Machine learning models reveal neurocognitive impairment type and prevalence are associated with distinct variables in HIV/AIDS

Wei Tu, Patricia A. Chen, Noshin Koenig, Daniela Gomez, Esther Fujiwara, M. John Gill, Linglong Kong & Christopher Power

Journal of NeuroVirology

ISSN 1355-0284
Volume 26
Number 1

J. Neurovirol. (2020) 26:41-51
DOI 10.1007/s13365-019-00791-6

Springer

Journal of NeuroVirology (2020) 26:41–51
https://doi.org/10.1007/s13365-019-00791-6

Machine learning models reveal neurocognitive impairment type and prevalence are associated with distinct variables in HIV/AIDS

Wei Tu[1] · Patricia A. Chen[2] · Noshin Koenig[3] · Daniela Gomez[4] · Esther Fujiwara[4] · M. John Gill[3] · Linglong Kong[1] · Christopher Power[2,4]

Received: 8 April 2019 / Revised: 2 August 2019 / Accepted: 12 August 2019 / Published online: 13 September 2019

Abstract

Neurocognitive impairment (NCI) among HIV-infected patients is heterogeneous in its reported presentations and frequencies. To determine the prevalence of NCI and its associated subtypes as well as predictive variables, we investigated patients with HIV/AIDS receiving universal health care. Recruited adult HIV-infected subjects underwent a neuropsychological (NP) test battery with established normative (sex-, age-, and education-matched) values together with assessment of their demographic and clinical variables. Three patient groups were identified including neurocognitively normal (NN, $n = 246$), HIV-associated neurocognitive disorders (HAND, $n = 78$), and neurocognitively impaired-other disorders (NCI-OD, $n = 46$). Univariate, multiple logistic regression and machine learning analyses were applied. Univariate analyses showed variables differed significantly between groups including birth continent, quality of life, substance use, and PHQ-9. Multiple logistic regression models revealed groups again differed significantly for substance use, PHQ-9 score, VACS index, and head injury. Random forest (RF) models disclosed that classification algorithms distinguished HAND from NN and NCI-OD from NN with area under the curve (AUC) values of 0.87 and 0.77, respectively. Relative importance plots derived from the RF model exhibited distinct variable rankings that were predictive of NCI status for both NN versus HAND and NN versus NCI-OD comparisons. Thus, NCI was frequently detected (33.5%) although HAND prevalence (21%) was lower than in several earlier reports underscoring the potential contribution of other factors to NCI. Machine learning models uncovered variables related to individual NCI types that were not identified by univariate or multiple logistic regression analyses, highlighting the value of other approaches to understanding NCI in HIV/AIDS.

Keywords Neurocognitive impairment · HIV-associated neurocognitive disorders · Machine learning · Neuropsychology · Comorbidity

Introduction

Within days of primary infection by human immunodeficiency virus type 1 (HIV-1), the virus can be detected within the central nervous system (CNS) (Davis et al. 1992). CNS infection by HIV-1 results in the production and release of host immune molecules and viral proteins, accompanied by frontal-striatal neuronal damage and death (Robertson et al. 2009). The cumulative effects of these pathways contribute to neurocognitive impairment (NCI) observed in a subset of patients with HIV-1 infection although systemic immunosuppression or inflammation also contribute to neuropathogenesis (Saylor et al. 2016). Recent studies indicate these processes are incompletely abrogated by combination antiretroviral therapy (cART) regimens. In fact, there is concern that some drugs comprising cART might contribute to NCI (Mothobi and Brew 2012).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13365-019-00791-6) contains supplementary material, which is available to authorized users.

✉ Linglong Kong
lkong@ualberta.ca

✉ Christopher Power
chris.power@ualberta.ca

[1] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada

[2] Department of Medicine (Neurology), University of Alberta, 6-11 Heritage Medical Research Centre, Edmonton, AB, Canada

[3] Department of Medicine, University of Calgary, Calgary, AB, Canada

[4] Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

NCI involves usually executive, motor, attention, memory, and less commonly language impairments with varying severity and frequency in patients with HIV/AIDS. The cardinal manifestations of HIV-driven NCI have been deemed the diagnostic entity, HIV-associated neurocognitive disorders (HAND) (Antinori et al. 2007). The HAND syndrome is classified as asymptomatic neurocognitive impairment (ANI), minor neurocognitive disorder (MND), or HIV-associated dementia (HAD) (Antinori et al. 2007). The reported prevalence of NCI in patients with HIV-1 infection despite the use of cART is wide-ranging, affecting 15–60% of HIV-infected patients, depending on the individual study (van den Dries et al. 2017; Belete et al. 2017; Sacktor et al. 2016; Heaton et al. 2010). These wide-ranging prevalence values suggest the evaluation and diagnosis of NCI in HIV-infected patients has become increasingly complex. The overall mortality has declined in patients with HIV-1 infection receiving in active care (Nakagawa et al. 2013; Roehr 2015; Puhan et al. 2010). Indeed, HIV-infected patient cohorts now exhibit more age-related cardiovascular and neurological comorbidities, are exposed to different cART regimens with specific off-target effects, and may be more demographically diverse due to migration. Other factors affect NCI including mood disorders, substance abuse, opportunistic infections, and unrecognized neurodevelopmental disorders (reviewed in (Smail and Brew 2018)). Given the increasing importance of these latter variables in the development of HIV-associated NCI, it is important to assess their associations with NCI in the context of contemporary diagnostic guidelines.

The present study assessed the prevalence and associated variables in HIV-infected patients with or without NCI in a contemporary clinical setting in which patients received long-term clinical follow-up, access to universal health care, and were of multiple ethnic and racial backgrounds. We hypothesized that implementing diverse statistical approaches to well-defined datasets from HIV/AIDS patients with and without NCI might yield new insights into variables contributing to NCI including HAND. To gain a deeper and more robust understanding of the relative importance of individual clinical and demographic variables involved in NCI occurrence and its subtypes, we compared univariate, logistic regression to machine learning analyses of these present datasets.

Methods and materials

Patient cohort HIV-1 seropositive patients at the Southern Alberta (HIV) Clinic (SAC) in Calgary, Alberta, were randomly invited to enroll in the study during routine clinical care with neurocognitive status as the primary outcome measure with defined exclusion criteria. The University of Calgary Ethics Committee (REB #-130615) approved the study and written consent was obtained from all patients.

Study setting The SAC serves all people with HIV-1 infection in active care in Southern Alberta and is a multidisciplinary clinic which opened in 1989 and offers regular clinical follow-up visits, laboratory investigations, free and modern antiretroviral therapies, all without costs to patients. A multidisciplinary team including physicians, nurses, social workers, dieticians, and pharmacists staffs the clinic. The SAC also maintains an in-house computerized database of all HIV-infected patients containing relevant patient diagnostic and treatment data (Kim et al. 2001; Pettersen et al. 2006; Vivithanaporn et al. 2010; McCombe et al. 2013).

Patient clinical and demographic variables A wide range of variables was extracted from the clinical database (Table 1) that included sex, age, continent of birth, years of education, current employment status, sexual orientation, duration of HIV-1 infection, presence or absence of AIDS, current and nadir CD4+ T cell counts, current/peak plasma viral load, cART treatment status, cART side effects (including cART toxicity), CNS penetration effectiveness (CPE) rank of current cART, and comorbidities (e.g., cardiovascular disease, hepatitis C virus seropositivity) as well as the VACS index; past and present substance use/abuse of alcohol, marijuana, cocaine, heroin, and other illicit substances; medical conditions: diabetes, hypothyroidism, head injury, and psychiatric disorder; HIV-related polyneuropathy; dyslipidemia, lipodystrophy, and toxoplasma serology. In addition, "polypharmacy" summarized all prescription and over-the-counter medications in addition to cART but excluded nutritional supplements. Birth continent was designated as the site of birth and included North America (NA), South America (SA), Asia, Africa, and Europe.

On the test day, patients were asked about their current health-related quality of life (HQoL) (Crane et al. 2006): "How has your health been in the past five days?" (1 = poor, 2 = fair; 3 = good; 4 = very good; 5 = excellent), the number of days with cART non-adherence in the past 5 days, concerns about sleep quality and number of hours of sleep per night. The presence and severity of depressive symptoms was assessed by the Patient Health Questionnaire (PHQ-9) assay and stratified in an established manner (Kroenke et al. 2001). All patients with NCI or declared neurocognitive symptomology underwent cranial magnetic resonance imaging (MRI) and/or CSF analyses.

Neuropsychological testing A standard neuropsychological testing battery lasting 45–90 min was conducted at the SAC and was comprised of tests of attention/processing speed, memory, motor, language, and executive functions (Gomez et al. 2017). The neuropsychological battery included the Symbol Digit Modalities Test (SDMT; correct responses) (Smith 1973), Trail Making Test Number Sequencing subtest from the Delis-Kaplan Executive Function System (D-KEFS)

Table 1 Comparisons of clinical and sociodemographic variables for the neurocognitively normal (NN), HIV-associated neurocognitive disorders (HAND), and neurocognitively impaired-other disorders (NCI-OD) groups

Variable[§]	NN ($n = 246$)	HAND ($n = 78$)	NCI-OD ($n = 46$)	Univariate test	Multiple logistic
Age (years)	47.6 (11.3)	47.4 (10.3)	45.5 (10.9)		
Sex (female)	36 (14.6%)	8 (10.3%)	3 (6.5%)		
Birth continent				a^{++}, b^{++}	A^{++}, B^{+}
North America	212 (86.2%)	43 (55.1%)	30 (65.2%)		
South America	2 (0.8%)	5 (6.4%)	2 (4.3%)		
Asia	11 (4.5%)	8 (10.3%)	6 (13.0%)		
Africa	12 (4.9%)	20 (25.6%)	6 (13.0%)		
Europe	9 (3.7%)	2 (2.6%)	2 (4.3%)		
Sexual orientation				a^{+}	
Bisexual	31 (12.6%)	5 (6.4%)	7 (15.2%)		
Heterosexual	70 (28.5%)	36 (46.2%)	18 (39.1%)		
Homosexual	145 (58.9%)	37 (47.4%)	21 (45.7%)		
Education (years)	14.0 (2.5)	14 (2.5)	13.2 (3.6)		
Employment (current) employed)	171 (69.5%)	49 (62.8%)	26 (56.5%)		B^{+}
Worked hours/week	28.0 (22.7)	25.0 (20.1)	22.5 (24.2)		
Quality of life	3.7 (0.9)	3.7 (1.0)	3.2 (1.3)	b^{+}, c^{+}	B^{+}, C^{+}
Cigarette use	55 (22.4%)	14 (17.9%)	16 (34.8%)	c^{+}	
Sleep (hours/night)	6.8 (1.4)	6.4 (1.8)	6.0 (2.0)	b^{+}	B^{+}
Current substance use current					
Alcohol (binge)	11 (4.5%)	1 (1.3%)	5 (10.9%)	a^{++}	A^{+}, C^{+}
Marijuana	71 (28.9%)	16 (20.5%)	17 (37.0%)		
Crack/cocaine	22 (8.9%)	4 (5.1%)	3 (6.5%)		
Past substance use					
Alcohol (binge)	3 (1.2%)	5 (6.4%)	1 (2.2%)	a^{+}	A^{+}
Marijuana	70 (28.5%)	16 (20.5%)	16 (34.8%)		
Crack/cocaine	50 (20.3%)	12 (15.4%)	19 (41.3%)	b^{++}, c^{++}	B^{+}, C^{++}
Peak viral load	4.7 (1.0)	4.6 (1.3)	4.7 (1.1)		
Current viral load copies/ml (log10)	1.8 (0.6)	1.8 (0.6)	2.0 (0.6)		C
Current CD4+ T cell (count/mm-)	584.9 (257.9)	525.3 (235.7)	543.8 (244.9)		
Nadir CD4+ T cell	216.7 (171.7)	221.6 (163.9)	234.5 (197.0)		A^{+}
AIDS-defined	114 (46.3%)	36 (46.2%)	24 (52.2%)		
Duration of HIV (years)	11.5 (8.0)	11.5 (8.4)	9.7 (6.9)		
cART at NP testing	232 (94.3%)	76 (97.4%)	43 (93.5%)		
cART side effects	48 (19.5%)	12 (15.4%)	3 (6.5%)	b^{+}	
cART non-adherence	22 (8.9%)	11 (14.1%)	3 (6.5%)		
VACs index	13.3 (13.1)	13.4 (12.7)	15.7 (14.4)		B^{++}, C^{++}
Depressive symptoms (PHQ-9)	5.6 (5.3)	6.4 (6.3)	11.3 (7.7)	b^{++}, c^{++}	B^{++}, C^{++}
CPE rank	6.8 (2.1)	7.3 (1.8)	6.8 (2.4)		
Polypharmacy	4.2 (3.4)	4.1 (3.5)	5.3 (4.8)		
Diabetes	18 (7.3%)	10 (12.8%)	5 (10.9%)		
Cardiac conditions	48 (19.5%)	12 (15.4%)	8 (17.4%)		
Lipodystrophy	19 (7.7%)	7 (9.0%)	3 (6.5%)		
Dyslipidemia	64 (26.0%)	16 (20.5%)	8 (17.4%)		
Seroconversion illness	44 (17.9%)	6 (7.7%)	6 (13.0%)	a^{+}	
Toxoplasmosis (sero + ve)	26 (10.6%)	18 (23.1%)	7 (15.2%)	a^{++}	
Head trauma	58 (23.6%)	21 (26.9%)	20 (43.5%)	b^{+}	B^{+}
Polyneuropathy	44 (17.9%)	18 (23.1%)	7 (15.2%)		

[§] Viral load (log_{10} copies/ml); CD4+ T cell (cell/mcl), toxoplasmosis (seropositivity), depressive symptoms (mean PHQ-9 score), diabetes (types 1 and 2). Data in parentheses indicate mean and standard deviation (continuous variables) or occurrence and percentages (categorical variables)

a, b, and c refer to the significant variables ($p < 0.05$) when comparing NN versus HAND groups, NN versus NCI-OD groups, and HAND versus NCI-OD groups, respectively, using univariate tests; A, B, and C refer to the significant variables ($p < 0.05$) when comparing NN versus HAND groups, NN versus NCI-OD groups, and HAND versus NCI-OD groups, respectively, using multiple logistic regression. The subscripts + and ++ indicate significance at confidence level 0.05 and 0.01 respectively; for example, a^{++} indicates p value smaller than 0.01 between NN and HAND groups using univariate test

(Delis et al. 2001), Grooved Pegboard (dominant and non-dominant hand) (Trites 1977), Hopkins Verbal Learning Test-Revised (HVLT-R; immediate and 25-min delayed recall) (Brandt and Benedict 2001), D-KEFS (letter fluency and category fluency, correct responses) (Delis et al. 2001), Wisconsin Card Sorting Test (WCST-64; perseverative and non-perseverative errors) (Kongs et al. 2000), and Trail Making Test Number-Letter Switching subtest from the D-

KEFS (Delis et al. 2001). All test scores were interpreted relative to established normative results from healthy age-, sex-, and education-matched controls (Kongs et al. 2000; Schretlen et al. 2010). Exclusion criteria included age less than 18 years, non-fluency in English, education less than 9 years, hepatitis C virus (HCV) infection with documented hepatic encephalopathy, untreated psychosis/schizophrenia, major depressive disorder, anxiety disorders, seizures, or other mental or neurological illness requiring past hospitalization, traumatic brain injury requiring hospitalization or loss of consciousness of greater than 5 min, current alcohol or drug abuse that might affect neuropsychological test outcomes, documented developmental disability, uncorrected vision or hearing disorders, and inability to consent (e.g., under legal guardianship).

Neurocognitive impairment (NCI) was defined as neuropsychological performance of at least 1.0 standard deviation ($z < -1.0$) below the normative means in at least two of the tested domains. The diagnosis and staging of HAND was established by the presence of impairment on neuropsychological testing coupled with a review of each patient by the SAC team to ensure that there was an apparent functional decline in patients' daily performance at home or in the workplace as well as excluding any confounding comorbidities that might contribute to neurocognitive impairment (Antinori et al. 2007). The diagnosis of neurocognitive impairment-other disorder (NCI-OD) was based on neurocognitive impairment in at least two tested domains ($z < -1.0$) and did not require evidence of a functional decline. In contrast to HAND, clinical assessment of NCI-OD reflected immediate causes of NCI that were not apparent or unrecognized at the time of recruitment including mood disorders, substance use, traumatic head injury, or neurodevelopmental disorders.

Statistical analyses Demographic and clinical comparisons between groups were performed using univariate and multivariate methods, as well as a principal component analysis. Univariate tests were conducted using Mann-Whitney *U* test for continuous data and Fisher's exact test for categorical data. For multivariate methods, a logistic regression model was first applied using all 37 demographic and clinical variables to predict NCI status of patients. To assess the limited predictive performance of logistic regression, an exploratory principal component analysis was implemented to assess the possibility of linear separation of these three groups, and the results prompted the application of a more advanced machine learning method. A random forest model was implemented because of its balance between robust prediction performance and interpretability, using the package "randomForest" in the R project for statistical computing (Version 3.6.0) (www.r-project.org). Mean Decreases in Accuracy (MDA) were computed to measure the importance level of each variable in the RF model. Partial plots were also provided for selected variables to illustrate the relationship between these variables and the neurocognitive status of patients.

Results

Study groups Of the 370 patients recruited, approximately two-thirds ($n = 246$) performed neuropsychologically within the normal ranges for their age, sex, and education, based on published normative data and thus were deemed neurocognitively normal (NN) (Table 1). Among patients who displayed mean neurocognitive deficits (<− 1.0 SD) in two or more neuropsychological domains, two groups with NCI were detected: most patients met the criteria for HAND ($n = 78$, 21%) with a predominance of MND (51%) compared to ANI (36%) and HAD (14%). A second group of patients were identified ($n = 46$, 12.5%) that displayed neurocognitive impairment due to other comorbid disorders (NCI-OD), which became evident during the clinical assessment and NP testing (e.g., mood or anxiety disorders, unrecognized developmental disorders). Notably, most NCI-OD patients (84%) reported neurocognitive symptoms.

Univariate analyses Comparisons of associated demographic and clinical variables among the three groups (NN, HAND and NCI-OD) revealed that multiple variables differed significantly between NN and HAND including birth continent, present and past alcohol binging, *Toxoplasma gondi* seropositivity, sexual orientation, and documented seroconversion illness (Table 1). The NN and NCI-OD groups differed significantly for the mean PHQ-9 score, past cocaine use, birth continent, minor traumatic head injury, hours of sleep per night, HQoL and reported cART side effects (Table 1). The HAND and NCI-OD groups differed significantly for PHQ-9 score, quality of life score, and history of cocaine use (Table 1).

Multiple logistic regression We next applied multiple logistic regression analysis to the data with all 37 variables for comparisons of NN versus HAND and NN versus NCI-OD. For the comparison of the NN versus HAND groups, the McFadden's R square was 24% and the fivefold cross-validation classification accuracy was 74%. Variables differing significantly in this analysis included birth continent, CD4 T cell nadir count, and current and past alcohol binging (Table 1). For the NN versus NCI-OD comparisons, the McFadden's R square was 32% and the fivefold cross-validation classification accuracy was 72%. Significant variables included PHQ9, VACS index, history of cocaine use, traumatic brain injury, employment, quality of life score and birth continent, and hours of sleep per night (Table 1). Finally, the McFadden's R square was 45% for the HAND versus NCI-OD groups, and the fivefold cross-validation

classification accuracy was 63%. The significant variables included PHQ-9 score, VACS index, history of cocaine use, recent viral load, quality of life, and current alcohol binging. Thus, intergroup comparisons differed for the individual variables that distinguished the three groups.

Principal component analysis The predictive performance of the multiple logistic regression model (Supplementary Table 1) provided evidence for the possibility of using the collected 37 variables to predict NCI status although there was limited prediction accuracy. To refine the performance of the logistic regression model, we conducted a principal component analysis (PCA) using all binary and continuous variables among all patients, and the first two components accounted for 20% of the variation (Fig. 1). If only continuous variables (12 out of 37) were used in the PCA, the first two components accounted for 30% of the variation. Notably, the variables were scaled before performing the PCA. The cohort failed to show dominant subgrouping in the PCA which precluded linear separation of the patients based on NCI status. This finding highlighted the limitations of the linear logistic regression modeling and prompted consideration of a more elaborate machine learning approach.

Random forest analyses Random forest models were used to classify NN from HAND, NN from NCI-OD, and the HAND from NCI-OD groups. Fivefold cross-validation (CV) was used for the parameter tuning (number of trees, try, and node size) and the Synthetic Minority Over-sampling Technique (SMOTE) [28] algorithm was applied to the training set to address the issue of imbalanced sample size present in the data. Accuracy and area under the curve (AUC) value were computed using fivefold CV in the following analyses. The final model was obtained using the entire data set with the parameters selected by CV. The accuracy was 81% and the area under the curve (AUC) value was 87% (Fig. 2a) for the NN versus HAND analysis. Of note, AUC was a better measure for unbalanced data sets as an accuracy of (246/324 = 75.9%) was achieved by classifying each patient into the NN group. The high AUC value indicated the RF model robustly discriminated between the NN and HAND groups (Supplementary Table 1). Similarly, for the NN versus NCI-OD groups, the accuracy was 84% (Supplementary Table 1) and the AUC value was 77% (Fig. 2b). For the HAND versus NCI-OD groups, the accuracy was 71% (Supplementary Table 1) and the AUC value was 73% (Fig. 2c). In summary, the RF model performed optimally for the comparison of NN versus HAND groups with the highest AUC value. Another advantage of the random forest model compared with logistic regression was its ability to deal with multicollinearity within the data due to the random selection of variables at each node creation.

To explore which variables were more important in the RF models, we computed the MDA values for each variable. The MDA reflected the relative loss of accuracy by the random permutation of one variable. The absolute values of MDA were non-quantitative but the rankings of the MDA values highlight which variables were more important in differentiating the two compared groups. For the NN versus HAND comparison, recent CD4+ T cell, current alcohol use, polypharmacy, hours of work per week, nadir CD4+ T cell counts, hours of sleep per night, age, PHQ-9 score, and VACs index are among the top nine variables identified in the relative importance plot (Fig. 3a). The full numeric analyses elaborated all the different variables' relative importance (Supplementary Table 2).

Partial plots were examined to assess the relationships between the chief variables and their response. The *y*-axis of the partial plots indicated the relative logit contribution of the variables on the probability of being classified in the NN group from the perspective of the RF model. For example, negative values indicate that the positive class is less likely for that value of the independent variable according to the model. Nine selected variables for the NN versus HAND comparisons illustrate threshold values for the likelihood of being the NN group including recent CD4+ T cell counts (Fig. 4a), current alcohol use (Fig. 4b), CPE rank score (Fig. 4c), nadir CD4+ T cell counts (Fig. 4d), hours worked per week (Fig. 4e), peak viral load (Fig. 4f), polypharmacy (Fig. 4g), PHQ-9 score (Fig. 4h), and birth continent (Fig. 4i).

For the NN versus NCI-OD comparison, the pattern of MDA distribution was less than the NN versus HAND

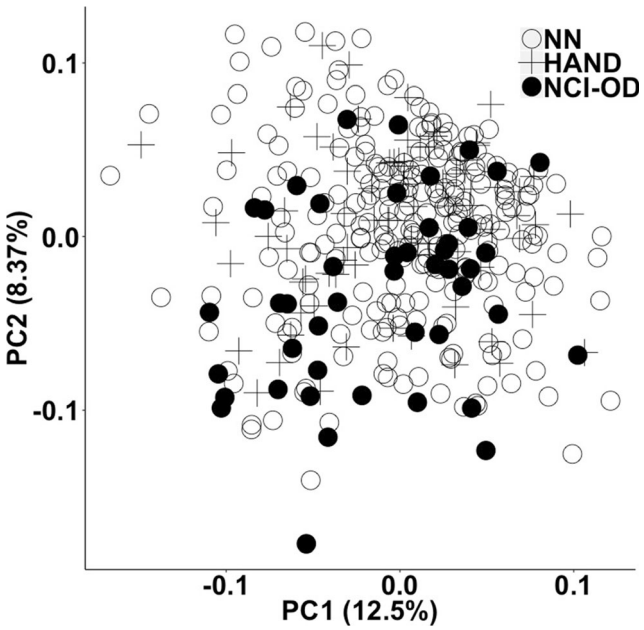

Fig. 1 Principal component analysis of cohort. The plot shows the first two principal components, which accounted for about 20% of variation within the entire cohort. Each point represents a single patient with different shapes indicating individual NCI status (NN, o; HAND, ⊕; NCI-OD, •)

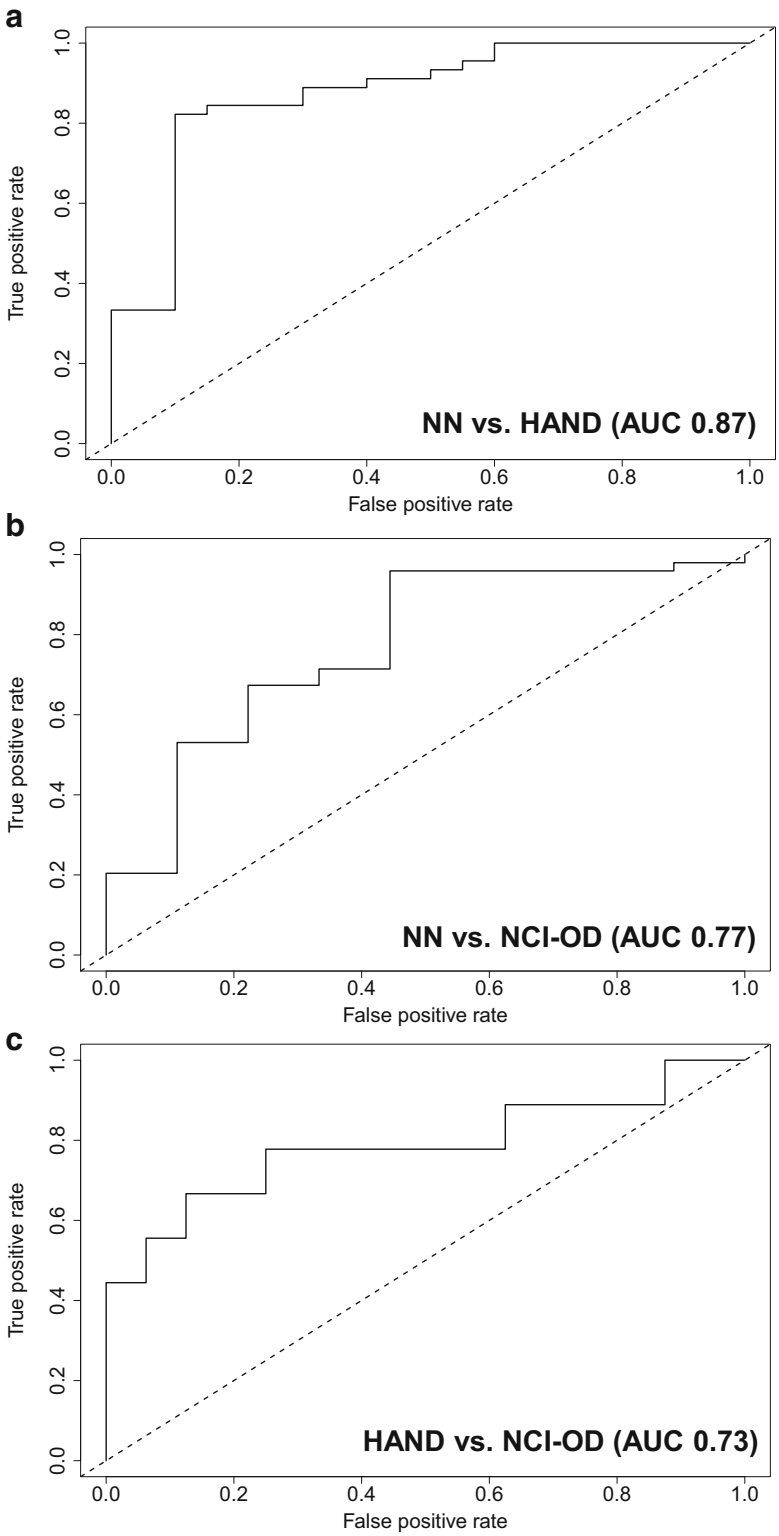

Fig. 2 Receiver operating characteristic (ROC) curves. **a** The NN versus HAND ROC curve showed an AUC value = 0.87. **b** The NN versus NC-OD ROC curve showed an AUC value = 0.77. **c** The HAND versus NC-OD ROC curve showed an AUC value = 0.73

because there were fewer dominant variables that were more important (Fig. 3b). Nonetheless, this comparison revealed specific discriminating variables: PHQ-9 score, recent viral load, hours of sleep per night, HQoL, birth continent, education, duration of HIV, age, and hours of work per week with the full numeric analysis also provided (Supplementary Table 2). Partial plots emphasized the differences between groups (Fig. 5). Finally, for the HAND versus NCI-OD relative importance plot, the variables at the top included quality of life, age, PHQ-9, and polypharmacy (Supplementary Fig. 1) with full numeric analyses (Supplementary Table 2) and representative partial plots (Supplementary Fig. 2). Thus, RF modeling yielded a robust

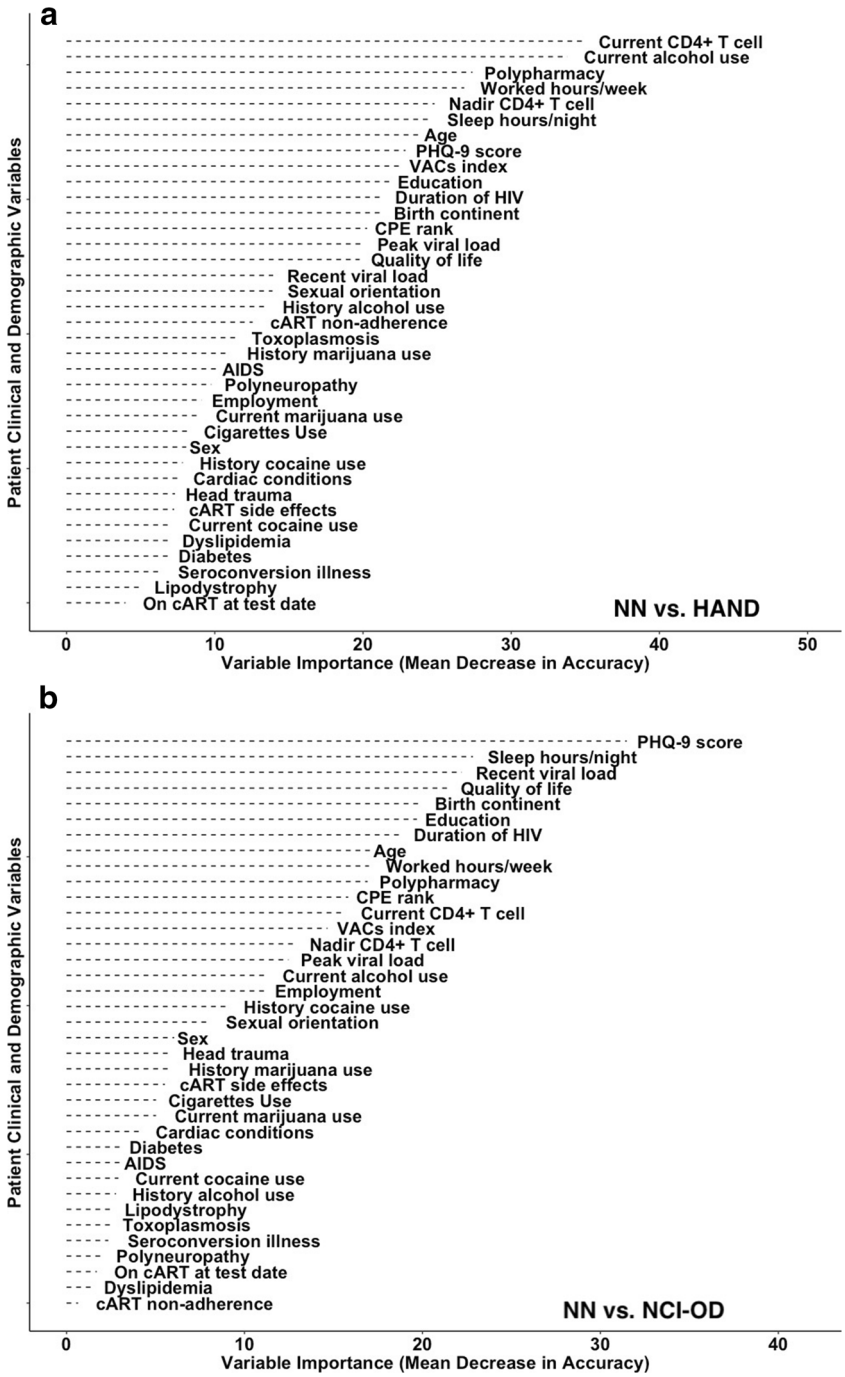

Fig. 3 Relative variable importance plots. The relative variable importance plots for NN versus HAND depicts the MDA values for both continuous and categorical cofactors that contributed to the differences between **a** NN versus the HAND and **b** NN versus the NCI-OD groups

differentiation between NN and HAND predicated on variables displaying a distinctive MDA profile relative to the NN versus NCI-OD or HAND versus NCI-OD comparisons.

Conclusions

To the best of our knowledge, this study represents the first analyses of prevalence and variables associated with common categories of NCI within a defined population of HIV-infected persons receiving modern cART comparing logistic regression and machine learning methodologies. While the overall prevalence of NCI was 33.5% and most of the patients were diagnosed with HAND, another group with NCI due to other causes was detected (NCI-OD). As expected, several variables were identified that distinguished the neurocognitively normal group from each of the NCI groups, evidenced by the univariate, multiple logistic regression and random forest analyses. Notably, the machine learning approach identified several distinguishing variables among groups that were not detected

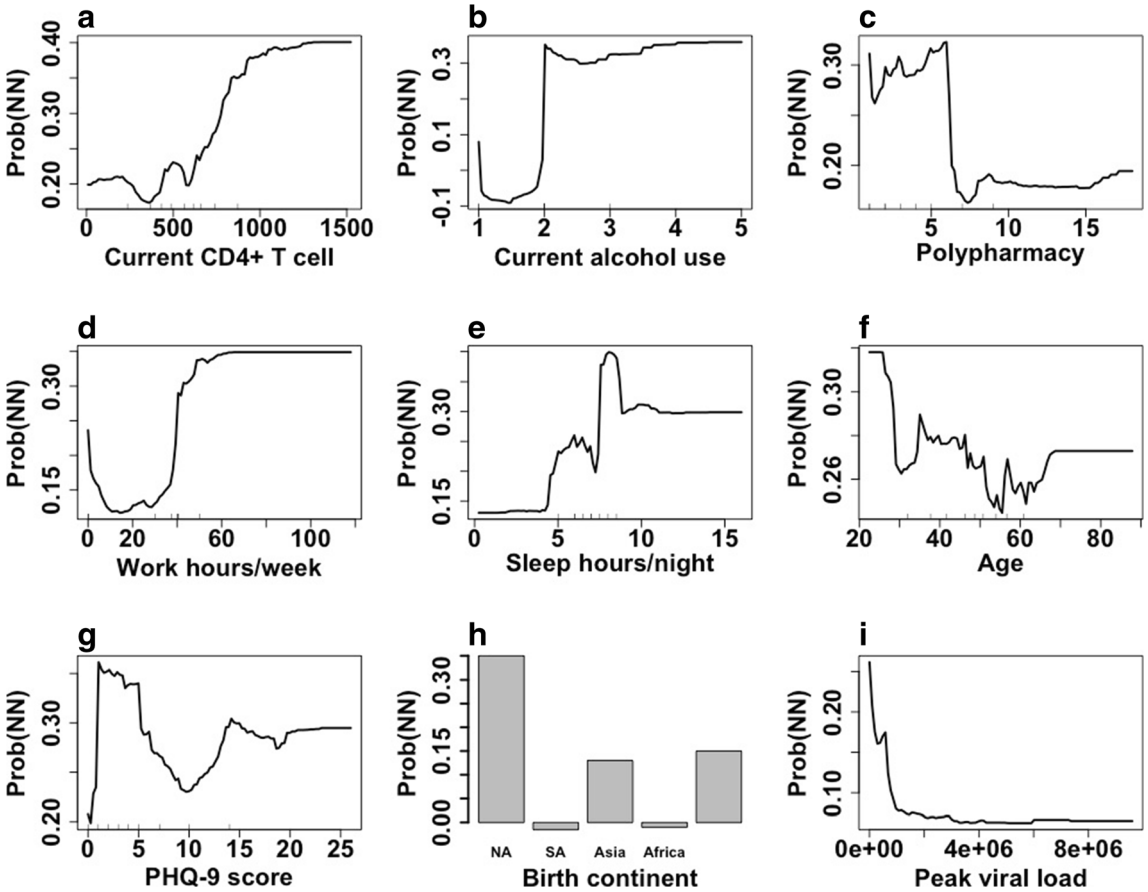

Fig. 4 Partial plots for the major importance variables for NN versus HAND groups. Both continuous and categorical variables disclosed the probability of being in the NN (Prob(NN) group versus the NCI-OD group based on threshold cofactor value for **a** current CD4+ T cell, **b** current alcohol use, **c** polypharmacy, **d** worked hours/week, **e** sleep hours/night, **f** age, **g** PHQ-9, **h** birth continent, and **i** peak viral load

in the logistic regression analyses. To date, machine learning, particularly random forest modeling, has not been applied to understanding the variables contributing to HAND. The present studies suggested that multiple demographic and clinical factors underlie the HAND diagnosis and are different from those associated with NCI-OD based on both the multiple logistic analyses (Table 1) as well as the RF modeling (Figs. 2 and 3). Indeed, the random forest model exhibited robust predictability for HAND (Supplementary Table 1) but was less efficient for NCI-OD, likely due to the greater heterogeneity within this latter group. A key outcome of the present studies was an appreciation of the growing diversity of variables associated with the development of HAND, despite implementing defined guidelines for its diagnosis.

Application of machine learning methods to dissect the different variables has been used previously to analyze other neurocognitive diseases with success but not for HAND versus other diagnostic entities (Gray et al. 2013; Patel et al. 2015; Boucekine et al. 2013). In the present study, logistic regression and machine learning approaches were complementary in that logistic regression measured the effects of individual cofactors while machine learning provided insights into the predictive properties of different cofactors. In fact, the selected variables were predictive of HAND or NCI-OD, based on the AUC values (Fig. 2) while the partial plots (Figs. 4 and 5) yielded insights into the threshold values for belonging to the NN group, which might be of clinical value. Nonetheless, counterintuitive findings emerged including the contribution of increased alcohol consumption in the NN versus HAND groups, which likely reflects the greater overall wellbeing of the NN group and perhaps prior substance use by the NCI groups. At the same time, variables such as current CD4+ T cell counts, polypharmacy, and age ranked highly in the MDA analyses of NN versus HAND but were identified by the multiple logistic regression model emphasizing the divergent results depending on the specific statistical application used. Multinomial logistic regression was not performed here because it was not germane to the overarching goal of the study although future analyses could address this question.

Neurocognitive dysfunction was recognized at the outset of the HIV epidemic, chiefly as HIV-associated dementia (Snider et al. 1983; McArthur 1987). With increasing availability of cART, there has been a greater appreciation of subtler neurocognitive deficits among patients with HIV-1 infection,

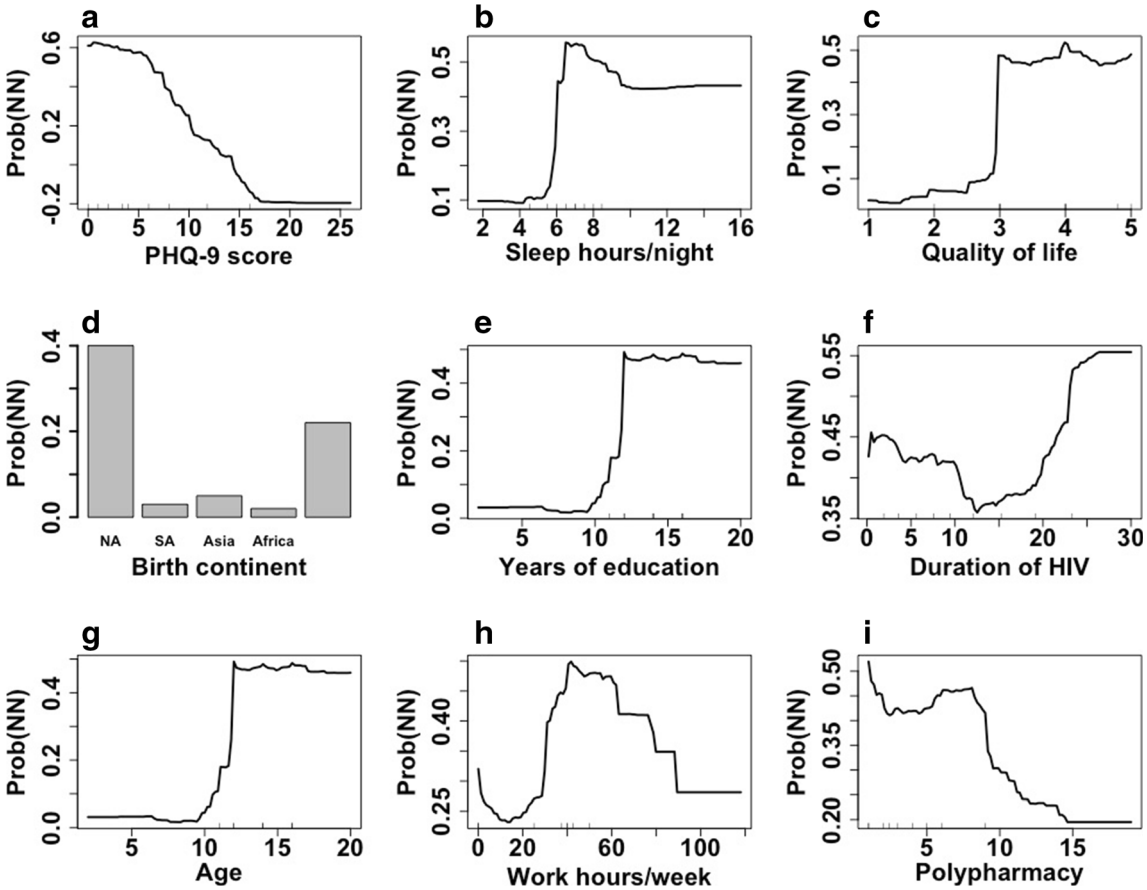

Fig. 5 Partial plots for the major importance variables for NN versus NCI-OD groups. Both continuous and categorical variables disclosed the probability of being in the NN (Prob(NN) group versus the NCI-OD group based on threshold cofactor value for **a** PHQ-9, **b** estimated hours of sleep per night, **c** HQoL value, **d** birth continent, **e** years of education, **f** duration of HIV-1 infection, **g** age, **h** hours worked per week, and **i** polypharmacy

collectively designated as HAND. There has been variation in how the criteria for HAND have been applied, particularly in the integration of comorbidities and emerging factors (e.g., aging, off-target effects of cART) that have prompted discussion within the literature (Gisslen et al. 2011; Spudich and Ances 2017). Several variables associated with developing symptomatic HAND in our previous studies (Vivithanaporn et al. 2010; McCombe et al. 2013) were similar to those identified in the present study including birth continent and nadir CD4+ T cell in blood that emerged in the random forest analyses. Substantial efforts have been invested in delineating the phenotypes and cut-off values for neuropsychological tests as well as the associated cofactors that underpinned HIV-associated NCI (Tierney et al. 2017; Su et al. 2015; Kamminga et al. 2017). While mood disorders are frequent among patients with HIV-1 infection, they do not account for the full extent of NCI in HIV-infected populations (Tymchuk et al. 2017). Efforts to encompass all these variables have prompted different definitions of NCI in people with HIV-1 infection together with implementation of a diversity of analytical tools and approaches. These range from population-based studies to selected patient cohorts (Rubin et al. 2017; Carvalhal et al. 2016), each with differing outcomes that are influenced by statistical approaches and evolving clinical circumstances.

Several challenges emerged in the present study. With strict use of diagnostic definitions and exclusion criteria, the proportion of patients with HAND within our cohort was only 21%. This limited group size constrained the type and extent of analyses that could be performed. Nonetheless, the current cohort was distinct from other reports in that there was a predominance of patients with HIV-associated mild neurocognitive disorder (MND); other studies have reported that asymptomatic neurocognitively impaired (ANI) patients were the largest group in the HAND category (Heaton et al. 2010; Carvalhal et al. 2016). This difference might reflect intensive questioning of patients and caregivers about neurocognitive symptoms at multiple stages (nurse, psychometrist, and physician) of the care delivery process within the present study. Another limitation to the present study is the lack of a definitive demarcation between the HAND versus NCI-OD groups based on ROC and the RF

analyses. The relative lack of distinction between these groups likely arises due to the small sample sizes of the HAND and NCI-OD groups. In fact, the poor accuracy performance might reflect the smaller sample sizes of HAND and NCI-OD groups while also indicating an overlap in some of the clinical and demographic factors within these two groups. Furthermore, the demographic differences within the cohort were also a challenge; for example, if the variable was sex with only 10% of the cohort as female then the MDA is related to the average loss of accuracy by randomly choosing 10% of the data points as female, which implicates an inherent bias in the data. Thus, comparative application of logistic regression and machine learning approaches to demographically different cohorts would provide validation to the present findings.

Although NCI remains a common and debilitating disorder among HIV-infected patients receiving contemporary care, the delineation of individual NCI phenotypes remains incomplete. This uncertainty might differ depending on geographic, ethnic, and clinical practice factors as well as individual disease course trajectories. Larger comparative studies that dissect specific neurocognitive phenotypes and their causal factors are essential to address these uncertainties. Moreover, the impact of specific cART drugs and multidrug regimens require deeper analyses that can only be achieved with larger and perhaps collaborative studies. Notwithstanding, the use of paired approaches such as logistic regression and machine learning generates new and valuable insights into the cofactors associated with NCI in HIV-infected persons that are potentially modifiable and amenable to treatment.

Acknowledgments A Canadian Institutes of Health Research, Emerging Team Grant (EF, MJG, and CP) supported these studies. The authors thank the patients and staff at the Southern Alberta Clinic for their willingness to participate in the study and Brenda Beckthold for the technical assistance.

Data availability All data presented within the present manuscript are available with accompanying accession numbers upon request to qualified investigators for secondary analyses.

Compliance with ethical standards

The University of Calgary Ethics Committee (REB #-130615) approved the study and written consent was obtained from all patients.

Conflict of interest The authors declare that they have no conflict of interest.

References

Antinori A, Arendt G, Becker JT, Brew BJ, Byrd DA, Cherner M, Clifford DB, Cinque P, Epstein LG, Goodkin K, Gisslen M, Grant I, Heaton RK, Joseph J, Marder K, Marra CM, McArthur JC, Nunn M, Price RW, Pulliam L, Robertson KR, Sacktor N, Valcour V, Wojna VE (2007) Updated research nosology for HIV-associated neurocognitive disorders. Neurology. 69(18):1789–1799

Belete T, Medfu G, Yemiyamrew E (2017) Prevalence of HIV associated neurocognitive deficit among HIV positive people in ethiopia: a cross sectional study at Ayder Referral Hospital. Ethiop J Health Sci 27(1):67–76

Boucekine M, Loundou A, Baumstarck K, Minaya-Flores P, Pelletier J, Ghattas B, Auquier P (2013) Using the random forest method to detect a response shift in the quality of life of multiple sclerosis patients: a cohort study. BMC Med Res Methodol 13:20

Brandt J, Benedict R (2001) Hopkins verbal learning test-revised (HVLT-R). Psychological Assessment Resources, Lutz

Carvalhal A, Gill MJ, Letendre SL, Rachlis A, Bekele T, Raboud J et al (2016) Central nervous system penetration effectiveness of antiretroviral drugs and neuropsychological impairment in the Ontario HIV Treatment Network Cohort Study. J Neuro-Oncol 22(3):349–357

Crane HM, Van Rompaey SE, Dillingham PW, Herman E, Diehr P, Kitahata MM (2006) A single-item measure of health-related quality-of-life for HIV-infected patients in routine clinical care. AIDS Patient Care STDs 20(3):161–174

Davis LE, Hjelle BL, Miller VE, Palmer DL, Llewellyn AL, Merlin TL, Young SA, Mills RG, Wachsman W, Wiley CA (1992) Early viral brain invasion in iatrogenic human immunodeficiency virus infection. Neurology. 42(9):1736–1739

Delis DC, Kaplan E, Kramer JH, D-KEFSSA (2001) TX: the psychological corporation. Delis-Kaplan executive function system. The Psychological Corporation, San Antonio

Gisslen M, Price RW, Nilsson S (2011) The definition of HIV-associated neurocognitive disorders: are we overestimating the real prevalence? BMC Infect Dis 11:356

Gomez D, Power C, Gill MJ, Fujiwara E (2017) Determinants of risk-taking in HIV-associated neurocognitive disorders. Neuropsychology. 31(7):798–810

Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D (2013) Alzheimer's disease neuroimaging I. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. Neuroimage 65:167–175

Heaton RK, Clifford DB, Franklin DR Jr, Woods SP, Ake C, Vaida F et al (2010) HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. Neurology. 75(23):2087–2096

Kamminga J, Bloch M, Vincent T, Carberry A, Brew BJ, Cysique LA (2017) Determining optimal impairment rating methodology for a new HIV-associated neurocognitive disorder screening procedure. J Clin Exp Neuropsychol 39(8):753–767

Kim D, Jewison DL, Milner GR, Rourke SB, Gill MJ, Power C et al (2001) Can J Neurol Sci 28:228–231

Kongs SK, Thompson LL, Iverson GL, Heaton RK (2000) Wisconsin card sorting test-64 card version (WCST-64) Lutz, FL:PAR

Kroenke K, Spitzer RL, Williams JB (2001) The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16(9):606–613

McArthur JC (1987) Neurologic manifestations of AIDS. Medicine (Baltimore) 66(6):407–437

McCombe JA, Vivithanaporn P, Gill MJ, Power C (2013) Predictors of symptomatic HIV-associated neurocognitive disorders in universal health care. HIV Med 14(2):99–107

Mothobi NZ, Brew BJ (2012) Neurocognitive dysfunction in the highly active antiretroviral therapy era. Curr Opin Infect Dis 25(1):4–9

Nakagawa F, May M, Phillips A (2013) Life expectancy living with HIV: recent estimates and future implications. Curr Opin Infect Dis 26(1):17–25

Patel A, Parikh R, Howell EH, Hsich E, Landers SH, Gorodeski EZ (2015) Mini-cog performance: novel marker of post discharge risk among patients hospitalized for heart failure. Circ Heart Fail 8(1):8–16

Pettersen JA, Jones G, Worthington C, Krentz HB, Keppler OT, Hoke A, Gill MJ, Power C (2006) Sensory neuropathy in human immunodeficiency virus/acquired immunodeficiency syndrome patients: protease inhibitor-mediated neurotoxicity. Ann Neurol 59(5):816–824

Puhan MA, Van Natta ML, Palella FJ, Addessi A, Meinert C (2010) Ocular complications of ARG. Excess mortality in patients with AIDS in the era of highly active antiretroviral therapy: temporal changes and risk factors. Clin Infect Dis 51(8):947–956

Robertson K, Liner J, Heaton R (2009) Neuropsychological assessment of HIV-infected populations in international settings. Neuropsychol Rev 19(2):232–249

Roehr B (2015) UNAIDS celebrates success in halting spread of HIV and sets goals for 2030. BMJ. 351:h3832

Rubin LH, Maki PM, Springer G, Benning L, Anastos K, Gustafson D, Villacres MC, Jiang X, Adimora AA, Waldrop-Valverde D, Vance DE, Bolivar H, Alden C, Martin EM, Valcour VG, For the Women's Interagency HIV Study (2017) Cognitive trajectories over 4 years among HIV-infected women with optimal viral suppression. Neurology. 89(15):1594–1603

Sacktor N, Skolasky RL, Seaberg E, Munro C, Becker JT, Martin E, Ragin A, Levine A, Miller E (2016) Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study. Neurology. 86(4):334–340

Saylor D, Dickens AM, Sacktor N, Haughey N, Slusher B, Pletnikov M, Mankowski JL, Brown A, Volsky DJ, McArthur JC (2016) HIV-associated neurocognitive disorder–pathogenesis and prospects for treatment. Nat Rev Neurol 12(4):234–248

Schretlen DJ, Testa SM, Pearlson GD (2010) GD. Clock-drawing test scoring approach from the Calibrated Neuropsychological Normative System. Lutz, FL: Psychological Assessment Resources

Smail RC, Brew BJ (2018) HIV-associated neurocognitive disorder. Handb Clin Neurol 152:75–97

Smith A (1973) Symbol digit modalities test. Western Psychological Services, Los Angeles

Snider WD, Simpson DM, Nielsen S, Gold JW, Metroka CE, Posner JB (1983) Neurological complications of acquired immune deficiency syndrome: analysis of 50 patients. Ann Neurol 14(4):403–418

Spudich SS, Ances BM (2017) CROI 2017: neurologic complications of HIV infection. Top Antivir Med 25(2):69–76

Su T, Schouten J, Geurtsen GJ, Wit FW, Stolte IG, Prins M, Portegies P, Caan MW, Reiss P, Majoie CB, Schmand BA, AGEhIV Cohort Study Group (2015) Multivariate normative comparison, a novel method for more reliably detecting cognitive impairment in HIV infection. AIDS. 29(5):547–557

Tierney SM, Sheppard DP, Kordovski VM, Faytell MP, Avci G, Woods SP (2017) A comparison of the sensitivity, stability, and reliability of three diagnostic schemes for HIV-associated neurocognitive disorders. J Neuro-Oncol 23(3):404–421

Trites RL (1977) Neuropsychological test manual. Royal Ottawa Hospital, Ottawa

Tymchuk S, Gomez D, Koenig N, Gill MJ, Fujiwara E, Power C (2017) Associations between depressive symptomatology and neurocognitive impairment in HIV/AIDS. Can J Psychiatr 706743717737029

van den Dries LWJ, Wagener MN, Jiskoot LC, Visser M, Robertson KR, Adriani KS, van Gorp ECM (2017) Neurocognitive impairment in a chronically well-suppressed HIV-infected population: the Dutch TREVI Cohort Study. AIDS Patient Care STDs 31(8):329–334

Vivithanaporn P, Heo G, Gamble J, Krentz HB, Hoke A, Gill MJ, Power C (2010) Neurologic disease burden in treated HIV/AIDS predicts survival: a population-based study. Neurology. 75(13):1150–1158