Cardiometabolic markers to identify cardiovascular disease risk in HIV-infected black South Africans

J M van Rooyen,1 DSc; C M T Fourie,1 RN, PhD; H S Steyn,2 DSc; G Koekemoer,2 PhD; H W Huisman,1 PhD; R Schutte,1 PhD; L Malan,1 RN, PhD; M Glynn,1 PhD; W Smith,1 PhD; C Mels,1 PhD; A E Schutte,1 PhD

1 Hypertension in Africa Research Team, School of Physiology, Nutrition and Consumer Sciences, Potchefstroom Campus, North-West University, Potchefstroom, South Africa
2 Statistical Consultation Services, Potchefstroom Campus, North-West University, Potchefstroom, South Africa

Corresponding author: J M van Rooyen (johannes.vanrooyen@nwu.ac.za)

Background. The prevalence of HIV is highest in sub-Saharan Africa; South Africa (SA) is one of the most affected countries with the highest number of adults living with HIV infection in the world. Besides the traditional risk factors for cardiovascular disease (CVD) in the general population, in people living with HIV there are specific factors – chronic inflammation, metabolic changes associated with the infection, therapy, and lipodystrophy – that potentially increase the risk for developing CVD.

Objective. This study proposes a screening discriminant model to identify the most important risk factors for the development of CVD in a cohort of 140 HIV-infected black Africans from the North West Province, SA.

Methods. Anthropometric measures, systolic blood pressure, diastolic blood pressure and the carotid-dorsalis pedis pulse wave velocity were determined. Blood was analysed to determine the levels of total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol, triglycerides (TGs) and glucose. Partial least squares discriminant analysis was performed as a supervised pattern recognition method. Independent Student's t-tests were further employed to compare the means of risk factors on interval scales; for comparison of categorical risk factors between groups, χ2 tests were used.

Results. A TG:HDL-C ratio ≥1.49, TC:HDL-C ratio ≥5.4 and an HDL-C level ≤0.76 mmol/l indicated CVD risk in this cohort of patients living with HIV.

Conclusion. The results have important health implications for black Africans living with HIV as these lipid levels may be a useful indicator of the risk for CVD.


- HIV prevalence is highest in sub-Saharan Africa, with South Africa (SA) being one of the most affected countries with the highest number of adults living with HIV in the world. The use of antiretroviral therapy (ART) has improved the quality and life expectancy in the HIV-infected population worldwide. The HIV-infected population in SA gained free access to ART through the antiretroviral (ARV) roll-out programme, which has made HIV a manageable disease in SA.

- However, infection with HIV and the treatment thereof may affect the heart and vasculature, leading to the development of cardiovascular disease (CVD), which currently, after cancer, is the most frequent cause of death among HIV-infected individuals. In SA, CVDs such as hypertension, atherosclerotic disease and heart failure are increasing among the black African population, which could be exacerbated by the high number of people living with HIV in SA. Furthermore, the HIV-infected population is ageing and, since age is a major non-modifiable risk factor for CVD, the risk for CVD may be expected to increase progressively in this population. In relation to CVD, HIV infection is associated with an increase in insulin resistance, dyslipidaemia, lipodystrophy, endothelial dysfunction, accelerated atherosclerosis and coagulation disorders. Lipid abnormalities, characterised by low high-density lipoprotein cholesterol (HDL-C) and high triglyceride (TG) levels form a large part of the risk for CVD in the HIV-infected population, and may be related to the viral infection, ART or both. Even though black Africans normally exhibit lower fasting TG and higher HDL-C levels than whites, the HDL-C levels of HIV-infected black South Africans were found to be lower than the levels normally associated with increased risk for the development of CVD. Besides the risk factors for CVD seen in the general population, there are specific factors in people living with HIV that could increase risk, i.e. chronic inflammation, metabolic changes associated with the infection, its therapy and lipodystrophy.

- HIV-infected people are managed in clinics through the ARV roll-out programme in SA. Early identification and management of CVD complications is essential, but also difficult as resources are not always readily available. The development of a non-invasive tool to help identify individuals who are most likely to derive clinical benefit from CVD risk-reduction therapy is therefore of great importance. This study aimed to identify the most important risk factors for the development of CVD in a cohort of 140 HIV-infected black Africans from the North West Province, SA.

Methods

Study design and participants

This cross-sectional sub-study is embedded in the international Prospective Urban and Rural Epidemiology (PURE) study, which was designed to address questions regarding the cause and development of cardiovascular risk factors and disease. In the present sub-study we included 154 HIV-uninfected and 140 HIV-infected persons known to be infected for at least three years. Data collection took place in 2008 and the 140 HIV-infected participants were used in a hierarchical cluster analysis to separate participants with and without risk for the development of CVD. HIV-infected participants were stratified into risk and no risk groups and the following variables were assessed: age;
body mass index (BMI); waist-to-hip ratio; systolic blood pressure (SBP); diastolic blood pressure (DBP); carotid-dorsalis pedis pulse wave velocity (cPWV); total cholesterol (TC); HDL-C; low-density lipoprotein cholesterol (LDL-C); TG; TG:HDL-C ratio; TC:HDL-C ratio; and glucose. The group identified by the analysis as exhibiting CVD risk factors was cross validated for the cut-off values for CVD risk according to the SA guidelines. All the participants in the CVD risk group exhibited one or more CVD risk factors. The cluster analysis was based on the cardiometabolic profile of the HIV-infected participants and was used to separate the HIV-infected participants into a group exhibiting risk factors for the development of CVD and a group without CVD risk factors.

Of the participants, 63 received ART consisting of two nucleoside reverse transcriptase inhibitors ( stavudine and lamivudine) plus one non-nucleoside reverse transcriptase inhibitor ( efavirenz or nevirapine). Twenty of the 63 participants had received ART for >2 years. The mean CD4+ cell count for those receiving ART was 260 cells/μl and for those not receiving ART was 264 cells/μl.

All participants provided signed informed consent after all procedures were explained in their home language. The study protocol complies with the Declaration of Helsinki as revised in 2008 and was approved by the Ethics Committee of the North-West University, Potchefstroom, SA (code 04M10).

Experimental protocol

During data collection, the participants arrived at the research locality of the rural or urban areas at 08h00 each morning after a 10 - 15-min drive (provided by the research team) from their communities. The participants were introduced to the planned research and, after the procedures were explained, they signed informed consent forms and received HIV pre-counselling from trained counsellors. In the course of the morning, demographic and lifestyle questionnaires were completed with the help of specially trained field workers in each subject’s home language. Lifestyle data included self-reported current tobacco and alcohol use as well as medical history. Referrals to the local clinic or hospital, if necessary, and feedback on HIV status and cardiovascular variables were provided to participants during individual post-counselling.

Anthropometric measurements

Height, weight, hip and waist circumference (WC) were measured (Precision Health Scale, A & D Company, Japan; Invicta Stadiometer, IP 1465, UK; Holtain unstretchable metal tape) using standardised procedures.

Demographic questionnaire

Demographic and lifestyle information were obtained from the above-mentioned questionnaire.

Cardiovascular measurements

SBP and DBP were obtained with the validated OMRON HEM-757 apparatus (Omron Healthcare, Japan). After a 10-min rest period, blood pressure measurements were performed twice (5 min apart) on the right arm while the participant was seated upright and relaxed with his/her right arm supported at heart level. Appropriate sized cuffs were used for obese participants.

Pulse wave velocity, a measure of arterial stiffness, was measured on the left side of each participant in the supine position. The measurement was determined using non-invasively accessible superficial pulses with the Compilor SP device (Artech-Medical, France) in a segment over the carotid-dorsalis pedis vessels.

Blood, serum and plasma samples

The subjects were required to fast overnight. Blood was drawn from the brachial vein using a sterile winged infusion set and syringes. Serum and plasma samples were prepared according to appropriate methods and stored at -80°C. In rural areas, samples were stored at -18°C (for no longer than five days) until transported to the laboratory facility and then stored at -80°C until analysis.

Biochemical analysis

Quantitative determinations of the TC, HDL-C, TG and glucose were performed with the Beckman Coulter DxC 800 Synchron Clinical System (Beckman Coulter Inc., USA). LDL-C was calculated by using the Friedewald formula. HIV status was determined with the First Response rapid HIV card test (FMC Medical, India) using whole blood. If tested positive, the test was repeated with the Sensa card test (Seyama Solutions, SA).

Statistical analysis

Statistical analysis was performed using Statistica (version 10.0) and R (version 2.13). Independent Student's t-tests were used to compare means of risk factors on interval scales; for comparison of categorical risk factors between groups, χ² tests were used.

Multivariate analysis was performed in the form of pattern recognition methods. Pattern recognition methods are commonly divided into supervised and unsupervised methods, where supervised methods make use of the grouping of information in the data. Prior to multivariate analysis, the input data were scaled using the transformation \[ Y = \log\frac{X}{\bar{X}} \], where \( \bar{X} \) is the arithmetic mean. This transformation was applied to make the scales of the CVD risk factors more comparable, so that the factors with larger scales would not dominate in the multivariate analysis. Hierarchical cluster analysis using the method of Ward with an euclidean distance measure was then performed as an unsupervised pattern recognition method to create two clusters of cases with similar multivariate profiles concerning the CVD risk factors measured. These clusters contain cases with and without CVD risk within the 140 HIV-infected participants. Using these clusters as dependent variables and the risk factors as independent predictors, a partial least squares (PLS) discriminant analysis (DA) was performed as a supervised pattern recognition method. Although the PLS-DA model can be used to predict a person’s cluster membership given the CVD risk factors as input, our aim in fitting this model was to identify those risk factors that contribute most (in a multivariate sense) to the existence of the grouping (clusters) formed by the cluster analysis.

The variable importance in the projection (VIP) from PLS-DA is a measure that is commonly used in this regard. Risk factors with a VIP value >1 were used as the criteria for variable selection. To validate the interpretations made from the discriminant model, the following fit statistics are reported: percentage of variance explained for the risk factors (R²X), the percentage variance explained for the dependent variable (R²Y), the predictive R²Y values (Q’) and cross-validated estimates of sensitivity (percentage true positives) and specificity (percentage true negatives). The model dimension was selected using the Q’ statistics. The cross-validated estimates of sensitivity and specificity are calculated to confirm that the PLS-DA model did not overfit the data. Hence, whether the model could make accurate cluster predictions for cases unseen by the model was investigated. These estimates were calculated as follows:

A PLS-DA model was constructed using all 140 cases. This model was used to determine a cut-off point for the model predicted value (used to make cluster predictions) by calculating the Youden index. Then, 10 000 unique stratified samples of size 39 (24 from cluster 1 and 15 from
cluster 2) were randomly selected and excluded from the data. This equated to an exclusion of ~30% of the data. For each of the 10 000 samples, a PLS-DA model was constructed using the remaining cases, after which the cluster membership of the withheld cases was predicted. For each of the 10 000 trials, the sensitivity and specificity were recorded and finally averaged to find the cross-validation estimates of sensitivity and specificity.

Receiver-operating characteristic (ROC) analysis was used to determine the optimum cut-off values (based on the Youden index) for risk and no risk in HIV-infected participants to develop CVD for each of the selected variables.

Results
In Table 1, the characteristics of HIV-infected participants are compared with the uninfected control participants. The HIV-infected participants had significantly lower SBP, TC and HDL-C levels. However, their TC:HDL-C and TG:HDL-C ratios were significantly higher compared with the uninfected control group.

The dendrogram of the cluster analysis of the 140 infected participants is shown in Fig. 1, from which a clear separation of the cases is visible. To identify the risk factors that contribute most to the separation, a PLS-DA model was constructed using the clusters formed (Fig. 1) as dependent variables and the risk factors as independent variables. One component was extracted for the PLS-DA model based on the Q² value. The summary fit statistics of the PLS-DA model indicated that the model accounted for only 23.2% of the total variance observed in the scaled risk factors; however, the variance explained for the dependent variable is 61.5%. The Q² value was 0.602, which indicated that the model did not overfit the data, and that good prediction performance could be expected if required. This was further confirmed by the cross-validation estimates of sensitivity and specificity, which were 94.4% and 93.3%, respectively. The respective standard errors of these estimates were 6% and 4.7%. From these statistics we concluded that the PLS-DA model fits reasonably well and could be used to identify risk factors that separated the group into no- and high-CVD risk within the HIV-infected participants.

There was no statistically significant difference between the risk and no-risk groups concerning age, gender, BMI, ART, alcohol or tobacco use. The CD4+ count (291 v. 225 cells/μl; p=0.08) also did not differ significantly when no-risk and risk groups were compared.

VIP values are presented in Table 2, from which TG:HDL-C ratio, HDL-C and

<table>
<thead>
<tr>
<th>Demographic</th>
<th>HIV-uninfected (N=154)</th>
<th>HIV-infected (N=140)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (±SD)</td>
<td>48.3 (±8.25)</td>
<td>47.3 (±7.01)</td>
<td>0.28</td>
</tr>
<tr>
<td>Gender (male), n (%)</td>
<td>57 (37)</td>
<td>45 (32)</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Anthropometric, mean (±SD)

| BMI (kg/m²) | 23.7 (±5.91) | 22.7 (±5.55) | 0.15 |
| WC (cm) | 78.2 (±10.5) | 76.1 (±11.7) | 0.11 |
| Waist:hip ratio | 0.83 (±0.08) | 0.82 (±0.09) | 0.59 |

Cardiovascular, mean (±SD)

| SBP (mmHg) | 134 (±21.6) | 129 (±18.9) | 0.05 |
| DBP (mmHg) | 88.6 (±13.2) | 86.8 (±11.6) | 0.23 |
| cdPWV (m/s) | 8.81 (±1.70) | 8.52 (±1.60) | 0.15 |

Biochemical, mean (±SD)

| TC (mmol/l) | 4.25 (±1.18) | 3.93 (±1.14) | 0.02 |
| HDL-C (mmol/l) | 1.19 (±0.55) | 0.92 (±0.45) | <0.01 |
| LDL-C (mmol/l) | 2.54 (±0.93) | 2.49 (±0.91) | 0.71 |
| TG (mmol/l) | 1.13 (±0.78) | 1.19 (±0.26) | 0.59 |
| TC:HDL-C ratio | 4.15 (±2.68) | 4.91 (±1.94) | <0.01 |
| TG:HDL-C ratio | 1.17 (±1.51) | 1.66 (±1.44) | <0.01 |
| Glucose (mmol/l) | 4.60 (±1.03) | 4.57 (±0.77) | 0.78 |
| hsCRP (mg/l), mean (95% CI) | 2.62 (0.5 - 27.3) | 3.05 (0.5 - 34.4) | 0.34 |

Lifestyle, n/N (%)†

| Tobacco use | 68/142 (48) | 37/118 (31) | <0.01 |
| Alcohol use | 72/142 (51) | 46/118 (39) | 0.06 |

SD = standard deviation; BMI = body mass index; WC = waist circumference; SBP = systolic blood pressure; DBP = diastolic blood pressure; cdPWV = carotid-dorsalis pedis pulse wave velocity; TC = total cholesterol; HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; TG = triglyceride; hsCRP = high-sensitivity C-reactive protein; CI = confidence interval.

*P-values were obtained with independent Students’ t-tests or χ²-tests.
†N = total participants who indicated their alcohol/tobacco use on the questionnaire.

![Fig. 1. Dendrogram of the HIV-infected participants to show the clustering of variables.](image-url)
TC:HDL-C ratio were identified as the risk factors that primarily cause group separation. Tobacco and alcohol use were not included in the PLS-DA as it was obtained from a self-reported questionnaire. High-sensitivity C-reactive protein (hsCRP) was included primarily in the PLS-DA but did not contribute to the cluster formation. To test the discriminatory ability of the three main risk factors identified, a PLS-DA model containing only HDL-C, TG:HDL-C and TC:HDL-C ratios (as independent variables) was fitted and cross-validated. The estimated sensitivity was 96.4% and the estimated specificity was 92.1%. The respective standard errors of these estimates were 4.9% and 4.8%. This confirms that the three risk factors identified contained almost all the discriminatory information, since similar results were obtained when using all 13 variables in the PLS-DA model.

The distributional properties and the relationship among these three risk factors are illustrated as a matrix-plot (Fig. 2). Categorised boxplots of each risk factor are on the diagonal and scatterplots with appropriate labels are in the off-diagonal entries. Each box’s height indicates the variability (difference between third and first quartile), while the median is the central line. The whiskers give the range of values except for the extremes displayed as individual points.

From ROC analysis, optimum cut-off values were determined for the participants. The cut-off values between risk and no-risk for CVD in participants were: HDL-C 0.76 mmol/l; TC:HDL-C ratio 5.40; and TG:HDL-C ratio 1.49.

**Discussion**

By making use of the PLS-DA statistical analysis method, we identified those CVD risk variables that contributed most (in a multivariate sense) to the existence of the grouping (clusters) observed in the total HIV-infected group that consisted of ARV-treated and never-treated participants. The most important variables identified by the PLS-DA were TG:HDL-C, HDL-C and TC:HDL-C with values of VIP > 1. ROC analysis indicated that a TG:HDL-C ratio ≥1.49, TC:HDL-C ratio ≥5.4 and an HDL-C level ≤0.76 mmol/l constituted risk values with the potential to contribute to the development of CVD in HIV-infected black Africans.

The statistical PLS-DA method was employed to discriminate between HIV-infected participants with risk and no risk in the same group. Wold et al.\[22\] reported that PLS regression is useful for the analysis of noisy, incomplete data where multicollinearity is present. The PLS-DA method is also useful when a large number of variables are investigated simultaneously. In addition, the

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG:HDL-C ratio</td>
<td>2.66</td>
</tr>
<tr>
<td>HDL-C (mmol/l)</td>
<td>1.68</td>
</tr>
<tr>
<td>TC:HDL-C ratio</td>
<td>1.16</td>
</tr>
<tr>
<td>TG (mmol/l)</td>
<td>0.98</td>
</tr>
<tr>
<td>LDL-C (mmol/l)</td>
<td>0.65</td>
</tr>
<tr>
<td>TC (mmol/l)</td>
<td>0.52</td>
</tr>
<tr>
<td>Glucose (mmol/l)</td>
<td>0.22</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.07</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>0.06</td>
</tr>
<tr>
<td>Waist:hip ratio</td>
<td>0.05</td>
</tr>
<tr>
<td>cdPWV (m/s)</td>
<td>0.04</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>0.03</td>
</tr>
<tr>
<td>DBP (mmHg)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

VIP = variable importance in the projection; CVD = cardiovascular disease; TG = triglyceride; HDL-C = high-density lipoprotein cholesterol; TC = total cholesterol; LDL-C = low-density lipoprotein cholesterol; BMI = body mass index; cdPWV = carotid-dorsalis pedis pulse wave velocity; SBP = systolic blood pressure; DBP = diastolic blood pressure.
PLS/VIP method is used to assist in the identification of relevant predictors (Table 2), while Chong and Jun found that this outperforms other variable selection methods.

HIV-related CVD is under-recognised and the clinical assessment thereof is a critical challenge for practitioners, especially in SA. Bearing in mind the enormous task facing the practitioners managing the ARV roll-out programme and the high number of people in SA living with HIV, the results of this study should simplify the identification of those individuals who exhibit a CVD risk. The discriminant model showed that the determination of HDL-C levels, and TG:HDLC and TC:HDLC ratios could be of clinical benefit for the identification of individuals at risk. It is well known that people living with HIV exhibit dyslipidaemia characterised by low HDLC and high TG levels. The latter was reported in newly identified HIV-infected South Africans, although black Africans normally exhibit lower fasting TG and higher HDLC levels than whites. The HDLC levels seen in the total HIV-infected group in this study (0.92 mmol/l) were below the level (1.28 mmol/l) generally regarded as an increased risk of developing CVD.

The cut-off values of TG:HDLC (1.49) and HDLC levels (0.76 mmol/l) seen in our study indicated that risk is lower than that described in the literature. However, modified CVD risk equations and models such as the Framingham Risk Score have not been validated for use in black Africans living with HIV. Thus, in light of the ethnic differences in TG and HDLC levels, and the findings of this study, different cut-off values may be needed to define the dyslipidaemia in HIV-infected black Africans that indicates risk of developing CVD. However, to evaluate individual risk the composite risk factor profile of traditional risk factors such as gender, lifestyle (smoking and alcohol abuse) and family history of CVD must be considered. The strength of this study is that it contributes to CVD risk data in HIV-infected black South Africans, which are sparse. It would benefit practitioners, and be clinically advantageous to use metabolic variables such as TG:HDLC and TC:HDLC ratios as well as HDLC levels, to discriminate between those individuals, who are at risk, and those who are not, for the risk of developing CVD.

Study limitations
Although there were no gender or treatment differences between the risk and no-risk groups, a limitation of the study is that this model did not differentiate between gender and those who were treated and those who did not receive treatment. It is difficult to differentiate between risk associated with HIV and that associated with ART. As previously reported, not all HIV-infected individuals eligible for treatment have been enrolled in the ART programme and, in fact, in 2009, the treatment coverage in SA was <40%. When the interaction between treatment and no treatment was determined statistically in the risk groups, interaction was seen only in the TC:HDLC ratio. ART (including nevirapine and efavirenz) were associated with the normalisation of HDLC levels and an increase in TC levels.

Conclusion
Our study involved a cohort of people living with HIV in the North West Province, SA. Future research in larger prospective studies is warranted to assess the predictive power of lipid ratios and HDLC levels for CVD development in people living with HIV in sub-Saharan Africa. Despite the limitations and the relatively small size of this study, we propose that by employing lipid ratios and HDLC levels for screening, early identification of South Africans living with HIV and at risk for CVD may be achieved. A TG:HDLC ratio ≥1.49, TC:HDLC ratio ≥5.4 and an HDLC level ≤0.76 mmol/l were indicative of CVD risk. These results have important health implications for blacks living with HIV, as these lipid levels may be useful indicators of CVD risk in clinical settings.

Acknowledgements. We thank Prof. A Kruger; the PURE-SA research team; the field workers and office staff at the Africa Unit for Transdisciplinary Health Research (AUTHer), North-West University; SA; Dr S Yusuuf (PURE-International) and the PURE project staff at the Population Health Research Institute (PHRI), Hamilton Health Sciences and McMaster University, Canada. This work was financially supported by the South Africa-Netherlands Research Programme on Alternatives in Development, South African National Research Foundation (NRF GUN no. 200939 and FA20060700010), North-West University, PHRI and the Medical Research Council of South Africa.

References
22. We thank Prof. A Kruger; the PURE-SA research team; the field workers and office staff at the Africa Unit for Transdisciplinary Health Research (AUTHer), North-West University; SA; Dr S Yusuuf (PURE-International) and the PURE project staff at the Population Health Research Institute (PHRI), Hamilton Health Sciences and McMaster University, Canada. This work was financially supported by the South Africa-Netherlands Research Programme on Alternatives in Development, South African National Research Foundation (NRF GUN no. 200939 and FA20060700010), North-West University, PHRI and the Medical Research Council of South Africa.