classifier uses the assay disposition as the ground truth. (c) Confusion matrix for the performance of the clinical classifier compared to clinical disposition. The clinical classifier uses the clinical disposition as the ground truth

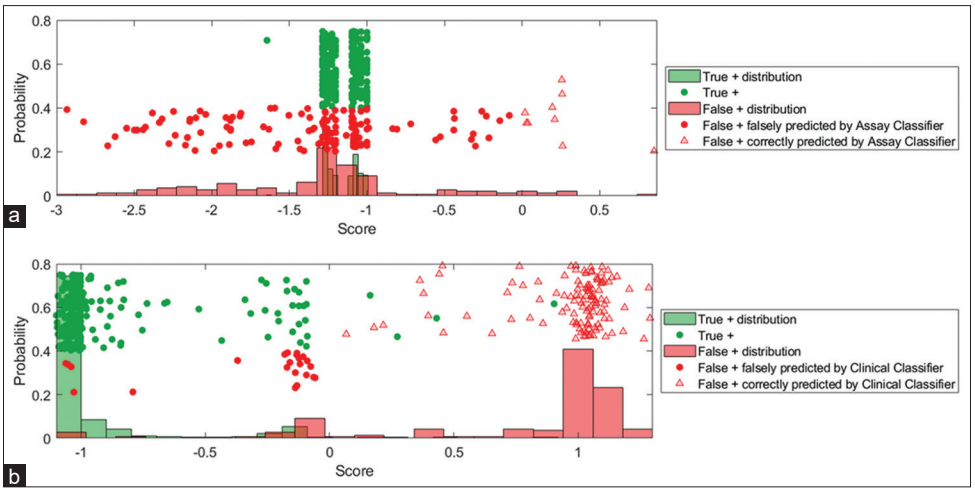

Figure 3: Normalized histograms of support vector machines scores of assay-positive cases assuming (a) assay and (b) clinical dispositions as the ground truths, respectively. Green and red bars demonstrate the true and false human immunodeficiency virus-positive cases, respectively. Scores for each assay-positive sample are also illustrated as a scatter plot. Samples corresponding to true and false human immunodeficiency virus positives correctly identified by each classifier are demonstrated by green circles and red triangles, respectively. Samples corresponding to false human immunodeficiency virus positives falsely identified as positive by each classifier are shown by red circles

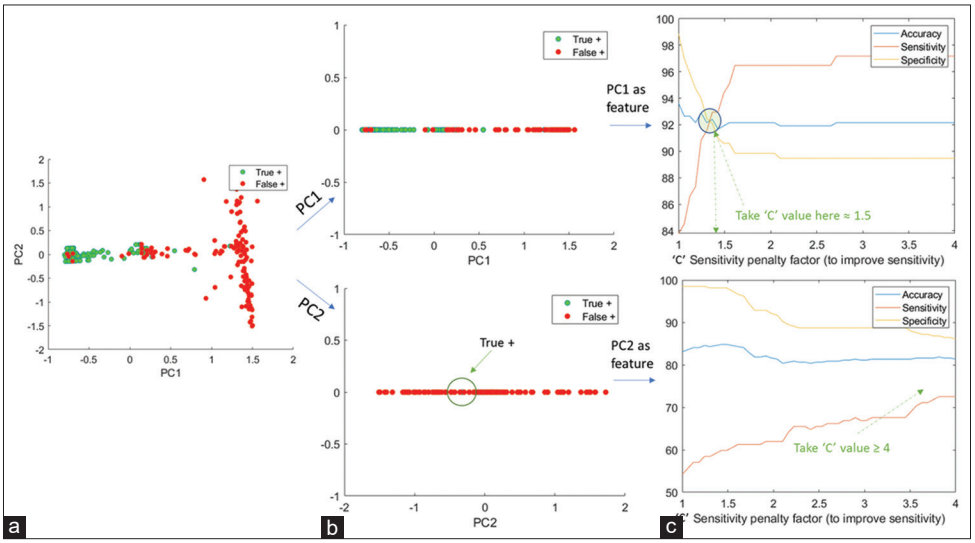

Figure 4: Principal component analysis and classification performance. (a) Joint and (b) individual scatter plots of principal components corresponding to the support vector machines scores obtained in the 2nd step of the classification method. Red and green circles represent the principal component values for true and false assay-positive samples. (c) Optimization of the true versus false-positive classification by selecting the sensitivity penalty factor using the (top) first and (bottom) second principal component as the feature in the classifier. Propagation of classification accuracy, specificity, and sensitivity as a function of the sensitivity penalty factor are shown in each figure

Conclusions

Here, we demonstrate the usefulness of machine learning to aid in triaging positive HIV screens for further assessment. We had observed that in our low prevalence population, approximately one-third of HIV5G-positive screen results confirmed by NAAT or clinical diagnosis as negative, highlighting the limitations of this rules-based approach. We had also observed that many of these screens appeared to have interference as there would be results >1.0 for HIV-1 Ab, HIV-1 Ag, and HIV-2 Ab or had low index values. We considered that machine learning classification systems could aid us in creating a workflow that would maximize turnaround times for true-positive cases to prevent further unknowing HIV spread and reduce the false-positive cases reported by appropriately reflexing testing.

The current CDC algorithms for testing do not yet include HIV5G which, unlike HIV4G, allows the laboratory to screen and differentiate which analyte is positive (HIV-1 Ag, HIV-1 Ab, HIV-2 Ab), potentially eliminating the need for a separate differentiation step. It remains unclear if an orthogonal screening test is necessary for all positive HIV5G screens. Given that, regardless of the result of the differentiation

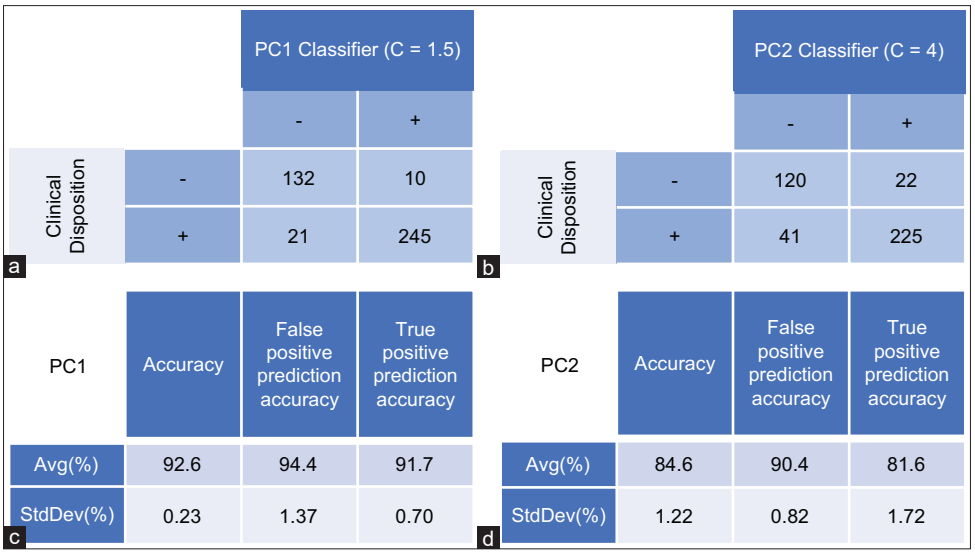

Figure 5: Classification performance of the PC1 and PC2 classifiers. Human immunodeficiency virus true versus false-positive classification confusion matrices based on (a) PC1 and (b) PC2 using assay-positive cases and assuming clinical disposition as the ground truth. Average values and standard deviation of the overall classification accuracy, and true and false-positive prediction accuracies obtained through 10 repetitions of 10-fold cross validation using (c) PC1 and (d) PC2 for the human immunodeficiency virus true versus false-positive classification

assay, an HIV5G screen with any analyte ≥1.0 will need to be confirmed with NAAT either to resolve the discrepancy or a viral load result will be needed to baseline therapy, it seems the orthogonal testing may be simply a delay in patient diagnosis and additional laboratory cost.[9]

Our institution requires a different sample type and an unopened tube for HIV nucleic acid testing compared to screening, and many institutions send out HIV NAAT. Therefore, creating an algorithm that optimizes true-positive screening results with minimal delay and reduced need for patients to return for testing is desirable. Conventionally, to establish appropriate reference intervals, we use numeric thresholds as cutoffs based on Gaussian statistical models, most often in the form of a 95% or 99% healthy population threshold. When sufficient healthy samples are not available, expert evaluation and biological basis of disease can also be used to infer a reasonable reference interval. However, these methods by definition eliminate outliers, which are often the cases of most import. Utilizing machine learning allows us to create models that can assess both normal and outlier cases simultaneously. Machine learning allows us to extract additional information above a single numeric value in isolation and assess them against a higher dimensional space classification which includes the full historical data. Data-driven decision rules reduce the limitations of existing decision rules which are based on a limited cohort or specialist knowledge. This reduces potential human error, cohort biases, and accounts for shifts in reagents and patient population when data are used in an ongoing data stream. We can use these algorithms to flag results that may be problematic, highlight instrument errors, or assess trends in assays as tools to support decision-making among clinical experts.[22-24]

This current algorithm serves as a proof of principle for improved HIV screening, and an additional assessment on inter-institutional data is desirable. While we have a large multi-year dataset the low prevalence of HIV in our population, while key to the sensitivity issues we experience, also limits our sample size. In machine learning, imbalanced classes are handled with corrections, but due to the low threshold for risk in both medicine and HIV testing, independent assessment of this algorithm in a similar or larger data set is required.

Continued work to refine this classifier and full testing algorithm include recruiting additional datasets and exploration of other laboratory tests where machine learning can serve as an aid to workflow and provide decision support. Our classifier provides robust prediction of false-positive test results for HIV5G testing. Future work should also assess the incorporation of other available laboratory tests to further improve the robustness of this classifier.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References
1. U.S. Department of Health and Human Services. HIVgov. Available from: https://www.hiv.gov/. [Last accessedon 2020 Nov].
2. Branson BM, Handsfield HH, Lampe MA, Janssen RS, Taylor AW, Lyss SB, *et al*. Revised recommendations for HIV screening of pregnant women. MMWR Morb Mortal Wkly Rep 2001;50:1-62.
3. Centers for Disease Control and Prevention. Vol. 27. HIV Surveillance Report; 2015. Available from: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html. [Last accessed on 2020 Nov].
4. INSIGHT START Study Group, Lundgren JD, Babiker AG, Gordin F, Emery S, Grund B, *et al*. Initiation of antiretroviral therapy in early asymptomatic HIV infection. N Engl J Med 2015;373:795-807.
5. Guarner J. Human immunodeficiency virus: Diagnostic approach. Semin Diagn Pathol 2017;34:318-24.

6. Salmona M, Delarue S, Delaugerre C, Simon F, Maylin S. Clinical evaluation of BioPlex2200 HIV Ag-Ab, an automated screening method providing discrete detection of HIV-1 p24 antigen, HIV-1 antibody, and HIV-2 antibody. J Clin Microbiol 2014;52:103-7.
7. Centers for Disease Control and Prevention. Laboratory Testing for the Diagnosis of HIV Infection: Updated Recommendations. Available from: http://dx.doi.org/10.15620/cdc. 23447 [Last accessed on 2020 Nov].
8. Centers for Disease Control and Prevention. Suggested Reporting Language for the HIV Laboratory Diagnostic Testing Algorithm. Available from: https://npin.cdc.gov/publication/suggested-reporting-language-hiv-laboratory-diagnostic-testing-algorithm. [Last accessed on 2020 Nov].
9. Branson BM. HIV diagnostics: Current recommendations and opportunities for improvement. Infect Dis Clin North Am 2019;33:611-28.
10. Altmann A, Däumer M, Beerenwinkel N, Peres Y, Schülter E, Büch J, *et al.* Predicting the response to combination antiretroviral therapy: Retrospective validation of geno2pheno-THEO on a large clinical database. J Infect Dis 2009;199:999-1006.
11. Altmann A, Rosen-Zvi M, Prosperi M, Aharoni E, Neuvirth H, Schülter E, *et al.* Comparison of classifier fusion methods for predicting response to anti HIV-1 therapy. PLoS One 2008;3:e3470.
12. Bisaso KR, Karungi SA, Kiragga A, Mukonzo JK, Castelnuovo B. A comparative study of logistic regression based machine learning techniques for prediction of early virological suppression in antiretroviral initiating HIV patients. BMC Med Inform Decis Mak 2018;18:77.
13. Paul RH, Cho KS, Luckett P, Strain JF, Belden AC, Bolzenius JD, *et al.* Machine learning analysis reveals novel neuroimaging and clinical signatures of frailty in HIV. J Acquir Immune Defic Syndr 2020;84:414-21.
14. Kwak NK, Lee C. A neural network application to classification of health status of HIV/AIDS patients. J Med Syst 1997;21:87-97.
15. Lee JS, Paintsil E, Gopalakrishnan V, Ghebremichael M. A comparison of machine learning techniques for classification of HIV patients with antiretroviral therapy-induced mitochondrial toxicity from those without mitochondrial toxicity. BMC Med Res Methodol 2019;19:216.
16. Larder B, Wang D, Revell A. Application of artificial neural networks for decision support in medicine. Methods Mol Biol 2008;458:123-36.
17. Wang D, Larder B. Enhanced prediction of lopinavir resistance from genotype by use of artificial neural networks. J Infect Dis 2003;188:653-60.
18. Singh Y, Mars M. Support vector machines to forecast changes in CD 4 count of HIV-1 positive patients. Sci Res Essays 2010;5:2384-90.
19. Betechuoh BL, Marwala T, Tettey T. Autoencoder networks for HIV classification. Current Science 2006;91:1467-73.
20. Xiang Y, Fujimoto K, Schneider J, Jia Y, Zhi D, Tao C. Network context matters: Graph convolutional network model over social networks improves the detection of unknown HIV infections among young men who have sex with men. J Am Med Inform Assoc 2019;26:1263-71.
21. Ahlström MG, Ronit A, Omland LH, Vedel S, Obel N. Algorithmic prediction of HIV status using nation-wide electronic registry data. EClinicalMedicine 2019;17:100203.
22. Sourati J, Erdogmus D, Akcakaya M, Kazmierczak SC, Leen TK. A Novel Delta Check Method for Detecting Laboratory Errors. Boston, MA: 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP); 2015. p. 1-6.
23. Sourati J, Kazmierczak SC, Akcakaya M, Dy JG, Leen TK, Erdogmus D. Assessing subsets of analytes in context of detecting laboratory errors. Annu Int Conf IEEE Eng Med Biol Soc 2016;2016:5793-6.
24. Higger M, Akcakaya M, Erdogmus D. A robust fusion algorithm for sensor failure. IEEE Signal Process Lett 2013;20:755-8.