

Novel Machine Learning Identifies Brain Patterns Distinguishing Diagnostic Membership of HIV, Alcoholism, and Their Comorbidity of Individuals

Ehsan Adeli¹, Natalie M. Zahr^{1,2}, Adolf Pfefferbaum^{1,2}, Edith V. Sullivan¹, Kilian M. Pohl^{*2}

¹Department of Psychiatry and Behavioral Sciences,
Stanford School of Medicine, Stanford, CA 94305

²Center for Biomedical Sciences, SRI International, Menlo Park, CA 94025

Abstract

Objective: The incidence of Alcohol Use Disorder (AUD) in human immunodeficiency virus (HIV) infection is twice that of the rest of the population. This study documents complex radiologically-identified, neuroanatomical effects of AUD+HIV comorbidity by identifying structural brain systems that predicted diagnosis on an individual basis.

Method: Applying novel machine learning analysis to 549 participants (199 controls, 222 AUD, 68 HIV, 60 AUD+HIV), 298 MRI brain measurements were automatically reduced to small subsets per group. Significance of each *diagnostic pattern* was inferred from its accuracy in predicting diagnosis and performance on 6 cognitive measures.

Results: While all three *diagnostic patterns* predicted the learning and memory score, the AUD+HIV pattern was the largest and had the highest predication accuracy (78.1%).

Conclusions: Providing a roadmap for analyzing large, multimodal data sets, the machine learning analysis revealed imaging phenotypes that predicted diagnostic membership of MRIs of individuals with AUD, HIV, and their comorbidity.

1. Introduction

Alcohol Use Disorder (AUD) is common (1), and its comorbidity in individuals with human immunodeficiency virus (HIV) infection is high (2-4), occurring at a rate twice that of the general population (5). AUD and HIV-infection each disrupts brain structural integrity with the likely outcome of reducing health-related quality of life and cognition (3, 4). AUD targets, among other regions, frontal cortices (6-8) and cerebellum (9, 10). HIV similarly compromises frontal cortices, but also cingulate and parietal regions, e.g., (11). AUD and HIV are independently associated with volume deficits in thalamus, hippocampus, caudate and putamen (e.g., AUD: (4, 10, 12); HIV: (4, 11, 13)). Relatively few brain studies have examined the heightened burden of disease comorbidity (2, 14), which has the potential to exacerbate the untoward effects on neural systems through synergistic or additive processes (15, 16). HIV+AUD show moderate to severe abnormalities especially in frontal cortices and thalamus (17).

One possible solution to enhance understanding of the complex neurological effects of AUD+HIV comorbidity is to encode the architecture of the whole brain of an individual through large numbers of measures extracted from a fine-grained parcellation of brain regions. To date, however, morphometric brain studies on AUD and HIV have relied on univariate testing of relatively few MRI metrics that are separately related to diagnostic groups (2, 3, 11, 14-16, 18-21). Alternatively, machine learning approaches (22-25) can jointly analyze a large number of metrics by combining them into a single score. To highlight the power of such multivariate analysis (22, 26) for expanding knowledge about neuropsychiatric disorders, we derived *diagnostic scores* to predict diagnosis from MRIs collected in individuals with AUD, HIV, or their comorbidity.

To test the hypothesis that MRIs can be used to predict diagnosis and cognitive measures in individuals with AUD, HIV, or their comorbidity, we first improved on a technology called sparse classification (27-29). We then applied the corresponding novel machine learning analysis to a

dataset of 549 MRIs of controls and individuals diagnosed with one of the three disorders, taken from our previous report on 30 regional volumes (17). Each MRI was now quantified in terms of 298 brain regional metrics of volume, surface area, thickness, and curvature. For each disorder, our data-driven approach first identified a *diagnostic pattern* by automatically reducing the large number of MRI metrics to a small subset affected by the disorder. The subset of metrics was then applied to each individual MRI to compute a *diagnostic score* to predict diagnosis and cognitive ramifications in the corresponding participant. By doing so, our machine learning analysis not only identified structural differences from controls in AUD individuals (as in (30)) but also in HIV and their comorbidity.

2. Materials and Methods

2.1 Participants

The four groups comprised 199 healthy controls (CTRL), 222 AUD individuals, 68 HIV-infected individuals (HIV), and 60 subjects with both AUD and HIV infection (AUD+HIV). Participants ranged in age between 25 – 75 years.

AUD participants were screened to meet DSM-IV criteria for Alcohol Dependence or Abuse and DSM5 criteria for AUD, to have ≥ 10 years of heavy drinking, and habitually consume ≥ 150 drinks a month for men or ≥ 90 for women. The study recorded their Days After Last Drink (DALD), total lifetime alcohol consumed in kilograms (Alc Kg), and alcohol consumed in the past year (Alc py).

Participants in non-AUD groups (e.g., HIV and CTRL) did not meet DSM-IV criteria for Alcohol Dependence or Abuse or DSM5 for AUD. HIV subjects were seropositive for the HIV-infection with CD4 count $> 100 \frac{\text{cells}}{\mu\text{L}}$ (average: 303.0) and had a Karnofsky score ≥ 70 (31). Participants in the CTRL group had never met DSM-IV or DSM5 criteria for any neuropsychiatric disorder and

tested negative for HIV infection. The two HIV groups had higher mean Veterans Aging Cohort Study (VACS) Indices (32) than either the control or AUD groups (Table 1). Note that VACS defines a score based on pre-assigned points for age, HIV indicators (CD4 count and HIV-1 RNA), and general indicators of organ system injury (see Supplement A for details).

Participants completed neuropsychological tests assessing six cognitive, motor, and social functional domains. Composite scores of each domain were derived from age- and education-corrected z-scores based on control performance: Verbal Language (VL), Executive Function (EXF), Learning and Memory (LM), Speed of Information Processing (SIP), Motor Skills (MS), and Quality of Social Functioning (QSF). Table 1 lists their mean \pm SD as well as the medication history of participants (see Supplement A for detail).

The neuropsychological scores were based on the means of composite scores representing 6 functional domains (cf., 33, 34) described previously (Supplemental Material 17). Raw scores from tests included in each composite score were age-corrected based on 66 male and 85 female healthy controls, aged 20 to 67 at their first examination and expressed as standardized z-scores. All metrics (e.g., speed scores such as Trails A and B) were transformed so that higher scores were in the direction of better performance. Each of the composite scores was the mean of the z-scores of available test measures for each participant: **EXF** comprised Trails B (35) or Color Trails 2 (36) time, Wechsler Memory Scale-Revised (WMS-R) (37) or MicroCog (38) forward and backward digit span, and the Golden Stroop Color Word raw score (39); **LM** comprised the Rey-Osterrieth Complex Figure Test immediate recall raw score (40) and WMS-R Logical Memory (immediate recall total raw score) (37) or MicroCog Memory (immediate recognition score) (38); **VL** comprised FAS letter fluency total score (41) and National Adult Reading Test (42), Peabody (43), or Wechsler Test of Adult Reading (44) total score; **SIP** comprised Trails A (35) or Color Trails 1 (36) time, Digit Symbol (45) or Symbol Digit (46) raw score, and Golden Stroop Color raw score (40); **MS** comprised Grooved Pegboard mean of left

and right hand scores (47), Fine Finger Movement mean of all conditions (48), and Ataxia mean score of standing on the left and right legs separately (49); **QSF** comprised Quality of Life SF-21 total raw score (50), Global Assessment of Functioning score (current) (51), and Activities of Daily Living (combined Performance and Instrumental scores) (52). Participants underwent different cognitive tests for some domains as some tests were replaced during the longitudinal study.

2.2 MRI Data Acquisition and Preprocessing

Imaging data were acquired on a 3T General Electric (GE; Waukesha, WI) SIGNA system using an 8-channel Array Spatial Sensitivity Encoding Technique (ASSET) coil for parallel and accelerated imaging. Inversion Recovery-SPOiled Gradient Recalled (IR-SPGR) echo sequence (TR=7.068ms, TI=300ms, TE = 2.208ms, flip angle=15°, matrix=256 × 256, slice dimensions=1.25 × 0.9375 × 0.9375mm, 124 slices) were collected in the sagittal plane.

Processing of a T1-weighted (T1w) MR image (see Supplement B for detail) resulted in the supratentorial volume (svol) according to the SRI24 atlas (53) and the z-scores of 298 morphometric measurements extracted by Freesurfer (54-56). Note that baseline volumetric MRI data of the 549 participants were previously published (17) but were derived solely from 30 ROIs of the SRI24 atlas (57) rather than the FreeSurfer atlas used herein.

The morphometric measurements of the CTRL group varied significantly with age, sex, and svol (Pearson correlation p -value < 0.005). These confounding factors were regressed out from the morphometric measurements by parameterizing a general linear model (58) on the controls of the training data. Details can be found in Supplement B.

Table 1: Demographic information and the statistics of the cognitive and clinical measures (mean ± standard deviation) for each group. Each of the composite scores was the mean of available test measures for each participant (refer to the text for details). Group differences are measured for each diagnosis group with respect to the CTRL group by χ^2 -test (for sex, ethnicity, HAART medication, and AIDS status) and t-test (for other measures), and are considered not significant (NS) if p-value > 0.05.

Measure	CTRL	AUD	p-value (vs. CTRL)	HIV	p-value (vs. CTRL)	AUD+HIV	p-value (vs. CTRL)	Pair-wise group differences*
Total Subjects	199	222	-	68	-	60	-	-
Demographic								
Sex F/M (F% / M%)	92/107 (46% / 54%)	66/156 (30% / 70%)	4.8108 x 10 ⁻⁴	21/47 (31% / 69%)	0.0017	22/38 (37% / 63%)	< 0.00001	$\chi^2=13.4$, p=0.004
Age (years)	46.7±14.2	48.4±9.9	NS	51.4±8.7	0.0102	51.7±6.9	0.0118	CTRL = AUD < HIV = AUD+HIV
Education (years)	16.0±2.3	13.4±2.4	< 0.00001	13.5±2.4	< 0.00001	13.0±2.1	< 0.00001	CTRL > AUD = HIV = AUD+HIV
Socioeconomic status (lower is better)	25.5±11.6	40.9±14.4	< 0.00001	40.7±14.2	< 0.00001	45.2±12.2	< 0.00001	CTRL < AUD = HIV = AUD+HIV
Body mass index	25.9±4.2	26.8±4.8	0.0408	26.6±4.7	NS	26.8±4.9	NS	CTRL = AUD = HIV = AUD+HIV
Ethnicity Asian/African American/Caucasian/Other	28/28/127/16	4/71/117/30	< 0.00001	0/31/34/3	< 0.00001	0/38/17/5	< 0.00001	$\chi^2=97.80$, p<0.0001
Clinical								
Days After Last Drink (DALD)	-	196.9±507.9	-	-	-	398.9±1126.8	-	AUD=AUD+HIV
Total Alcohol Consumed (Alc Kg)	34.0 ±57.0	1206.2±885.7	7.4523 x 10 ⁻⁴⁰	110.5±240.7	NS	1081.0±916.1	4.9219 x 10 ⁻²⁸	CTRL = HIV < AUD = AUD+HIV
Alcohol consumed in past year	-	42.2±45.7	-	-	-	14.0±17.7	-	AUD > AUD+HIV
VACS Index	13.8±12.3	15.0±12.7	NS	27.3±18.1	4.6496 x 10 ⁻⁷	32.8±22.5	5.1148 x 10 ⁻⁹	CTRL = AUD < HIV = AUD+HIV
CD4 cell count (100/mm ³)	-	-	-	303.0±188.7	-	278.9±216.5	-	HIV = AUD+HIV
Nadir CD4 (100/mm ³)	-	-	-	202.0±176.4	-	208.5±174.4	-	HIV = AUD+HIV
Viral Load	-	-	-	2.13±1.14	-	2.24±1.18	-	HIV = AUD+HIV
Percentage on HAART Medication	-	-	-	88%	-	87%	-	HIV = AUD+HIV ($\chi^2=0.0457$)
Percentage with AIDS Status	-	-	-	53%	-	60%	-	HIV = AUD+HIV ($\chi^2=0.9968$)
Cognitive								
Verbal Language (VL)	-0.10±0.86	-0.70±0.96	2.0424 x 10 ⁻⁵	-0.92±1.14	3.7039 x 10 ⁻⁶	-0.71±0.99	2.5637 x 10 ⁻⁴	CTRL > AUD = HIV = AUD+HIV
Executive Function (EXF)	-0.04±0.90	-0.70±1.12	3.7239 x 10 ⁻⁶	-0.98±1.55	4.7593 x 10 ⁻⁶	-1.12±1.43	9.3434 x 10 ⁻⁸	CTRL > AUD = HIV = AUD+HIV
Learning and Memory (LM)	-0.08±0.80	-0.95±0.86	8.5698 x 10 ⁻¹¹	-0.84±0.97	2.8087 x 10 ⁻⁶	-1.08±0.88	7.0479 x 10 ⁻¹⁰	CTRL > AUD = HIV = AUD+HIV
Speed of Information Processing (SIP)	0.14±0.69	-0.39±0.90	5.8493 x 10 ⁻⁶	-0.51±0.85	8.5445 x 10 ⁻⁷	-0.82±1.01	4.6449 x 10 ⁻¹⁰	CTRL > AUD = HIV = AUD+HIV
Motor Skills (MS)	-0.10±0.77	-0.72±0.90	4.1689 x 10 ⁻⁷	-0.93±1.36	5.6897 x 10 ⁻⁶	-0.96±1.44	1.0281 x 10 ⁻⁵	CTRL > AUD = HIV = AUD+HIV
Quality of Social Functioning (QSF)	0.15±0.75	-2.00±1.63	2.2572 x 10 ⁻³⁸	-1.94±1.47	1.6003 x 10 ⁻³⁰	-2.80±1.96	4.8170 x 10 ⁻³⁶	CTRL > AUD = HIV = AUD+HIV

* '=' not significantly different; '<' or '>' significantly different (p < 0.05)

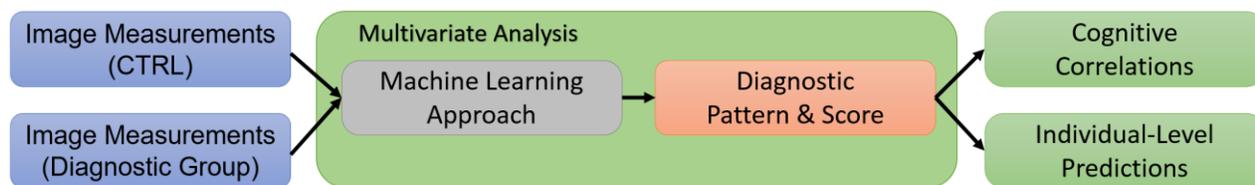


Figure 1: The analysis approach used for identifying diagnostic pattern and score specific to AUD, HIV, and AUD+HIV. It includes three major steps: multivariate analysis for identification of diagnostic pattern and score (Step 1), cognitive correlations (Step 2), and individual-level predictions (Step 3).

2.3 Machine Learning for Statistical Analysis

The data were divided into three diagnosis-specific sets: AUD(N=222) vs. CTRL(N=199), HIV(N=68) vs. CTRL(N=199), and AUD+HIV(N=60) vs. CTRL(N=199). With respect to each data set, our analysis followed the three steps outlined in Figure 1. Specifically, we first identified a *diagnostic pattern* by applying a multivariate machine learning method (see (27) and Supplement D for detail) to the entire data set (Step 1), which also automatically matched the sample size of the two cohorts. The resulting pattern was applied to individual MRIs producing a *diagnostic score*, which was the prediction of an individual having the diagnosis based solely on brain MRI measurements. The perfect *diagnostic score* for CTRLs was 0.0; for any individual from the diagnostic group, the perfect *diagnostic score* was 1.0. The *diagnostic pattern* was also correlated with the 6 cognitive scores recorded for members of that diagnostic group (Step 2). Correlations that were positive and had a p -value < 0.05 were reported. Step 3 measured the accuracy of the machine learning method via 10-fold cross-validation (59). The balanced accuracy (BAcc) (27), specificity, and sensitivity in predicting the diagnosis of each test subject were recorded. Furthermore, the significance of the accuracy (p -value < 0.001) was inferred using the Fisher exact test (60). A detailed description of the three steps appears in the Supplement C.

3. Results

3.1 Multivariate Analysis: Diagnostic Patterns

The *diagnostic patterns* are listed in Table 2 and visualized in Figure 2 (omitting white matter hypointensities). The AUD-specific pattern consisted of measures from 13 unique brain ROIs, the HIV-specific pattern of 15 brain ROIs, and the AUD+HIV-specific pattern of 25 brain ROIs. All patterns included the surface area of the posterior cingulate and the volumes of the WM hypointensities, precentral gyurs, thalamus, and hippocampus. 7 additional measures from the AUD+HIV pattern overlapped with either the AUD- or the HIV-pattern and included the mean curvature of the banks of the superior temporal sulcus and the average thickness of the superior frontal and superior parietal gyri.

Table 2: Measures associated with each diagnostic pattern. Bold-Italic entries denote measures in the AUD or HIV pattern that were also in the AUD+HIV pattern.

	Volume	Surface Area	Mean Curvature	Average Thickness
AUD	<i>superior frontal</i> <i>precentral gyrus</i> <i>thalamus</i> caudate <i>hippocampus</i> accumbens <i>WM hypointensities</i>	<i>posterior cingulate</i>	<i>bankssts*</i> lingual insula	lateral orbitofrontal pars triangularis postcentral insula
HIV	parahippocampal posterior cingulate <i>precentral gyrus</i> superior parietal <i>thalamus</i> <i>hippocampus</i> accumbens <i>WM hypointensities</i>	<i>supramarginal</i> <i>inferior temporal</i> <i>temporal pole</i> <i>pericalcarine</i> <i>posterior cingulate</i> insula	precuneus posterior cingulate insula	<i>superior frontal</i> <i>superior parietal</i> inferior temporal middle temporal insula
AUD+HIV	<i>superior frontal</i> frontal pole pars opercularis pars triangularis <i>precentral gyrus</i> precuneus <i>thalamus</i> caudate putamen <i>hippocampus</i> <i>WM hypointensities</i>	superior frontal caudal middle frontal paracentral medial orbitofrontal <i>supramarginal</i> <i>bankssts*</i> entorhinal <i>inferior temporal</i> <i>temporal pole</i> lateral occipital <i>pericalcarine</i> <i>posterior cingulate</i> rostral anterior cingulate	medial orbitofrontal paracentral pars orbitalis <i>bankssts*</i> entorhinal parahippocampal pericalcarine	<i>superior frontal</i> frontal pole medial orbitofrontal <i>superior parietal</i>

* bankssts = banks of the superior temporal sulcus

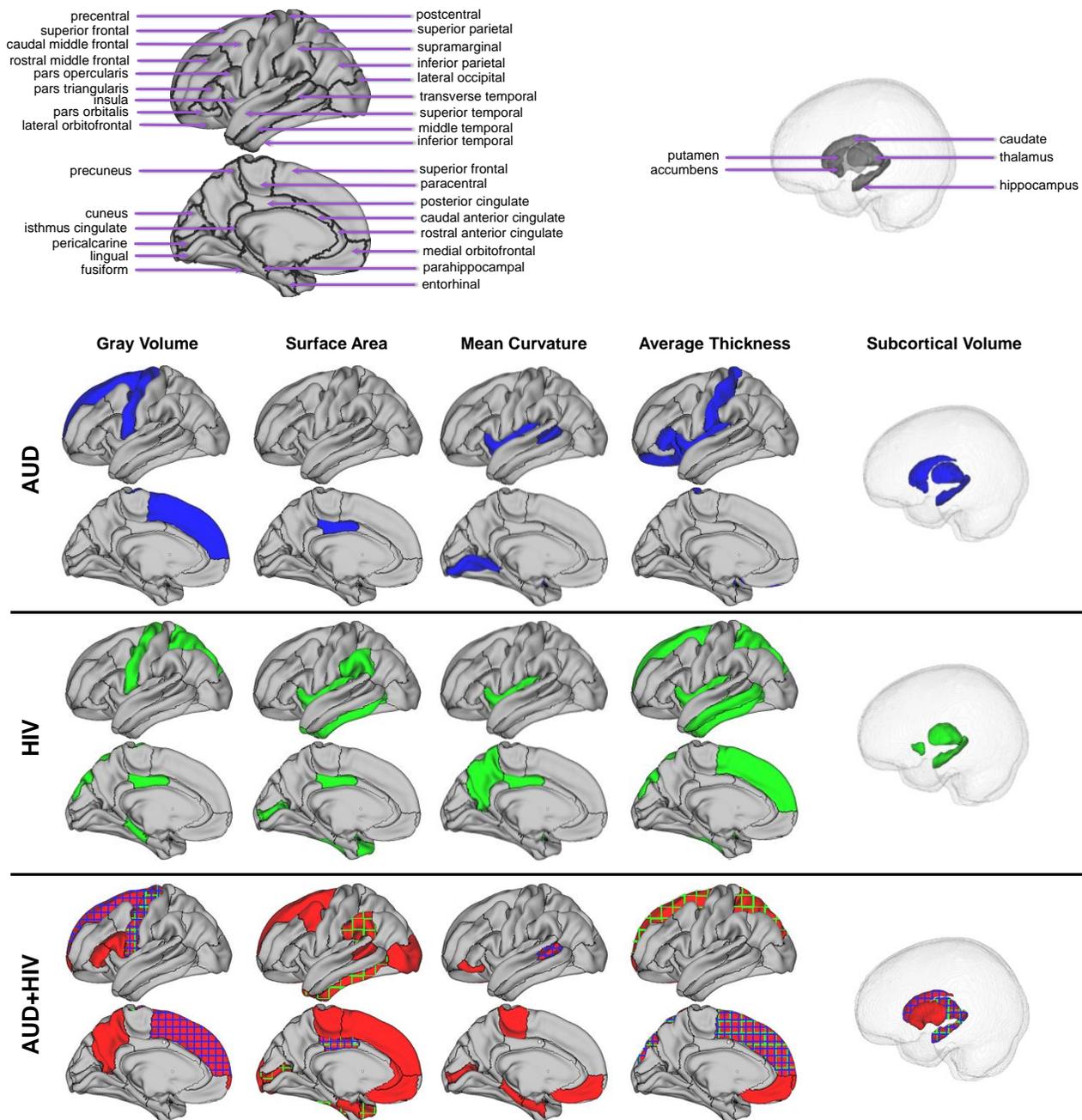


Figure 2: The diagnostic patterns (see also Table 2) for alcohol use dependency (AUD), HIV-infection (HIV), and the comorbidity (AUD+HIV). Measures in the HIV (purple) or AUD pattern (green) also appearing in the comorbidity pattern are shown in plaid in AUD+HIV.

3.2 Cognitive Correlations

Table 3 summarizes the correlations between the *diagnostic patterns* and the 6 cognitive measures. All three *diagnostic patterns* significantly predicted lower Learning/Memory (LM) scores. Furthermore, executive function (EXF) was significantly correlated with the AUD-pattern, whereas speed of information processing (SIP) was significantly correlated with the HIV and AUD+HIV patterns. As shown in Figure 3, these performance scores had lower values with higher *diagnostic scores*, i.e., greater certainty of an individual to be diagnosed with the condition. Testing the *diagnostic pattern* of AUD with lifetime alcohol consumption and that of HIV with VACS and CD4 measures did not reveal significant correlations.

3.3 Individual-Level Prediction

All three *diagnostic scores* predicted the diagnosis of individuals with significantly high accuracies. The balanced accuracy of the *diagnostic score* for AUD vs. CTRL was 70.1%, for HIV vs. CTRL was 76.2%, and for AUD+HIV vs. CTRL was 78.1%. The ranking of the three accuracy scores agreed with the size of the *diagnostic pattern*, i.e., the AUD pattern was the smallest (with 15 regional measurements), followed by HIV (22 measurements), and AUD+HIV (35 measurements) (Table 2). As each pattern inferred its own *diagnostic score*, each subject of the control group had a *diagnostic score* specific to each diagnosis. As expected, *diagnostic scores* of the CTRLs were generally lower than those of disease-affected participants (Figure 4).

Table 3: Correlation between the pattern identified for each diagnosis group and the six cognitive scores. Cognitive scores were measured through tests outlined in the Supplement A.

	Verbal Language	Executive Function	Learning and Memory	Speed of Information Processing	Motor Skills	Quality of Social Functioning
AUD		✓	✓			
HIV			✓	✓		
AUD+HIV			✓	✓		

✓ significant correlation (p-value < 0.05)

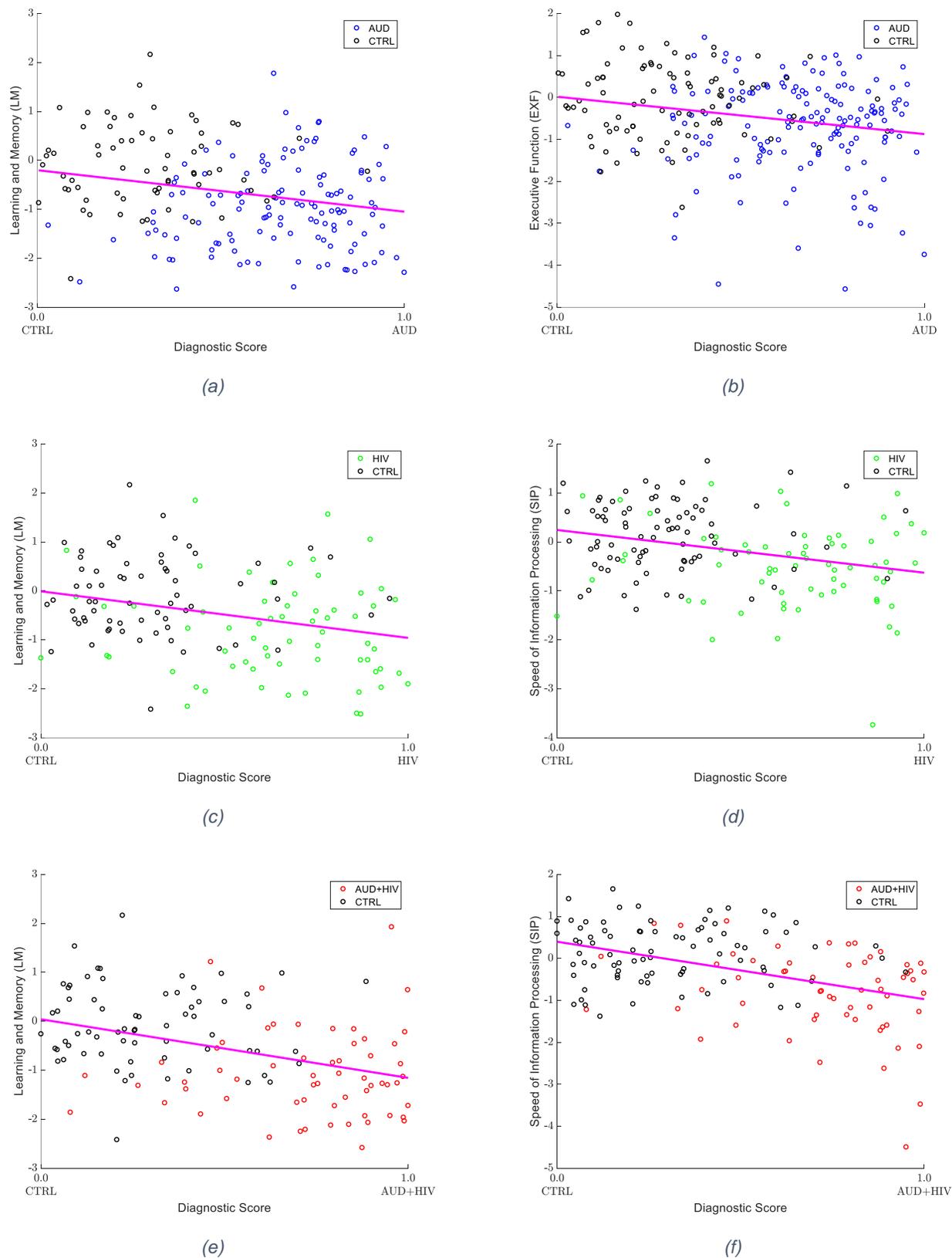


Figure 3: Cognitive measures significantly correlating with the diagnostic scores. The cognitive scores decline with increasing diagnostic scores. (a) Learning and Memory (LM) measure as a function of diagnostic score, for AUD vs. CTRL (b) Executive Function (EXF) measure as a function of diagnostic score, for AUD vs. CTRL (c) Learning and Memory (LM) measure as a function of diagnostic score, for HIV vs. CTRL (d) Speed of Information Processing (SIP) measure as a function of diagnostic score, for HIV vs. CTRL (e) Learning and Memory (LM) measure as a function of diagnostic score, for AUD+HIV vs. CTRL (f) Speed of Information Processing (SIP) measure as a function of diagnostic score, for AUD+HIV vs. CTRL

4. Discussion

The outcome of this machine learning analysis supported the hypothesis that MRIs alone can predict diagnosis and cognitive scores of AUD, HIV, or AUD+HIV. Our approach reduced the 298 brain measures to those most informing diagnosis, *i.e.*, the *diagnostic pattern* and then applied the pattern to the MRIs of individuals to compute each person's *diagnostic score*. Doing so preserved the statistical power of the data because testing for significance did not require multiple comparison correction as it is the case for conventional univariate analysis.

Conventional studies minimize the number of comparisons by pre-selecting measurements (e.g., gray matter volume (11) or average thickness (61)) or creating summary scores (17) deemed informative according to expert domain-knowledge. Scores then rely on univariate testing that entails a group-level analysis to correlate each score with each diagnosis. A powerful alternative to this conventional analysis is the proposed machine learning technology, which, as noted in Supplement E, was also more accurate in predicting diagnosis than other

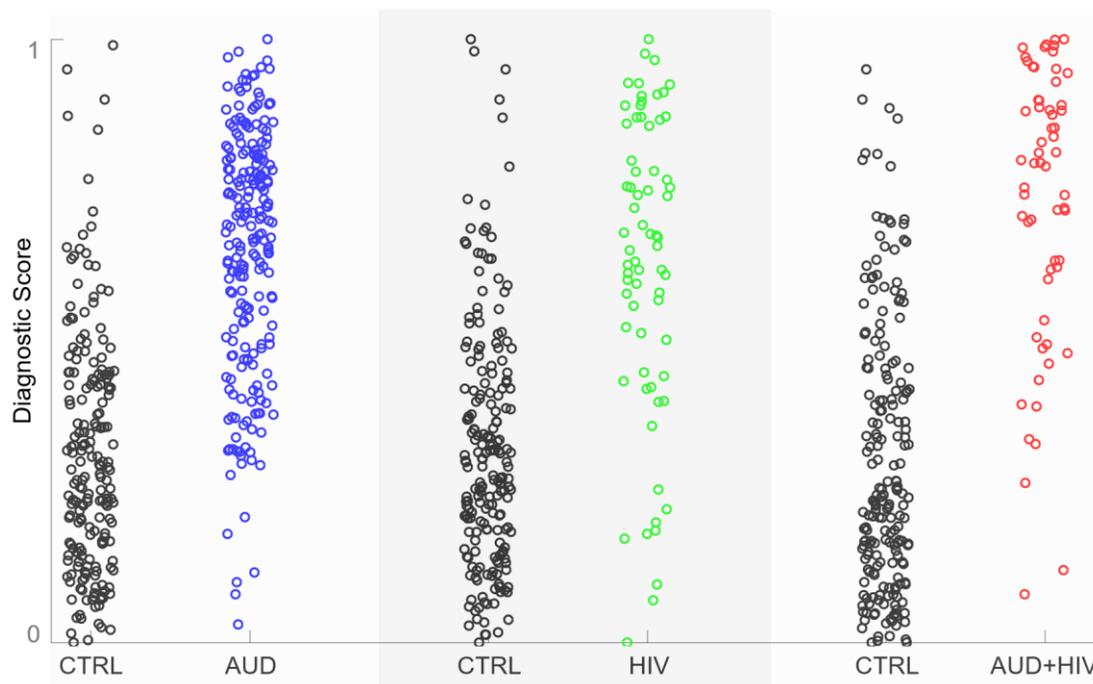


Figure 4: Diagnostic scores for each sample with respect to the three diagnosis-specific group comparisons (CTRL vs. AUD, CTRL vs. HIV, CTRL vs. AUD+HIV).

multivariate approaches. Not only did the novel machine learning approach predict AUD or HIV diagnosis of individuals based solely on their MRIs, but the predictive power of the identified patterns was measured on *unseen data* (*i.e.*, data not used for optimizing the approach) so that the patterns could serve as imaging phenotypes in other MRI studies of AUD or HIV.

We recognize that this paper presents a secondary analysis of data previously published in (17). Unlike the original study we omitted Hepatitis C coinfection from the analysis but our analysis was still confined by the number of HIV-infected patients (HIV and AUD+HIV) being much smaller than the control cohort and the AUD group. For our analysis to be impartial to this issue, the machine learning method automatically selected an equal number of samples from each group and trained its model on this balanced data set. Furthermore, we measured the prediction power using the balanced accuracy (BAcc) metric, which accounted for unequal sample sizes. The scores indicated that our approach produced accurate findings even in case of imbalanced data.

Human MRI studies of HIV or AUD have used machine learning analysis to predict the age of participants (62, 63) or to select MRI metrics related to a diagnosis (64-68). A hallmark of the proposed analysis was the *diagnostic score* of individuals (see Figure 4), which was a continuous score directly linking variation in MRI metrics to diagnosis. This link enabled a refined interpretation of significant correlations between *diagnostic patterns* and functional ramifications of the condition. For example, all three *diagnostic patterns* predicted the *Learning/Memory score* (see Table 3) and consisted of regions closely linked to this brain function, namely, the hippocampus (69-71), thalamus (72, 73), and posterior cingulate cortex (74, 75). The hippocampus and thalamus have been identified as targets of AUD (76), HIV (13, 77, 78), and AUD+HIV comorbidity has been shown to principally affect the thalamus (4, 79). Cingulate volume is more frequently reported as compromised in the HIV relative to the AUD literature, e.g., (79-83). While the imaging literature has typically reported on gray matter volume

effects, studies that assess cortical thickness rather than cortical volume can show different results. For example, HIV has been shown to compromise cortical thickness of areas such as the insula and temporal cortices (84-86).

Also reported for all three conditions was the link between lower *learning and memory scores* and higher *diagnostic scores* (Figure 3). According to our machine learning model, a higher *diagnostic score* reflected a greater impact of the disorder on the regional measures defining the *diagnostic pattern*. Thus, the *diagnostic scores* and *patterns* accurately summarized the magnitude of the impact that each disorder had on an affected metric.

An interesting observation was the inclusion of the hippocampus in the patterns of all groups. As it is reported in the literature (13, 87), this region supports learning and memory. Based on the statistics reported in Table 1, this cognitive measure was also significantly impaired across all three groups compared to the CTRL cohort.

Critically, our findings support a compounding effect of AUD and HIV on the neural systems of individuals diagnosed with both conditions. Among the three conditions, the *diagnostic pattern* of the comorbidity was the largest consisting of 35 regional measures. Several of these measures featured as part of the AUD and HIV pattern and were selected by the AUD+HIV pattern with the exception of the mean curvature and average thickness of the insula (see Table 2); however, neither mean curvature nor average thickness measure was selected by all three patterns. Rather, the patterns of the three diagnoses converged on volumes of four regions (precentral gyrus, hippocampus, thalamus, WM hypointensities) and one surface area (posterior cingulate cortex). Consistent with the size of the *diagnostic patterns* was the prediction accuracy of the *diagnostic scores*, which was most accurate for AUD+HIV comorbidity (78.1% BAcc). While the accuracy scores might be further improved based on the discussion in Supplement F, all these findings indicate that the combined impact of AUD and HIV on the brain system was more extensive than either condition alone.

In addition to overlap among the three *diagnostic patterns*, the comorbidity pattern contained elements specific each single diagnosis. The HIV and AUD+HIV patterns were highly accurate in predicting *speed of information processing performance*, which is known to decline faster in patients with HIV (88, 89) and alcohol (90) than the healthy individuals. Speed of information processing has been also linked to regions that were part of both patterns, notably, the superior frontal cortex (91, 92), precentral gyrus (93), superior parietal lobe (94, 95), inferior temporal lobe (91, 96, 97), pericalcarine gyrus (91), supramarginal gyrus (98, 99), and the temporal pole (100). Featured in both patterns was the thalamus, whose volume has been reported to be significantly smaller in the HIV population, with (4, 81) or without (4, 11, 13) AUD comorbidity. Appearing in both the AUD and the comorbidity pattern was the volume of the superior frontal cortex (17, 101), which has been observed to be smaller in AUD than controls (18, 102). The significant correlation between executive function and the diagnostic pattern was only reported with respect to the AUD group but not for the comorbidity cohort. This inconsistency might be explained by the executive functioning being negatively affected by alcohol consumption, which was more recent and prevalent in the AUD cohort. Their “days after last drink” was significantly shorter ($p=0.0012$; two-sided t-test) and the “alcohol consumed in past year” was significantly greater ($p < 0.001$; two-sided t-test) than for the AUD+HIV cohort. Interestingly, these findings further supported the compounding effect of AUD and HIV as their combined effects lead to a higher prediction accuracy than the AUD-specific pattern extracted on a cohort with higher alcohol consumption.

Note that we only reported on the compounding effect as the machine learning analysis could not quantitatively assess the additive or interactive characteristic of an effect. Furthermore, the subject-level inference from this type of analysis often does not accord to the results of group-level analysis (103). For instance, “quality of social functioning” was the cognitive score most strongly differentiating between controls and the three cohorts. However, the score did not

significantly correlate with any of the diagnostic patterns or in distinguishing individuals. In this data set, a better predictor for identifying significant correlations was the variance of a score within a diagnostic group. For each diagnostic group, the corresponding diagnostic patterns significantly correlated with the cognitive scores having the smallest variation, which was Speed of Information Processing for the HIV cohort and Learning/Memory for the other two diagnoses. This observation is in line with the analysis performed in Step 2 of *Figure 1* as the corresponding correlation was sensitive towards the within-class covariance.

Another limitation of this study was the assumption that samples were healthy or diagnosed with HIV, AUD, or their comorbidity. One could thus increase the significance of the data-driven predictions by measuring the prediction accuracies of the *diagnostic scores* on the MRIs of participants with other diagnoses. However, the prediction accuracy was determined on *unseen data* so that the findings of this study should apply to other MRI studies adhering to this assumption. Specifically, we measured the accuracy of our machine learning method using 10-fold cross-validation (Step 3 in *Figure 1*). To avoid reporting overly optimistic findings, cross-validation parameterized the z-scores (with respect to the controls) and the proceeding method (including sample selection) on a subset of the data (training) and then the accuracy of the method was measured on the remaining data, which avoided reporting overly optimistic findings. One drawback of this process is that our method identified a unique diagnostic pattern for each training run. Discussing the common denominator of the 10 different patterns is complex and requires statistics over the entire data set. To simplify, we focused the discussion for each diagnosis on an example of a diagnostic pattern, which was created by applying the machine learning approach on the entire diagnostic data set (Step 1 in *Figure 1*).

Finally, we caution against drawing conclusions about measurements omitted from diagnostic patterns presented here as these measurements can also be informative with respect to a diagnosis. However, they were not picked by the machine learning approach, which identified a

constellation of measurements that achieved a higher accuracy in labeling the individuals of the training data. Changing the training data (as done in cross-validation) can lead to selecting a different pattern. Thus, the *diagnostic patterns* presented in this article should be viewed as an example of a family of patterns that lead to similar prediction accuracy.

5. Conclusion

We report on the *diagnostic patterns* and *scores* based on MRI data that predicted diagnostic classification of individuals with AUD, HIV, or their comorbidity relative to control patterns. Novel machine learning technology automatically reduced 298 MRI brain measures to small subsets implicated by each diagnostic group, eliminating the need for expert-driven input. The impact of a disorder on the diagnostic pattern was summarized by a *diagnostic score*, which revealed an exacerbated effect of AUD+HIV comorbidity. The *diagnostic patterns* and *scores* also predicted cognitive performance of individuals and their accuracy was measured on unseen data. Thus, they could serve as imaging phenotypes for studies investigating AUD, HIV, and their comorbidity. The entire analysis was data-driven so that the novel machine learning approach is readily applicable to MRI studies of other neuropsychiatric conditions also enabling repurposing of multi-metric data.

Acknowledgments and Disclosures

Funding for this study was received from the U.S. National Institute on Alcohol Abuse and Alcoholism (AA017347, AA005965, AA010723, AA017168, AA026762), and the Moldow Women's Hope and Healing Fund.

None of the authors have conflicts of interest with the reported data or their interpretation.

References

1. Organization WH, Unit WHOMoSA (2014): *Global status report on alcohol and health, 2014*. World Health Organization.
2. Gongvatana A, Morgan EE, Iudicello JE, Letendre SL, Grant I, Woods SP, et al. (2014): A history of alcohol dependence augments HIV-associated neurocognitive deficits in persons aged 60 and older. *J Neurovirol.* 20:505-513.
3. Fama R, Sullivan EV, Sassoos SA, Pfefferbaum A, Zahr NM (2016): Impairments in Component Processes of Executive Function and Episodic Memory in Alcoholism, HIV Infection, and HIV Infection with Alcoholism Comorbidity. *Alcohol Clin Exp Res.* 40:2656-2666.
4. Pfefferbaum A, Rosenbloom MJ, Sassoos SA, Kemper CA, Deresinski S, Rohlfing T, et al. (2012): Regional brain structural dysmorphology in human immunodeficiency virus infection: effects of acquired immune deficiency syndrome, alcoholism, and age. *Biol Psychiatry.* 72:361-370.
5. Justice A, Sullivan L, Fiellin D, Veterans Aging Cohort Study Project T (2010): HIV/AIDS, comorbidity, and alcohol: can we make a difference? *Alcohol Res Health.* 33:258-266.
6. Cardenas VA, Studholme C, Meyerhoff DJ, Song E, Weiner MW (2005): Chronic active heavy drinking and family history of problem drinking modulate regional brain tissue volumes. *Psychiatry Res.* 138:115-130.
7. Pfefferbaum A, Sullivan EV, Mathalon DH, Lim KO (1997): Frontal lobe volume loss observed with magnetic resonance imaging in older chronic alcoholics. *Alc Clin Exp Research.* 21:521-529.
8. Cardenas VA, Studholme C, Gazdzinski S, Durazzo TC, Meyerhoff DJ (2007): Deformation-based morphometry of brain changes in alcohol dependence and abstinence. *Neuroimage.* 34:879-887.
9. Sullivan EV, Rose J, Pfefferbaum A (2006): Effect of vision, touch and stance on cerebellar vermian-related sway and tremor: a quantitative physiological and MRI study. *Cerebral cortex.* 16:1077-1086.
10. Sullivan EV, Pfefferbaum A (2009): Neuroimaging of the Wernicke-Korsakoff syndrome. *Alcohol Alcohol.* 44:155-165.
11. Pfefferbaum A, Rogosa DA, Rosenbloom MJ, Chu W, Sassoos SA, Kemper CA, et al. (2014): Accelerated aging of selective brain structures in human immunodeficiency virus infection: a controlled, longitudinal magnetic resonance imaging study. *Neurobiol Aging.* 35:1755-1768.
12. Le Berre AP, Pitel AL, Chanraud S, Beaunieux H, Eustache F, Martinot JL, et al. (2014): Chronic alcohol consumption and its effect on nodes of frontocerebellar and limbic circuitry: comparison of effects in France and the United States. *Hum Brain Mapp.* 35:4635-4653.
13. Wade BS, Valcour V, Busovaca E, Esmaeili-Firidouni P, Joshi SH, Wang Y, et al. (2015): Subcortical shape and volume abnormalities in an elderly HIV+ cohort. *Proc SPIE Int Soc Opt Eng.* 9417.
14. Fama R, Rosenbloom MJ, Nichols BN, Pfefferbaum A, Sullivan EV (2009): Working and episodic memory in HIV infection, alcoholism, and their comorbidity: baseline and 1-year follow-up examinations. *Alcohol Clin Exp Res.* 33:1815-1824.
15. Bernardin F, Maheut-Bosser A, Paille F (2014): Cognitive impairments in alcohol-dependent subjects. *Front Psychiatry.* 5:78.
16. Heaton RK, Franklin DR, Ellis RJ, McCutchan JA, Letendre SL, Leblanc S, et al. (2011): HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors. *J Neurovirol.* 17:3-16.
17. Pfefferbaum A, Zahr NM, Sassoos SA, Kwon D, Pohl KM, Sullivan EV (2018): Accelerated and Premature Aging Characterizing Regional Cortical Volume Loss in Human Immunodeficiency

- Virus Infection: Contributions From Alcohol, Substance Use, and Hepatitis C Coinfection. *Biol Psychiatry Cogn Neurosci Neuroimaging*.
18. Manzo G, De Gennaro A, Cozzolino A, Serino A, Fenza G, Manto A (2014): MR imaging findings in alcoholic and nonalcoholic acute Wernicke's encephalopathy: a review. *Biomed Res Int*. 2014:503596.
 19. Avants BB, Epstein CL, Grossman M, Gee JC (2008): Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*. 12:26-41.
 20. Crovitz HF, Zener K (1962): A group-test for assessing hand- and eye-dominance. *Am J Psychol*. 75:271-276.
 21. Rosenbloom MJ, Sullivan EV, Sasso SA, O'Reilly A, Fama R, Kemper CA, et al. (2007): Alcoholism, HIV infection, and their comorbidity: factors affecting self-rated health-related quality of life. *J Stud Alcohol Drugs*. 68:115-125.
 22. Bzdok D (2017): Classical Statistics and Statistical Learning in Imaging Neuroscience. *Front Neurosci*. 11:543.
 23. Smith SM, Nichols TE (2018): Statistical Challenges in "Big Data" Human Neuroimaging. *Neuron*. 97:263-268.
 24. Varoquaux G, Thirion B (2014): How machine learning is shaping cognitive neuroimaging. *Gigascience*. 3:28.
 25. Castiglioni I, Salvatore C, Ramirez J, Gorriz JM (2018): Machine-learning neuroimaging challenge for automated diagnosis of mild cognitive impairment: Lessons learnt. *J Neurosci Methods*. 302:10-13.
 26. Zhang Y, Kwon D, Esmaeili-Firidouni P, Pfefferbaum A, Sullivan EV, Javitz H, et al. (2016): Extracting patterns of morphometry distinguishing HIV associated neurodegeneration from mild cognitive impairment via group cardinality constrained classification. *Hum Brain Mapp*. 37:4523-4538.
 27. Adeli E, Li X, Kwon D, Zhang Y, Pohl KM (2018): Logisitc Regression Confined by Cardinality-Constrained Sample and Feature Selection. *Accepted in IEEE Transactions to Pattern Analysis and Machine Intelligence*.
 28. Adeli E, Shi F, An L, Wee CY, Wu G, Wang T, et al. (2016): Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *Neuroimage*. 141:206-219.
 29. Hastie T, Tibshirani R, Wainwright M (2015): *Statistical learning with sparsity: the lasso and generalizations*. CRC Press.
 30. Guggenmos M, Scheel M, Sekutowicz M, Garbusow M, Sebold M, Sommer C, et al. (2018): Decoding diagnosis and lifetime consumption in alcohol dependence from grey-matter pattern information. *Acta Psychiatr Scand*. 137:252-262.
 31. Ohya T, Kikuchi S (1982): [Clinical evaluation of chemotherapeutic agents in the treatment of primary liver cancer]. *Gan To Kagaku Ryoho*. 9:1623-1627.
 32. Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, et al. (2006): Veterans Aging Cohort Study (VACS): Overview and description. *Med Care*. 44:S13-24.
 33. Antinori A, Arendt G, Becker JT, Brew BJ, Byrd DA, Cherner M, et al. (2007): Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*. 69:1789-1799.
 34. Sacktor N, Skolasky RL, Seaberg E, Munro C, Becker JT, Martin E, et al. (2016): Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study. *Neurology*. 86:334-340.
 35. Tombaugh TN (2004): Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*. 19:203-214.

36. D'Elia L, Satz P (1989): *Color Trails 1 and 2*. Odessa, FL: Psychological Assessment Resources.
37. Bowden SC, Bell RC (1992): Relative usefulness of the WMS and WMS-R: a comment on D'Elia et al. (1989). *Journal of clinical and experimental neuropsychology*. 14:340-346.
38. Elwood RW (2001): MicroCog: assessment of cognitive functioning. *Neuropsychology review*. 11:89-100.
39. Golden C (1978): *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Chicago: Stoelling Co.
40. Bennett-Levy J (1984): Determinants of performance on the Rey-Osterrieth complex figure test: An analysis, and a new technique for single-case assessment. *British Journal of Clinical Psychology*. 23:109-119.
41. Borkowski JG, Benton AL, Spreen O (1967): Word fluency and brain damage. *Neuropsychologia*. 5:135-140.
42. Nelson HE (1982): *The National Adult Reading Test (NART)*. Windsor, Canada: Nelson Publishing Company.
43. Dunn LM, Dunn ES (1997): *Peabody Picture Vocabulary Test - Third Edition*. Circle Pines, MN: American Guidance Service.
44. Wechsler D (2001): *Wechsler Test of Adult Reading: WTAR*. San Antonio, TX: Pearson Education, Inc.
45. Hart R, Kwentus J, Wade J, Hamer R (1987): Digit symbol performance in mild dementia and depression. *Journal of Consulting and Clinical Psychology*. 55:236-238.
46. Smith A (1973): *The Symbol Digit Modalities Test Manual*. Los Angeles: Western Psychological Services.
47. Trites RL (1977): The Grooved Pegboard Test. *Neuropsychological Test Manual*. Ontario, Canada: Royal Ottawa Hospital.
48. Fama R, Eisen JC, Rosenbloom MJ, Sassoon SA, Kemper CA, Deresinski S, et al. (2007): Upper and lower limb motor impairments in alcoholism, HIV infection, and their comorbidity. *Alcohol Clin Exp Res*. 31:1038-1044.
49. Fregly AR (1968): An ataxia battery not requiring rails. *Aerospace Medicine*. 39:277-282.
50. Bozette SA, Hays RD, Berry SH, Kanouse DE, Wu AW (1995): Derivation and properties of a brief health status assessment instrument for use in HIV disease. *Journal of acquired immune deficiency syndromes and human retrovirology : official publication of the International Retrovirology Association*. 8:253-265.
51. Moos RH, McCoy L, Moos BS (2000): Global assessment of functioning (GAF) ratings: determinants and role as predictors of one-year treatment outcomes. *Journal of Clinical Psychology*. 56:449-461.
52. Katz S (1983): Assessing self-maintenance activities of daily living, mobility and instrumental activities of daily living. *Journal of the American Geriatric Society*. 31:721-727.
53. Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A (2010): The SRI24 multichannel atlas of normal adult human brain structure. *Hum Brain Mapp*. 31:798-819.
54. Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*. 9:179-194.
55. Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012): Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*. 61:1402-1418.
56. Fischl B (2012): FreeSurfer. *Neuroimage*. 62:774-781.
57. Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A (2010): The SRI24 multi-channel atlas of normal adult human brain structure. *Human Brain Mapping*. 31:798-819.
58. Madsen H, Thyregod P (2010): *Introduction to general and generalized linear models*. CRC Press.

59. Arlot S, Celisse A (2010): A survey of cross-validation procedures for model selection. *Statistics surveys*. 4:39.
60. Fisher RA (1935): The logic of inductive inference. *Journal of the Royal Statistical Society*. 98:43.
61. Thompson PM, Dutton RA, Hayashi KM, Toga AW, Lopez OL, Aizenstein HJ, et al. (2005): Thinning of the cerebral cortex visualized in HIV/AIDS reflects CD4+ T lymphocyte decline. *Proc Natl Acad Sci U S A*. 102:15647-15652.
62. Kuhn T, Kaufmann T, Doan NT, Westlye LT, Jones J, Nunez RA, et al. (2018): An augmented aging process in brain white matter in HIV. *Hum Brain Mapp*. 39:2532-2540.
63. Cole JH, Underwood J, Caan MW, De Francesco D, van Zoest RA, Leech R, et al. (2017): Increased brain-predicted aging in treated HIV disease. *Neurology*. 88:1349-1357.
64. Squeglia LM, Ball TM, Jacobus J, Brumback T, McKenna BS, Nguyen-Louie TT, et al. (2017): Neural Predictors of Initiating Alcohol Use During Adolescence. *Am J Psychiatry*. 174:172-185.
65. Cao B, Kong X, Kettering C, Yu P, Ragin A (2015): Determinants of HIV-induced brain changes in three different periods of the early clinical course: A data mining analysis. *Neuroimage Clin*. 9:75-82.
66. Wade BS, Valcour VG, Wendelken-Riegelhaupt L, Esmaili-Firidouni P, Joshi SH, Gutman BA, et al. (2015): Mapping abnormal subcortical brain morphometry in an elderly HIV+ cohort. *Neuroimage Clin*. 9:564-573.
67. Underwood J, Cole JH, Caan M, De Francesco D, Leech R, van Zoest RA, et al. (2017): Gray and White Matter Abnormalities in Treated Human Immunodeficiency Virus Disease and Their Relationship to Cognitive Function. *Clin Infect Dis*. 65:422-432.
68. Zhu X, Du X, Kerich M, Lohoff FW, Momenan R (2018): Random forest based classification of alcohol dependence patients and healthy controls using resting state MRI. *Neurosci Lett*. 676:27-33.
69. Deng W, Aimone JB, Gage FH (2010): New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? *Nat Rev Neurosci*. 11:339-350.
70. Petersen RC, Jack CR, Jr., Xu YC, Waring SC, O'Brien PC, Smith GE, et al. (2000): Memory and MRI-based hippocampal volumes in aging and AD. *Neurology*. 54:581-587.
71. Bird CM, Burgess N (2008): The hippocampus and memory: insights from spatial processing. *Nat Rev Neurosci*. 9:182-194.
72. Aggleton JP, O'Mara SM, Vann SD, Wright NF, Tsanov M, Erichsen JT (2010): Hippocampal-anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *Eur J Neurosci*. 31:2292-2307.
73. de Bourbon-Teles J, Bentley P, Koshino S, Shah K, Dutta A, Malhotra P, et al. (2014): Thalamic control of human attention driven by memory and learning. *Curr Biol*. 24:993-999.
74. Pearson JM, Heilbronner SR, Barack DL, Hayden BY, Platt ML (2011): Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci*. 15:143-151.
75. Leech R, Sharp DJ (2014): The role of the posterior cingulate cortex in cognition and disease. *Brain*. 137:12-32.
76. Sullivan EV, Marsh L (2003): Hippocampal volume deficits in alcoholic Korsakoff's syndrome. *Neurology*. 61:1716-1719.
77. Scott-Sheldon LA, Walstrom P, Carey KB, Johnson BT, Carey MP, Team MR (2013): Alcohol use and sexual risk behaviors among individuals infected with HIV: a systematic review and meta-analysis 2012 to early 2013. *Curr HIV/AIDS Rep*. 10:314-323.
78. Sanford R, Fernandez Cruz AL, Scott SC, Mayo NE, Fellows LK, Ances BM, et al. (2017): Regionally Specific Brain Volumetric and Cortical Thickness Changes in HIV-Infected Patients in the HAART Era. *Journal of acquired immune deficiency syndromes*. 74:563-570.

79. Janssen MA, Meulenbroek O, Steens SC, Goraj B, Bosch M, Koopmans PP, et al. (2015): Cognitive functioning, wellbeing and brain correlates in HIV-1 infected patients on long-term combination antiretroviral therapy. *AIDS*. 29:2139-2148.
80. Clark US, Walker KA, Cohen RA, Devlin KN, Folkers AM, Pina MJ, et al. (2015): Facial emotion recognition impairments are associated with brain volume abnormalities in individuals with HIV. *Neuropsychologia*. 70:263-271.
81. Fama R, Rosenbloom MJ, Sassoos SA, Rohlfing T, Pfefferbaum A, Sullivan EV (2014): Thalamic volume deficit contributes to procedural and explicit memory impairment in HIV infection with primary alcoholism comorbidity. *Brain Imaging Behav*. 8:611-620.
82. Agudelo M, Khatavkar P, Yndart A, Yoo C, Rosenberg R, Devieux JG, et al. (2014): Alcohol abuse and HIV infection: role of DRD2. *Curr HIV Res*. 12:234-242.
83. Ances BM, Ortega M, Vaida F, Heaps J, Paul R (2012): Independent effects of HIV, aging, and HAART on brain volumetric measures. *Journal of acquired immune deficiency syndromes*. 59:469-477.
84. Kallianpur KJ, Kirk GR, Sailasuta N, Valcour V, Shiramizu B, Nakamoto BK, et al. (2012): Regional cortical thinning associated with detectable levels of HIV DNA. *Cerebral cortex*. 22:2065-2075.
85. Sanford R, Fellows LK, Ances BM, Collins DL (2018): Association of Brain Structure Changes and Cognitive Function With Combination Antiretroviral Therapy in HIV-Positive Individuals. *JAMA Neurol*. 75:72-79.
86. Adeli E, Kwon D, Zhao Q, Pfefferbaum A, Zahr NM, Sullivan EV, et al. (2018): Chained regularization for identifying brain patterns specific to HIV infection. *Neuroimage*. 183:425-437.
87. Anderson EB, Grossrubatscher I, Frank L (2014): Dynamic Hippocampal Circuits Support Learning- and Memory-Guided Behaviors. *Cold Spring Harb Symp Quant Biol*. 79:51-58.
88. Fellows RP, Byrd DA, Morgello S (2014): Effects of information processing speed on learning, memory, and executive functioning in people living with HIV/AIDS. *Journal of clinical and experimental neuropsychology*. 36:806-817.
89. Anand P, Springer SA, Copenhaver MM, Altice FL (2010): Neurocognitive impairment and HIV risk factors: a reciprocal relationship. *AIDS Behav*. 14:1213-1226.
90. Tzambazis K, Stough C (2000): Alcohol impairs speed of information processing and simple and choice reaction time and differentially impairs higher-order cognitive abilities. *Alcohol Alcohol*. 35:197-201.
91. Dux PE, Tombu MN, Harrison S, Rogers BP, Tong F, Marois R (2009): Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*. 63:127-138.
92. Charlton RA, Landau S, Schiavone F, Barrick TR, Clark CA, Markus HS, et al. (2008): A structural equation modeling investigation of age-related variance in executive function and DTI measured white matter damage. *Neurobiol Aging*. 29:1547-1555.
93. Carter CS, Mintun M, Cohen JD (1995): Interference and facilitation effects during selective attention: an H215O PET study of Stroop task performance. *Neuroimage*. 2:264-272.
94. Gold BT, Powell DK, Xuan L, Jiang Y, Hardy PA (2007): Speed of lexical decision correlates with diffusion anisotropy in left parietal and frontal white matter: evidence from diffusion tensor imaging. *Neuropsychologia*. 45:2439-2446.
95. Nishiike S, Nakagawa S, Tonoike M, Takeda N, Kubo T (2001): Information processing of visually-induced apparent self motion in the cortex of humans: analysis with magnetoencephalography. *Acta Otolaryngol Suppl*. 545:113-115.
96. Ptak R, Valenza N (2005): The inferior temporal lobe mediates distracter-resistant visual search of patients with spatial neglect. *J Cogn Neurosci*. 17:788-799.

97. Mruczek RE, Sheinberg DL (2007): Activity of inferior temporal cortical neurons predicts recognition choice behavior and recognition time during visual search. *J Neurosci.* 27:2825-2836.
98. Perianez JA, Maestu F, Barcelo F, Fernandez A, Amo C, Ortiz Alonso T (2004): Spatiotemporal brain dynamics during preparatory set shifting: MEG evidence. *Neuroimage.* 21:687-695.
99. Lazeron RH, Rombouts SA, Machielsen WC, Scheltens P, Witter MP, Uylings HB, et al. (2000): Visualizing brain activation during planning: the tower of London test adapted for functional MR imaging. *AJNR Am J Neuroradiol.* 21:1407-1414.
100. Decety J, Cacioppo S (2012): The speed of morality: a high-density electrical neuroimaging study. *J Neurophysiol.* 108:3068-3072.
101. Sullivan EV, Zahr NM, Sassoon SA, Thompson WK, Kwon D, Pohl KM, et al. (2018): The Role of Aging, Drug Dependence, and Hepatitis C Comorbidity in Alcoholism Cortical Compromise. *JAMA Psychiatry.* 75:474-483.
102. Zahr NM (2014): Structural and microstructural imaging of the brain in alcohol use disorders. *Handb Clin Neurol.* 125:275-290.
103. Zhang Y, Kwon D, Pohl KM (2017): Computing group cardinality constraint solutions for logistic regression problems. *Med Image Anal.* 35:58-69.

Novel Machine Learning Identifies Brain Patterns Distinguishing Diagnostic Membership of Human Immunodeficiency Virus, Alcoholism, and Their Comorbidity of Individuals

Supplementary Information

Participants

Table 1 (in the main manuscript) also lists the medication history of the participants. With reference to (1), we recorded self-reported medication history in the current study. At the time of data acquisition and clinical assessment, the participants were asked if they were taking any kind of medication for alcohol addiction within the past 30 days of their visit. Based on those records, none of the participants with AUD diagnosis was taking naltrexone or any other anti-craving medication. The abstinence status of these participants, denoted by days after the last drink in Table 1, was also recorded for AUD groups. Furthermore, approximately 88% of the participants in HIV groups (HIV and AUD+HIV) were on HAART (highly active antiretroviral therapy) medications at the time of their scans. The VACS index defines a score by summing pre-assigned points for age, indicators of HIV disease (CD4 count and HIV-1 RNA), and general indicators of organ system injury including hemoglobin, platelets, aspartate and alanine transaminase (AST and ALT), creatinine, and viral hepatitis C infection (HCV). Alongside other variables, Table 1 shows the group differences between the diagnosis groups and CTRL. As expected, the AUD group does have significantly different VACS scores compared to the CTRL, while HIV and AUD+HIV groups are significantly different in terms of this index. Note, (2) provides additional details about the demographic factors and cognitive tests.

Past or current nicotine dependence had a higher incidence in the diagnostic groups than the CTRL. A large proportion in each diagnostic group also met historical DSM-IV criteria for substance dependence (to cannabis, cocaine, amphetamines, or opiates): 29 (42.6%) HIV, 46

(76.7%) AUD+HIV, and 128 (57.7%) AUD. One control individual developed cannabis dependence at a later MRI but had no drug diagnosis at her initial visit. Only participants with serologically-confirmed HCV status were considered in analysis including 23/66 (34.8%) HIV, 31/59 (52.5%) HIV+AUD, 37/152 (24.3%) AUD, and 4/93 control participants (only the 89 HCV-negative controls were included in analyses related to HCV).

MRI Processing

Preprocessing of the T1-weighted (T1w) MR images involved noise removal (3) and correcting field inhomogeneity via N4ITK (Version 2.1.0) (4). Next, the brain mask was segmented by majority voting (5) across maps extracted by FSL BET (Version 5.0.6) (6), AFNI 3dSkullStrip (Version AFNI_2011_12_21_1014) (7), FreeSurfer mri-gcut (Version 5.3.0) (8), and the Robust Brain Extraction (ROBEX) method (Version 1.2) (9); all methods were applied to bias and non-bias corrected T1w images. Based on the brain mask, image inhomogeneity correction was repeated and the skull was omitted from the images. FreeSurfer software (Version 5.3.0) (10-12) was applied to the resulting skull-stripped T1w MRI to measure the mean curvature (MeanCurv), surface area (SurfArea), gray matter volume (GrayVol), and average thickness (ThickAvg) of 34 bilateral cortical Regions Of Interest (ROIs) from the MNI template (13), the volumes of 8 bilateral subcortical ROIs (i.e., thalamus, caudate, putamen, pallidum, hippocampus, amygdala, accumbens, cerebellar cortex), the volumes of 5 subregions of the corpus callosum (i.e., posterior, mid-posterior, central, mid-central and anterior), and the volume of the white matter hyperintensities (12). In addition to FreeSurfer, the SRI24 atlas (14) was non-rigidly registered to the T1w MRI of the subject via ANTS (Version: 2.1.0) (15) to measure the supratentorial volume (svol) and the volumes of the ventricles. Each subject was then represented by svol and the z-scores of the remaining 298 morphometric measurements.

The morphometric measurements of the CTRL group varied significantly with age, sex, and svol (Pearson correlation p -value < 0.005). These confounding factors (v_{age} , v_{sex} , and v_{svol}) were

regressed out from the j^{th} morphometric measurements (v_j) by parameterizing β_0, \dots, β_3 of the following general linear model (GLM) (16) on the controls of the training data:

$$v_j \sim \beta_{j,0} + \beta_{j,1}v_{age} + \beta_{j,2}v_{sex} + \beta_{j,3}v_{svol}.$$

After parameterizing the GLM, the model was applied to the morphometric measurements of each sample of the test data set.

Detailed Steps of Statistical Analysis

Step 1: Multivariate Analysis - Identification of Diagnostic Pattern and Score

For each diagnosis-specific data set, our multivariate machine learning method selected a pattern of brain measures that was most informative for distinguishing individuals of the diagnostic group from CTRLs. The *diagnostic pattern* was identified by the machine learning method, which did so by first automatically matching the sample size of the two cohorts. Sample matching consisted of selecting individuals, whose measurements the algorithm judged as best in characterizing the cohort. By doing so, this entirely data-driven processes minimized the risk of biasing the analysis in favor of the group with more samples, a common problem with this type of technology (17). The method either selected a regional measure from both hemispheres or ignored it entirely because it was assumed that the effects of AUD and HIV infection on the brain are bilateral. Finally, the identified pattern was applied to individual MRIs resulting in a *diagnostic score*, which was the prediction of an individual's having the condition based solely on brain MRI measurements.

Step 2: Cognitive Correlations

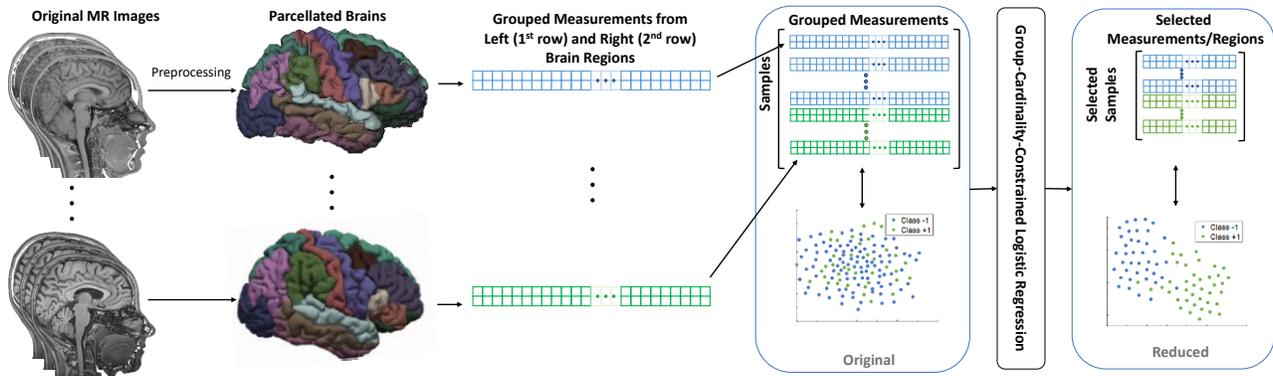
For each diagnostic-specific data set, the corresponding diagnostic pattern was correlated to 6 cognitive scores with respect to the samples of the diagnostic group. Separately for each cognitive score, the distance correlation metric (18) computed one correlation of the score to all measures

associated with a diagnostic pattern. A correlation was significant if it was positive and had a p -value < 0.05 .

Step 3: Individual-Level Predictions

Accuracy of each diagnostic score in predicting diagnosis was measured using 10-fold nested cross-validation (19). For each diagnosis, the corresponding dataset (CTRL vs. diagnostic group) was first divided into 10 folds. For each fold, a 5-fold inner-cross validation was performed on the remaining data to determine the optimal parameter setting for the multivariate analysis. For each subject in the single (test) fold, the parameterized machine learning approach then computed the *diagnostic score*. In other words, the diagnostic score was generated on *unseen data*, i.e., the data were not used for optimizing the approach so that findings on those data should translate to data sets from other studies with similar imaging characteristics. The perfect *diagnostic score* for CTRLs was 0.0; for any individual from the diagnostic group, the perfect *diagnostic score* was 1.0.

After the process of computing *diagnostic scores* was completed with respect to each of the 10 folds, the analysis recorded the balanced accuracy (BAcc) (20), specificity, and sensitivity of the *diagnostic score* predicting the diagnosis of each subject. The significance of the diagnostic score and BAcc score was determined by the Fisher exact test (21) (p -value < 0.001). We provide a detailed description about the machine learning algorithm for pattern identification and individual-level prediction below. For more technical insights on the approach, refer to (20).



Supplement Figure S1: Overview of our proposed classification method: Each subject in the study defines a sample, which is represented by its measures based on parcellations of the brain. Measures from the left and right brain ROIs are grouped together into a sample matrix. During training, the method simultaneously selects samples (rows) and groups of measures (columns) to determine the classifier most accurately predicting diagnose of individual subjects just based on their brain parcellations.

Technical Details of the Proposed Classification Framework

For each diagnosis-specific group comparisons, the novel machine learning approach identified a *diagnostic pattern* and *score* by selecting the most informative (residual) measures and samples for prediction (see Supplement Figure S1). As in (20), the selection process was based on minimizing a sparse logistic loss function (22), which dealt with noise, redundant measures, and skewed class distributions (imbalanced number of samples across classes); three common issues in analyzing MRI data. Further improving on (20), the approach assumed the effects of AUD and HIV infection on the brain to be bilateral by relying on group-sparsity (23), i.e., regional measure from both hemispheres were selected or this *group* of measures was entirely ignored. The approach enforced group-sparsity by first arranging the measures of a sample as a matrix (a.k.a. sample matrix). Each column in that sample matrix represented a group of measures, such as the volumes of the right and left precentral gyrus. Selecting group of measures than simplified to selecting columns in the sample matrix. Similarly, the matrices of all (training) samples were combined into a large matrix so that identifying the most informative samples simplified to selecting the corresponding rows in that matrix. The entire selection process was encoded by a *weight* matrix. The *weight* matrix was sparse, i.e., most entries in that matrix were zero. Finally,

the *diagnostic score* of a sample was then based on multiplying the weight matrix with the corresponding sample matrix.

Formulation

The prediction is defined as a classification problem with respect to n training samples $\mathbf{x}_i \in \mathbb{R}^p$ and their respective labels $y_i \in \{-1, +1\}, i = 1, \dots, n$. Samples selected for training of the classifier are encoded via the vector of indicator variables $\alpha \in \mathbb{Z}^n$ ($\mathbb{Z}^n := \{0,1\}^n$ indicates the set of all n -dimensional binary vectors). The indicator variable $\alpha_i = 1$ if the i^{th} sample (\mathbf{x}_i, y_i) is selected, and $\alpha_i = 0$ otherwise. Feature selection is encoded via the feature weight vector $\beta \in \mathbb{R}^n$. Now, let $\sigma(t) := \log(1 + \exp(-t))$ be the sigmoid function and $\beta_0 \in \mathbb{R}$ the bias term defining the linear log-odds function $\log\left(\frac{\Pr(y_i=1)}{\Pr(y_i=-1)}\right) = \beta^\top \mathbf{x}_i + \beta_0$. Then, the logistic regression loss function with an ℓ_2 - norm regularization on the classification weight vectors defining the binary classification model

$$\mathcal{L}(\alpha, \beta, \beta_0) := \sum_{i=1}^n \alpha_i \sigma(y_i(\beta^\top \mathbf{x}_i + \beta_0)) + \lambda \|\beta\|_2^2,$$

where λ is a regularization hyperparameter, encoding the weight of the regularization over the logistic loss function. Its value is tuned via cross-validation during the final prediction task similar to (20).

Training of the classifier determines the parameters that minimize the above cost function according to the constraints specific to sample and feature selection. As discussed before, we use group-sparsity first transforming the weight vector into a two-row matrix (shown also in Supplement Figure S1) so that each column represents the measurements from the left and right hemispheres. Hence, each group combines the weights associated with measurements from the same type and region (regardless of hemisphere). Let $\hat{\beta} \in \mathbb{R}^{2 \times g}$ denote this grouped matrix that contains g different groups. Similar to (20, 24), we constrain the model to select only $r < g$ groups of measurements (as for the feature selection module). With respect to sample selection, only k

samples are selected from each class. We deliberately enforce both classes to select same number of samples, to balance them in the process of building the classifier. We explicitly enforce this constraint by first introducing the set $\mathbf{I}(z) = \{i: y_i = z\}$ of indices associated with samples with label z and $\boldsymbol{\alpha}_{\mathbf{I}(z)}$, which is $\boldsymbol{\alpha}$ reduced to the indices of $\mathbf{I}(z)$.

Finally, the entire minimization problem is then defined as

$$\underset{\boldsymbol{\alpha}, \boldsymbol{\beta}, \beta_0}{\text{Argmin}} \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \beta_0) \quad \text{s. t.} \quad \|\tilde{\boldsymbol{\beta}}\|_0 \leq r \quad \text{and} \quad \|\boldsymbol{\alpha}_{\mathbf{I}(-1)}\| = k \quad \text{and} \quad \|\boldsymbol{\alpha}_{\mathbf{I}(+1)}\| = k$$

where $\tilde{\boldsymbol{\beta}} := \left(\|\hat{\boldsymbol{\beta}}^1\|_2^2, \dots, \|\hat{\boldsymbol{\beta}}^1\|_2^2 \right)^\top$ groups the weight vectors by computing the ℓ_2 -norm of the columns of $\hat{\boldsymbol{\beta}}$. This problem of joint sample and feature selection is solved using the Block Coordinate Decent (BCD) and Penalty Decomposition as in (20, 24).

Results of Comparison

Supplement Table S1 lists the accuracy of the diagnostic score generated by the proposed novel machine learning approach compared to two state-of-the art methods: the JFSS (25) and the SFS (26)+SVM (27). The former is a recent method for Joint Feature-Sample Selection (JFSS) applied to neuroimaging data, and the latter is a widely used method that applies Sparse Feature Selection (SFS) followed by a Support Vector Machine (SVM) classifier. While all approaches were significantly accurate ($p < 0.001$; Fisher exact test (21)) in predicting comorbidity, only the proposed score was significantly accurate in all three diagnoses. For all three diagnosis-specific group comparisons, the proposed novel machine learning approach was also significantly more accurate than the other two methods (DeLong test (28), p -value < 0.05). The difference in accuracy was especially larger for the two diagnosis-specific group comparisons, where the number of samples per group were highly imbalanced (*i.e.*, CTRLs (N=199) vs. HIV (N=68) and CTRLs (N=199) vs. AUD+HIV (N=60)). These findings further support the proposed data-driven analysis for identifying *diagnostic patterns* and *scores* from MRI data.

Supplement Table S1: Classification balanced accuracy (BAcc), Sensitivity (Sen) and Specificity (Spe) for each diagnosis-specific group comparisons. * indicates methods with $p < 0.001$ in a Fisher exact test.

	CTRL vs. AUD			CTRL vs. HIV			CTRL vs. AUD+HIV		
	BAcc	Sen	Spe	BAcc	Sen	Spe	BAcc	Sen	Spe
Our Method	70.1*	0.71	0.69	76.2*	0.78	0.75	78.1*	0.79	0.77
JFSS	66.0	0.71	0.61	71.5*	0.71	0.72	72.2*	0.74	0.70
SFS+SVM	64.1	0.56	0.72	68.4	0.59	0.79	70.1*	0.64	0.76

Discussion

An important factor in the prediction accuracy of the machine learning approach is the accuracy of the measurements extracted from the MRI data. To minimize the labor associated with the analysis, a fully automatic approach extracted the measurements based on FreeSurfer segmentations. These label maps are imperfect as, for example, they underestimated subcortical regions and white matter hypointensities. One could have overcome those deficiencies via manual editing. However, this would have also increased the risk of biasing the study to reporting overly optimistic findings as ensuring consistent quality across the resulting measurements is challenging.

A limitation of this work was that due to the longitudinal nature of data gathering for this study, some cognitive tests were replaced by others, as described in the main paper. Finally, our study restrained from relating our diagnostic specific patterns to genetic mechanisms. This relation could be of interest as the presence of certain genotypes can modulate the effect of alcohol of abuses on HIV disease progression (29).

References

1. Springer SA, Di Paola A, Barbour R, Azar MM, Altice FL (2018): Extended-release Naltrexone Improves Viral Suppression Among Incarcerated Persons Living with HIV and Alcohol use Disorders Transitioning to the Community: Results From a Double-Blind, Placebo-Controlled Trial. *J Acquir Immune Defic Syndr.* 79:92-100.
2. Pfefferbaum A, Zahr NM, Sassoon SA, Kwon D, Pohl KM, Sullivan EV (2018): Accelerated and Premature Aging Characterizing Regional Cortical Volume Loss in Human Immunodeficiency Virus Infection: Contributions From Alcohol, Substance Use, and Hepatitis C Coinfection. *Biol Psychiatry Cogn Neurosci Neuroimaging.*
3. Coupe P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C (2008): An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans Med Imaging.* 27:425-441.
4. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. (2010): N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging.* 29:1310-1320.
5. Rohlfing T, Russakoff DB, Maurer CR, Jr. (2004): Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE Trans Med Imaging.* 23:983-994.
6. Smith SM (2002): Fast robust automated brain extraction. *Hum Brain Mapp.* 17:143-155.
7. Cox RW (1996): AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 29:162-173.
8. Sadananthan SA, Zheng W, Chee MW, Zagorodnov V (2010): Skull stripping using graph cuts. *Neuroimage.* 49:225-239.
9. Iglesias JE, Liu CY, Thompson PM, Tu Z (2011): Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Trans Med Imaging.* 30:1617-1634.
10. Dale AM, Fischl B, Sereno MI (1999): Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage.* 9:179-194.
11. Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012): Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage.* 61:1402-1418.
12. Fischl B (2012): FreeSurfer. *Neuroimage.* 62:774-781.
13. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage.* 15:273-289.
14. Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A (2010): The SRI24 multichannel atlas of normal adult human brain structure. *Hum Brain Mapp.* 31:798-819.
15. Avants BB, Epstein CL, Grossman M, Gee JC (2008): Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal.* 12:26-41.
16. Madsen H, Thyregod P (2010): *Introduction to general and generalized linear models.* CRC Press.
17. Wan X, Liu J, Cheung WK, Tong T (2014): Learning to improve medical decision making from imbalanced data without a priori cost. *BMC Med Inform Decis Mak.* 14:111.
18. Székely GJ, Rizzo ML (2013): The distance correlation t-test of independence in high dimension. *Journal of Multivariate Analysis.* 117:20.
19. Arlot S, Celisse A (2010): A survey of cross-validation procedures for model selection. *Statistics surveys.* 4:39.
20. Adeli E, Li X, Kwon D, Zhang Y, Pohl KM (2018): Logisitc Regression Confined by Cardinality-Constrained Sample and Feature Selection. *Accepted in IEEE Transactions to Pattern Analysis and Machine Intelligence.*

21. Fisher RA (1935): The logic of inductive inference. *Journal of the Royal Statistical Society*. 98:43.
22. Zhang Y, Kwon D, Esmaeili-Firidouni P, Pfefferbaum A, Sullivan EV, Javitz H, et al. (2016): Extracting patterns of morphometry distinguishing HIV associated neurodegeneration from mild cognitive impairment via group cardinality constrained classification. *Hum Brain Mapp*. 37:4523-4538.
23. Hastie T, Tibshirani R, Wainwright M (2015): *Statistical learning with sparsity: the lasso and generalizations*. CRC Press.
24. Zhang Y, Kwon D, Pohl KM (2017): Computing group cardinality constraint solutions for logistic regression problems. *Med Image Anal*. 35:58-69.
25. Adeli E, Shi F, An L, Wee CY, Wu G, Wang T, et al. (2016): Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *Neuroimage*. 141:206-219.
26. Nie F, Huang H, Cai X, Ding CH (2010): Efficient and Robust Feature Selection via Joint L21-Norms Minimization. *Neural Information Processing Systems*.9.
27. Chang C-C, Lin C-J (2011): LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*. 2:27.
28. DeLong ER, DeLong DM, Clarke-Pearson DL (1988): Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 44:837-845.
29. Agudelo M, Khatavkar P, Yndart A, Yoo C, Rosenberg R, Devieux JG, et al. (2014): Alcohol abuse and HIV infection: role of DRD2. *Curr HIV Res*. 12:234-242.