Assessing the Performance of Three HIV Incidence Risk Scores in a Cohort of Black and White MSM in the South

Jeb Jones, PhD, MPH, MS1, Martin Hoenigl, MD2,3, Aaron Siegler, PhD, MHS1, Patrick Sullivan, DVM, PhD1, Susan Little, MD2, and Eli Rosenberg, PhD1

1Department of Epidemiology, Emory University, Atlanta, GA
2Division of Infectious Diseases, University of California, San Diego, San Diego, CA
3Division of Pulmonology and Section of Infectious Diseases, Medical University of Graz, Graz, Austria

Abstract

Background—Risk scores have been developed to identify men at high risk of HIV seroconversion. These scores can be used to more efficiently allocate public health prevention resources, such as pre-exposure prophylaxis. However, the published scores were developed with datasets that comprise predominantly white men who have sex with men (MSM) collected several years prior and recruited from a limited geographic area. Thus, it is unclear how well these scores perform in men of different races or ethnicities or men in different geographic regions.

Methods—We assessed the predictive ability of three published scores to predict HIV seroconversion in a cohort of black and white MSM in Atlanta, GA. Questionnaire data from the baseline study visit were used to derive individual scores for each participant. We assessed the discriminatory ability of each risk score to predict HIV seroconversion over two years of follow-up.

Results—The predictive ability of each score was low among all MSM and lower among black men compared to white men. Each score had lower sensitivity to predict seroconversion among black MSM compared to white MSM and low area under the curve values for the ROC curve indicating poor discriminatory ability.

Conclusions—Reliance on the currently available risk scores will result in misclassification of high proportions of MSM, especially black MSM, in terms of HIV risk, leading to missed opportunities for HIV prevention services.

Keywords

Men who have sex with men; HIV risk; risk scores

Corresponding author: Jeb Jones, PhD, MPH, MS, Department of Epidemiology, Emory University, 1518 Clifton Road, Atlanta, GA 30322, jeb.jones@emory.edu, Phone: 404-421-2928, Fax: 404-727-8737.

The authors have no conflicts of interest to declare.
Introduction

The HIV epidemic in the United States has disproportionately impacted men who have sex with men (MSM). Currently, MSM account for close to two thirds of new diagnoses in the U.S. annually (1). However, MSM do not represent a homogenous group with regard to HIV risk. Black MSM experience the highest rate of HIV infection of any risk group in the U.S., with a rate of diagnosis that is 6 times higher than that of white MSM and 2 times higher than Hispanic MSM (2). The Involve[men]t study is a recently completed cohort study of black and white MSM in Atlanta, GA (3, 4). Following two years of follow-up, the cumulative incidence of HIV among black MSM was 9.2% compared to 2.7% among white MSM(4). Previous self-report research indicates that higher frequencies of risk behaviors are not responsible for the increased rates of infection among black MSM (5, 6). Rather, black MSM tend to report similar or lower frequencies of HIV risk behaviors compared to MSM from other racial or ethnic backgrounds(5–7), suggesting that other factors are responsible for the racial disparities in HIV incidence.

Pre-exposure prophylaxis (PrEP) with daily tenofovir and emtricitabine (TDF/FTC) has been proven safe and effective in preventing HIV acquisition among MSM(8). In persons adherent to PrEP, this treatment has been shown to reduce the risk of HIV seroconversion by over 90%(8, 9). However, PrEP is also a costly intervention, estimated to cost $10,000 or more per year per individual on PrEP(10). Efficient allocation of PrEP to those individuals at highest risk is therefore paramount. Over recent years, risk indices have been developed to identify candidates at highest risk of HIV for targeting prevention services and PrEP prioritization(11–13). MSM-specific risk scores include the Menza score(13), the HIV Incidence Risk Index for MSM (HIRI-MSM)(11), and more recently the San Diego Early Test (SDET) score(12). Both the Menza score and the HIRI-MSM score were developed based on behavioral risk data collected from 1999–2003 in a clinical trial population that may not accurately represent the behavioral risks associated with HIV acquisition risk in a real-world setting over a decade later(11, 13). Also, use of methamphetamine or inhaled nitrites is weighted in both scores, while other drugs are not(11, 13), which may restrict the use of the score to settings where methamphetamines and inhaled nitrites are drivers of the HIV epidemic, but not other priority areas, such as the US South(14, 15). The SDET score was developed more recently and has an emphasis on sexual risk variables directly associated with HIV acquisition among MSM(12) and thus may be more broadly applicable to different MSM populations (since sexual risk behavior associated with substance use will still be captured(16)). However, the SDET score was developed and validated using data from a cohort of mostly white MSM and those with Hispanic ethnicity in southern California. The score may therefore be less applicable to other geographic settings or in areas where African Americans represent a substantial portion of the population at risk for HIV infection.

Each of these three HIV risk scores has a published criterion level that is recommended to indicate high risk for HIV acquisition to be used in clinical and community settings to screen men seeking HIV prevention services. The PrEP Guidelines released by the Centers for Disease Control and Prevention and the United States Public Health Service in 2014 recommend that providers use HIRI-MSM as a tool to identify potential candidates for
However, the clinical utility of each of these scores has not been demonstrated in racially diverse populations, nor have the scores been directly compared to assess their relative performance in predicting HIV seroconversion, particularly among black MSM, the most at-risk subgroup of HIV risk in the U.S. It is important to understand how well each of these three risk scores will perform in populations and contexts beyond those in which they were developed.

We examined the sensitivity, specificity, and area under the curve (AUC) of each score. Sensitivity provides an indication of how well the scores identify high-risk men, specificity provides an indication of how many low risk men the scores incorrectly classify as high risk, and the AUC provides a summary measure that combines the two. The objective of this study was to assess the performance of the three different HIV risk scores for predicting HIV seroconversion in the Involve[men]t cohort.

**MATERIALS AND METHODS**

**Involve[men]t Study Population**

The Involve[men]t study was a cohort study of black and white MSM in Atlanta, Georgia from July 2010–March 2014. The study methods have been described previously(3, 4). Briefly, HIV-negative, non-Hispanic, black and white MSM were enrolled in a prospective cohort study and followed for two years or until seroconversion. Participants were recruited using venue-based time-space sampling and via Facebook advertisements(18). At baseline and every 6 months thereafter (for a maximum of 24 months) participants completed extensive questionnaires on sexual risk behavior and drug use and were tested for HIV and other sexually transmitted infections (STIs). To reduce participant burden, a short subset of sexual behavior questions was asked if a new partner was not reported in a given follow-up study visit. To identify participants who were acutely infected at baseline, HIV testing was repeated after 3 months (no questionnaires were filled out during this brief study visit) and quantitative viral load testing was performed on frozen blood samples provided at baseline by men with a positive test result.

**Risk Score Calculations**

Data from the baseline questionnaire were used to calculate risk indices. The ability of each score to predict seroconversion during the full 2-year study period was determined.

The variables used to calculate each score, necessary modifications that were made based on the data available from the Involve[men]t study, and the cutoff score that is recommended to indicate high risk for HIV seroconversion for each risk score are displayed in Table 1. For each risk score, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for HIV seroconversion were calculated using the recommended cutoff. These values were calculated using baseline scores for all participants to predict seroconversion after two years (representing the full follow-up period). All analyses were conducted for the whole study cohort and in addition stratified by race. Chi-square tests were conducted to assess whether there were differences in classification as high risk by race. To assess the overall predictive ability of the scores, receiver operating characteristic
(ROC) curve analysis was conducted for all three indices and area under the curve (AUC) values (including 95% confidence intervals) determined. Analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC) and SPSS version 23 (SPSS Inc., Chicago, IL).

**Sensitivity Analyses**

We conducted three sensitivity analyses, which are presented in detail in the Appendix.

First, we examined the positive and negative predictive values of each risk score under a range of prevalence values. Whereas sensitivity and specificity are properties of the risk scores themselves, the utility of the predictive values is also a function of underlying prevalence. Therefore, we explored how each score would perform in contexts different than the Atlanta MSA.

Second, each score was tested and validated to predict six-month seroconversion. Therefore, we also conducted the main analyses with an outcome of seroconversion within six months of baseline.

Third, we used longitudinal data to update the risk score for each participant at follow-up study visits based on their survey responses at that visit. As described above, participants only completed the full battery of sexual behavior questions if a new partner was reported at a given visit. Thus, we were only able to update risk scores for participants reporting a new partner at a given visit.

**RESULTS**

**Involve[m]ent Cohort**

A total of 562 (260 black, 302 white) MSM were enrolled and followed prospectively. The average age of participants was 27 years (sd: 6.7). Including 6 men who were determined to be acutely infected with HIV at baseline, there were a total of 32 seroconversions observed – 24 among black men and 8 among white men. In addition to the acute infections, 5 seroconversions were detected at the month 3 visit, 3 at month 6, 3 at month 12, 8 at month 18, and 7 at month 24.

**Risk Score Performance at Published Cutoff Values**

The overall sensitivity of HIRI-MSM to predict HIV was 63% (95% CI: 44–79%) for the entire two-year study period (Table 2). Sensitivity was differential by race. Among white participants, baseline scores predicted 75% (95% CI: 35–97%) of seroconversions that occurred over the full study period. Sensitivity among black MSM was 58% (95% CI: 37–78%) for the full study period. Specificity of HIRI-MSM was 57% (95% CI: 52–61%) over the two-year follow-up. Among black MSM, specificity was 66% (95% CI: 60–72%); among white MSM specificity was 49% (95% CI: 43–55%).

The sensitivity of the Menza score was similar to HIRI-MSM. Menza scores also performed differently for black and white MSM. The Menza score had 88% (95% CI: 47–100%) sensitivity to predict seroconversion over the 2-year study among white MSM. Among black men, the Menza score had a sensitivity of 54% (95% CI: 33–74%) over the two-year follow-
Specificity was markedly lower than that observed for HIRI-MSM with a 41% (95% CI: 37–46%) true negative rate. Specificity was 42% (95% CI: 35–48%) among black MSM and 41% (95% CI: 35–47%) among white MSM, indicating that unlike sensitivity, the specificity of the Menza score did not differ by race.

The sensitivity of the SDET score was markedly lower than either HIRI-MSM or the Menza score. Among all participants, the SDET predicted 25% (95% CI: 12–43%) of seroconversions over the 2-year follow-up. Similar to both of the other scores, sensitivity of the SDET differed by race. Sensitivity was 50% (95% CI: 16–84%) over two years of follow-up for white MSM and 17% (95% CI: 5–37%) among black MSM. The specificity of the SDET score was higher than either HIRI-MSM or Menza. Overall, specificity of the SDET score was 84% (95% CI: 81–87%). Specificity was slightly higher among black MSM (89%, 95% CI: 84–92%) than among white MSM (80%, 95% CI: 75–85%).

Reflecting the overall incidence observed in the study population, positive predictive values were low and negative predictive values were high for each of the three risk scores, regardless of race.

Results of the chi-square analyses indicate that white MSM were more likely than black MSM to receive a score indicating high risk using HIRI-MSM, indicating potential PrEP eligibility (p < .05), despite the higher risk of HIV among black MSM. No racial differences were observed in assigned scores using the Menza model or SDET.

**Overall Performance of Risk Scores**

The overall performance of each score was determined using ROC curves. In general, higher AUC values were observed for white MSM compared to black MSM for each of the scores (Table 3). However, AUC was low for both groups.

The AUC for seroconversion over the full two years of follow-up was 0.62 (95% CI: 0.52–0.72) for HIRI, 0.51 (95% CI: 0.41–0.60) for Menza, and 0.55 (95% CI: 0.44–0.66) for SDET. AUC was higher among white MSM for HIRI (AUC = 0.67, 95% CI: 0.47–0.88), Menza (AUC = 0.60, 95% CI: 0.44–0.75), and SDET (AUC = 0.66, 95% CI: 0.46–0.87) than for black MSM (HIRI AUC = 0.63, 95% CI: 0.51–0.75, Menza AUC = 0.49, 95% CI: 0.36–0.62, SDET AUC = 0.52, 95% CI: 0.39–0.65).

**Sensitivity analyses**

We conducted three sensitivity analyses to assess the predictive ability of the three risk scores under different conditions. The results of each of these analyses is described below and presented in more detail in the supplementary appendix.

Because each of these three scores was derived with an outcome of six-month seroconversion we also assessed their performance to predict seroconversions within six months of the baseline visit. The results of these analyses were consistent with the results for two-year incidence; however, given the smaller number of events, there was a substantial loss of precision.
We also conducted a longitudinal analysis using the data from all study visits to update participants’ risk scores for each study visit. Risk scores could only be updated if participants reported a new partner at a given visit and therefore completed a full sexual behavior questionnaire. In crude models, none of the scores was statistically significantly associated with HIV seroconversion.

Finally, we evaluated the predictive value for each score under differing prevailing HIV prevalence. PPV and NPV are dependent on sensitivity, specificity, and prevalence. Thus, we examined the effect of different prevalence on the estimated predictive values. We found that, as HIV prevalence increased, PPV increased and NPV decreased.

DISCUSSION

We examined the discriminatory abilities of three different risk indices designed to predict HIV seroconversion in MSM: HIRI-MSM, the Menza model, and the SDET. All three risk indices were developed using data from populations that were majority white MSM, whereas the HIV epidemic in the United States is marked by disproportionate rates of HIV diagnoses among black MSM(2). Thus, we used a prospective cohort of black and white MSM from Atlanta, GA to assess the diagnostic performance of each score among MSM in the southeastern U.S. Further, we evaluated whether diagnostic performance was differential by race. Diagnostic performance was poor overall; however, the overall results should be interpreted with caution given the differences in performance that were observed across the race-specific strata. Sensitivity was substantially lower for black MSM compared to white MSM on all three scores, but specificity was lower among white MSM compared to black MSM. AUC values were poor for both black and white MSM for all three scores.

All three risk indices use behavioral inputs of risk behavior, such as unprotected sex and drug use, known to be high-risk behaviors for HIV transmission. Previous studies have consistently found that black MSM report similar or lower rates of risk behavior compared to white MSM(6) while continuing to experience dramatically higher rates of HIV incidence. This suggests that a criterion value indicating high HIV risk from a risk score primarily developed and validated using data from white MSM may not be applicable to black MSM. Indeed, all of the risk indices examined in this study demonstrated lower sensitivity to predict seroconversion in the subsequent two-year period for black MSM compared to white MSM. In the primary analyses in the present study we did not evaluate race as a predictor because the goal of the study was to evaluate the predictive performance of the three risk scores, none of which includes race as a criterion. However, race was a stronger indicator of HIV risk than any of the risk scores in repeated measures models using risk scores updated from each study visit with an outcome of six-month seroconversion (see Appendix). Race is an important indicator for HIV risk among MSM, as is evident in the disproportionate rates of diagnosis comparing black and white MSM.(2) The prevalence in the different sexual networks of black and white MSM likely accounts for some of the differences in HIV risk.

The results of this study indicate that none of these three scores is adequately calibrated to predict HIV risk among MSM. The risk scores provided slightly better discrimination based
on AUC among white compared to black MSM; however, discrimination was poor in both
groups. The small number of seroconversions observed, particularly among white MSM,
restricted our ability to estimate sensitivity and AUC with a high level of precision. Due to
the high number of non-events, specificity did not suffer from this problem.

It should be noted that 6 of the 32 seroconversions observed were actually acutely infected at
the baseline visit. Thus, in these cases the risk assessment followed HIV infection. However,
given that these men were unaware of their HIV infection, there is no reason to expect that
there risk level had recently changed any more in comparison to a participant who was not
acutely infected. That is, we expect the effect of acute infection to be non-differential and do
not hypothesize that these men affected the results of the current study.

The poor discriminative ability of each score could be attributed to the race of the
participants, the geographic differences in the populations used to derive the scores
compared to the Involve[men]t cohort, or both. The HIV epidemic among MSM in the South
is characterized by different risk factors and different prevalence estimates compared to
other regions in the U.S.(19). Alternative (i.e., lower) cutoff values for the scores might
improve their performance among black MSM; however, the AUC values from ROC
analyses of each score indicate performance only slightly above chance for each score,
suggesting that there might not be an ideal value that will provide a suitable balance between
sensitivity and specificity. The similar AUC values for the three scores, in contrast to the
disparate sensitivities and specificities across scores, indicate that the proposed cutoff values
are driving much of the differences observed for the three scores. That is, the observed
sensitivity of the SDET score might have been more comparable to HIRI-MSM and Menza
if a different threshold of high risk were used. Future studies should investigate whether
different cutoffs or other modifications (e.g., incorporating underlying population HIV
prevalence based on geographic location) to these scales will improve their performance in
this population.

The present analysis suggests that each of these risk scores is not generalizable to
populations with differing distribution of race or from different geographical locations
compared to those that were used for development and validation of each score. This implies
a fundamental mismatch between the risk scores and the population that is most
disproportionately affected by HIV, black MSM.

The results of this study highlight the interplay between sensitivity and specificity. Although
sensitivity was lower for all three scores for black MSM compared to white MSM, the
opposite was true for specificity. Thus, these results indicate that use of these risk scores
would result in under-identification of high risk black MSM and over-identification of high
risk white MSM.

PrEP is a highly efficacious yet expensive HIV prevention intervention that will need to be
brought to scale in order to have a meaningful impact on HIV incidence in the United States,
with coverage levels much higher than have currently been achieved(20). Efficient
identification of potential candidates for PrEP is a necessary component of bringing this
intervention to scale; however, if the tools that are available to identify high risk MSM
perform poorly then there will likely be a large number of missed opportunities to identify men for whom PrEP might be a beneficial HIV prevention strategy. This could be particularly problematic for black MSM given the higher rates of HIV diagnosis in black compared to white MSM. A PrEP continuum of care has been proposed (21) that proposes several factors that affect PrEP uptake and potential interventions to increase uptake of PrEP. These include recommendations that healthcare providers screen for risk and determine patient eligibility for PrEP. Currently, if providers were to rely on the published risk scores then a large proportion of high risk black MSM would be incorrectly classified as at low risk for HIV seroconversion.

In addition to the issues with sensitivity that affect the scores’ ability to identify men at highest risk of HIV seroconversion, specificity might play an important role in PrEP prescription as well. Although modeling studies suggest providing PrEP at high levels in an epidemic context such as Atlanta’s is cost-effective (10, 22), it will be necessary to continue to monitor individuals’ risk profiles while on PrEP. Individual risk is likely to fluctuate over time, such that PrEP might become unnecessary for a man who was previously at higher risk. Identifying patients who have reduced their risks and could discontinue PrEP will be an important aspect of maintaining the cost-effectiveness of this intervention. Cost-effectiveness will also be affected if low risk men are identified as PrEP candidates, given the diminishing returns predicted by mathematical modeling studies when PrEP is administered to low risk men. (22) Specificity was generally low for all three scores; however SDET performed better than HIRI-MSM or Menza. This suggests that additional information beyond these risk scores might need to be considered with regard to initiation or discontinuation of PrEP.

This analysis has several strengths. First, the study population was recruited using a combination of venue-based time-space (VBTS) sampling and Facebook advertisements to obtain a minimally-biased sample of MSM in Atlanta, Georgia. Previous analyses have demonstrated that the participants recruited via Facebook did not differ from those recruited via VBTS (18), and VBTS reduces a number of biases inherent in other recruitment methods such as the dependence on social networks in respondent driven sampling (23). Men were followed for two years or until HIV seroconversion and completed HIV testing every six months. Extensive sexual behavior questionnaires were completed at baseline and were updated every six months if a new partner was reported.

There are also limitations to the current analysis. An additional published risk score developed using a population from the Southeastern United States could not be assessed in this analysis (24). One of the inputs for this risk score is type of facility where testing occurred (e.g., jail, counseling and testing site); this is not a variable that is available in the current data given that all testing occurred in the context of a research study. A relatively small number of seroconversions (N = 32) occurred over the follow-up period; this restricted us from being able to develop a new risk score that might perform better than the three that were assessed and resulted in poor precision for our estimates of sensitivity and AUC. The Involve[men] study population was limited to non-Hispanic black and white MSM in the Atlanta metropolitan statistical area, limiting the generalizability of this analysis to men of other races and ethnicities and of other geographic locations where different factors might be
driving HIV incidence. However, the Involve[men]t study enrolled more black MSM than previous studies, including those used to develop Menza(25) and HIRI-MSM(25, 26).

Each of the risk scores was developed and validated to predict incident infection at baseline or HIV seroconversion after six months, rather than the two-year follow-up period that was used in the present analysis. In this study two-year follow-up data were used in order to maximize the number of seroconversions available for analysis, increase accuracy of sensitivity estimates, and because it was not possible to update the risk score for all participants for each follow-up study visit. Finally, it was necessary to make modifications to the inputs for HIRI-MSM and Menza scores given the data that were available. Drug use variables in these two scores used a six-month window, whereas the drug use data from the Involve[men]t baseline visit used a 12-month window. This modification would result in an increase in sensitivity (and corresponding decrease in specificity) given the extended time window for recalling drug use.

Overall, the results of this study indicate that the published risk scores do not perform well in MSM in general. Further, the sensitivity of each score is lower among black MSM than among white MSM. This disparity in sensitivity has the potential to result in disparities in provision of HIV prevention interventions. Although the disparate sensitivities raise concerns about continuing disparities in HIV incidence in black compared to white MSM, the AUCs of all three scores indicate that none has good discrimination to identify either white or black high-risk MSM. Other metrics should be used to prioritize HIV prevention services, such as PrEP, for this disproportionately affected population. Additional efforts should also be undertaken to develop tools to identify high risk MSM overall.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

Sources of support: National Institutes of Health

**References**


Summary

Three HIV risk scores for men who have sex with men were found to have low sensitivity to predict seroconversion among MSM overall, and lower sensitivity in black compared to white MSM.
### Table 1

Risk score items, associated scores, and recommended cutoff scores.

<table>
<thead>
<tr>
<th>Score</th>
<th>Variables</th>
<th>Cutoff Score</th>
<th>Necessary Modifications</th>
</tr>
</thead>
</table>
| HIRI-MSM | • Age (18–28, +8; 29–40, +5, 41–48, +2; 49+, +0)  
Each of the following over the past 6 months:  
• Total number of male partners (>10, +7; 6–10, +4, <6, +0)  
• Total number of HIV-infected male partners (>1, +8; 1, +4; 0, +0)  
• Number of episodes of CRAI, with any partner (1 or more, +10; 0, +0)  
• Number of episodes of CIAI, with HIV-infected partner (5 or more, +6; <5, +0)  
• Amphetamine (+5 if yes) and popper use (+3 if yes) | >=10         | Recall period for drug use at baseline was 12 months                                    |
| Menza | • Gonorrhea, chlamydia, or syphilis diagnosis at baseline (+4 if yes)  
• Methamphetamine or popper use, past 6 months (+11 if yes)  
• Number of male sex partners, past 12 months (>9, +3)  
• CAI with serodiscordant partner (+1 if yes) | >=1          | Recall period for drug use at baseline was 12 months                                    |
| SDET | Each of the following over the past 12 months:  
• 10 or more partners (+2 if yes)  
• Any CRAI and at least 5 partners (+3 if yes)  
• Any CRAI with a HIV-infected partner (+3 if yes)  
• Syphilis, gonorrhea, or chlamydia diagnosis (+2 if yes) | >=5          | N/A                                                                                     |
Table 2

Sensitivity, specificity, positive predictive value, and negative predictive value for HIRI-MSM, Menza, and SDET risk scores over two years of follow-up.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (95% Confidence Interval)</th>
<th>Overall</th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRI-MSM</td>
<td>62.5 (43.7, 78.9)</td>
<td>58.3 (36.6, 77.9)</td>
<td>75.0 (34.9, 96.8)</td>
<td></td>
</tr>
<tr>
<td>Menza</td>
<td>62.5 (43.7, 78.9)</td>
<td>54.2 (32.8, 74.4)</td>
<td>87.5 (47.3, 99.7)</td>
<td></td>
</tr>
<tr>
<td>SDET</td>
<td>25.0 (11.5, 43.4)</td>
<td>16.7 (4.7, 37.4)</td>
<td>50.0 (15.7, 84.3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Specificity (95% Confidence Interval)</th>
<th>Overall</th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRI-MSM</td>
<td>56.7 (52.4, 61.0)</td>
<td>66.4 (60.0, 72.4)</td>
<td>49.0 (43.1, 54.8)</td>
<td></td>
</tr>
<tr>
<td>Menza</td>
<td>41.1 (36.9, 45.5)</td>
<td>41.5 (35.2, 48.1)</td>
<td>40.8 (35.1, 46.7)</td>
<td></td>
</tr>
<tr>
<td>SDET</td>
<td>83.9 (80.5, 87.0)</td>
<td>88.5 (83.7, 92.3)</td>
<td>80.3 (75.3, 84.7)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Positive Predictive Value (95% Confidence Interval)</th>
<th>Overall</th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRI-MSM</td>
<td>8.0 (5.0, 12.1)</td>
<td>15.1 (8.5, 24.0)</td>
<td>3.8 (1.4, 8.2)</td>
<td></td>
</tr>
<tr>
<td>Menza</td>
<td>6.0 (3.7, 9.2)</td>
<td>8.6 (4.7, 14.3)</td>
<td>3.9 (1.6, 7.8)</td>
<td></td>
</tr>
<tr>
<td>SDET</td>
<td>8.6 (3.8, 16.2)</td>
<td>12.9 (3.6, 29.8)</td>
<td>6.5 (1.8, 15.7)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Negative Predictive Value (95% Confidence Interval)</th>
<th>Overall</th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRI-MSM</td>
<td>96.2 (93.4, 98.0)</td>
<td>94.0 (89.2, 97.1)</td>
<td>98.6 (95.1, 99.8)</td>
<td></td>
</tr>
<tr>
<td>Menza</td>
<td>94.8 (91.1, 97.3)</td>
<td>89.9 (82.7, 94.9)</td>
<td>99.2 (95.5,100.0)</td>
<td></td>
</tr>
<tr>
<td>SDET</td>
<td>94.9 (92.5, 96.7)</td>
<td>91.2 (86.8, 94.6)</td>
<td>98.3 (95.8, 99.5)</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3
Area under the curve (95% CI) values from ROC analysis of Menza, HIRI-MSM, and SDET risk scores.

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Black</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIRI-MSM</td>
<td>0.62 (0.52, 0.72)</td>
<td>0.63 (0.51, 0.75)</td>
<td>0.67 (0.47, 0.88)</td>
</tr>
<tr>
<td>Menza</td>
<td>0.51 (0.41, 0.60)</td>
<td>0.49 (0.36, 0.62)</td>
<td>0.60 (0.44, 0.75)</td>
</tr>
<tr>
<td>SDET</td>
<td>0.55 (0.44, 0.66)</td>
<td>0.52 (0.39, 0.65)</td>
<td>0.66 (0.46, 0.87)</td>
</tr>
</tbody>
</table>