STUDY PROTOCOL

ClinicalTrials.gov **NCT07590154**

Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders

---

Cross-sectional Functional Stratification Based on Psychometric Profiling and Machine Learning in Patients With Substance Use Disorders (SUD)

SISAP-TUS

Document Type:	**Study Protocol**
Protocol ID:	**SISAP-TUS-EFTP-ML-2024**
ClinicalTrials.gov:	**NCT07590154**
UTN:	**U1111-1346-9830**
Document Version:	**1.0**
Document Date:	**14 March 2024**
Documentation Consolidation Date:	**19 September 2026**

Sponsor-Investigator

Lauro Gutiérrez Castro
Under The Tree Therapeutic Community

Supporting Information for ClinicalTrials.gov

Document Control

Documentary nature.
This document is a documented version of the executed study protocol. It describes the study as executed and presents its objectives, design, population, procedures, variables, outcome measures, analysis plan, ethical considerations, limitations, and data-governance provisions in a structured format.
The protocol version identified in this document was approved by the institutional ethics committee on 14 March 2024. The present document was subsequently consolidated and formatted for ClinicalTrials.gov supporting documentation on 19 September 2026.
The present documentation distinguishes the original executed protocol from subsequent dataset expansions used for later analyses. No substantive protocol amendment was introduced through the documentary consolidation.

Field	Value
Protocol ID	SISAP-TUS-EFTP-ML-2024
ClinicalTrials.gov ID	NCT07590154
UTN	U1111-1346-9830
Official Title	Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders
Brief Registered Title	Cross-sectional Functional Stratification Based on Psychometric Profiling and Machine Learning in Patients With Substance Use Disorders (SUD)
Acronym	SISAP-TUS
Document Type	Study Protocol
Document Version	1.0
Document Date	14 March 2024
Original Executed Protocol Date	1 March 2024
Ethics Approval Date	14 March 2024
Documentation Consolidation Date	19 September 2026
Sponsor-Investigator	Lauro Gutiérrez Castro, Under The Tree Therapeutic Community
Document Status	Current documented version of the executed study protocol
Replaces	None; first documented version of the executed protocol
Distribution Status	Public supporting information for ClinicalTrials.gov

Document History

Table 2: Document history.

Version	Date	Status	Description
1.0	14 March 2024	Approved / Executed	Ethics-approved protocol version governing the executed study.

The present PDF was subsequently consolidated for ClinicalTrials.gov supporting documentation on 19 September 2026 without introducing a substantive protocol amendment.

Contents

Document Control **1**

Document History **1**

List of Abbreviations **6**

1 Protocol Synopsis **8**

2 Background and Rationale **9**
2.1 Study Publication 9

3 Study Objectives **9**
3.1 Primary Objective 9
3.2 Secondary Objectives 10
3.3 Exploratory Objectives 10

4 Study Design **10**

5 Population and Eligibility Criteria **10**
5.1 Description 10
5.2 Inclusion Criteria 11
5.3 Exclusion Criteria 11
5.4 Sampling 11
5.5 Sample Size 11
5.6 Population and Sampling Considerations 11

6 Recruitment and Participant Flow **12**

7 Instruments and Variables **12**
7.1 Psychometric Instruments 13
7.2 Scoring Rules and Normative Verification by Instrument 13
7.2.1 BIS-11 13
7.2.2 BSL-23 and Behavioral Supplement 14
7.2.3 GAD-7 14
7.2.4 STAXI-2 (Reduced Administered Form) 14
7.2.5 BADS 15
7.2.6 EROS 15
7.3 OCI-R (Descriptive Marker, Not Included in the Autoencoder) 15
7.4 Baseline Clinical, Sociodemographic, and Substance-Use History Variables 16
7.5 Scoring Integrity and Data Provenance 16
7.6 Demographic and Clinical Variables 16
7.7 Model Input Variables 17

8 Study Procedures **17**

9 Outcomes and Operational Definitions **18**
9.1 Temporal Framework of Outcome Measures 19
9.2 Nature of the Outcome Measures 20

10 Data Management and Quality Control **20**

11 Statistical Analysis **20**
11.1 Data Preparation . . . 20
11.2 Autoencoder Representation . . . 21
11.3 Comparison with Principal Component Analysis . . . 21
11.4 Clustering Using Gaussian Mixture Models . . . 21
11.5 Interpretation of Latent Dimensions . . . 21
11.6 Inferential Nature of the Analysis . . . 21

12 Ethical Considerations **21**
12.1 Ethical and Regulatory Principles . . . 22
12.2 Institutional Ethics Review . . . 22
12.3 Informed Consent . . . 22
12.4 Privacy and Confidentiality . . . 23
12.5 Risks and Protection Measures . . . 23
12.6 Use of Results and Machine Learning Model . . . 23

13 Risks and Benefits **24**

14 Data Monitoring and Follow-up **25**
14.1 Data Monitoring . . . 25
14.2 Data Verification and Quality Control . . . 25
14.3 Participant Follow-up . . . 26
14.4 Adverse Events and Unanticipated Events . . . 26

15 Protocol Amendments **26**
15.1 Modification History . . . 27

16 Publication, Data Sharing, and Intellectual Property **27**
16.1 Publication . . . 27
16.2 Data Sharing . . . 27
16.3 Access Criteria and Procedure . . . 28
16.4 Intellectual Property and Restricted Materials . . . 28

17 Timeline **28**

18 Anticipated Methodological Limitations **29**

19 References **30**

20 Appendices **32**
20.1 Appendix A — Supporting Study Documentation . . . 32
20.2 Appendix B — Publication Reference . . . 32
20.3 Appendix C — Administrative Information . . . 33

List of Tables

2 Document history. 1
3 Protocol synopsis. 8
4 Participant recruitment and flow. 12
5 Registered outcomes and operational definitions. 19
6 Risks, mitigation measures, and potential benefits. 24
7 Study timeline. 28

List of Figures

List of Abbreviations

Abbreviation	Meaning
BIC	Bayesian Information Criterion
CSR	Clinical Study Report
CV	Coefficient of Variation
GMM	Gaussian Mixture Model
ICF	Informed Consent Form
MSE	Mean Squared Error
NCT	ClinicalTrials.gov Identifier
PCA	Principal Component Analysis
SAP	Statistical Analysis Plan
SD	Standard Deviation
SUD	Substance Use Disorder
UTN	Universal Trial Number

1 Protocol Synopsis

Table 3: Protocol synopsis.

Element	Description
Study type	Observational, non-interventional, cross-sectional, single-center study based on a patient registry.
Observational model	Case-Only series.
Temporal perspective	Cross-sectional.
Population	Adults aged ≥ 18 years with substance use disorder (SUD) according to DSM-5, receiving residential treatment.
Study site	Under The Tree residential therapeutic community, Ajijic, Jalisco, Mexico.
Sample size	**155 participants** (actual enrollment); initial planned sample size: 155.
Study start	25 March 2024 (actual).
Recruitment period	February 2024 to December 2025.
Primary completion	18 February 2026 (actual).
Study completion	22 April 2026 (actual).
Intervention	None. No drug, device, behavioral therapy, or other intervention was administered.
Biological samples	None retained.
Primary objective	To learn a 12-dimensional latent representation from standardized psychometric and clinical variables using an autoencoder and to evaluate the stability of that representation across repeated training runs.
Secondary objectives	To compare reconstruction performance with principal component analysis (PCA); characterize the clinical meaning of the latent dimensions; explore subgroups using Gaussian mixture models (GMM); and assess their stability and interpretability.
Registered primary outcome	Scores on the 12 latent dimensions obtained from the bottleneck layer of the autoencoder trained on 21 standardized clinical variables, using an ensemble of the five best-performing runs.
Registered secondary outcomes	(2) GMM cluster membership; (3) reconstruction pseudo-R^2; (4) reconstruction mean squared error (MSE); (5) coefficient of variation (CV) of the MSE across 30 runs; (6) cross-validated R^2 using 5-fold cross-validation repeated 3 times; and (7) variance explained by the first 12 principal components.
Statistical analysis	Separate document: *Statistical Analysis Plan (SAP)*, version 1.0, aligned with the audited analysis code.
Risks	Mild emotional discomfort or fatigue while completing questionnaires; no physical risk. Psychological support was available.
Benefits	No direct benefit to participants. Potential indirect benefit through transdiagnostic functional characterization.
Ethics review	As declared in the public registry (NCT07590154, *Oversight* section). Committee and approval number: `[TO BE COMPLETED]`.
Applicable declaration	Declaration of Helsinki and applicable local regulations.
Data sharing	Yes, following publication and de-identification (see section 16).

2 Background and Rationale

Substance use disorders (SUDs) exhibit clinical heterogeneity that extends beyond traditional diagnostic categories. Patients with the same formal diagnosis may present markedly different psychopathological profiles and treatment needs. Recent literature has highlighted two central transdiagnostic processes, **emotion dysregulation** and **impulsivity**, because of their consistent associations with the onset, maintenance, and relapse of substance use [1, 2]. This perspective converges with the *Research Domain Criteria* framework, known as RDoC, which proposes organizing psychopathology around functional domains and dimensional constructs [3].

In parallel, machine learning has opened new approaches for analysis in mental health research. However, the addiction literature remains largely dominated by **supervised** approaches that require clinical labels and do not fully exploit the latent structure of the data [4]. The **autoencoder**, as an unsupervised learning method, learns compact and nonlinear representations that may preserve complex relationships among psychological and behavioral variables more effectively than some classical linear methods, such as principal component analysis (PCA) [5].

The latent space obtained through an autoencoder can be combined with probabilistic clustering methods such as **Gaussian mixture models** (GMM). Unlike k-means clustering, GMMs estimate membership probabilities and can represent groups with different sizes, shapes, and covariance structures [9]. In SUD, these subgroups may differ in impulsivity, emotion regulation, or comorbidity, with potential implications for treatment personalization. However, their clinical interpretation requires rigorous validation of stability, replicability, and clinical relevance in order to avoid interpreting noise as clinically meaningful structure [11].

Study gap. There is limited application of unsupervised deep learning in SUD, limited integration of autoencoders with probabilistic clustering, and a need for deep phenotyping strategies based on detailed psychometric data. In response to this gap, the study proposes a two-stage framework: first, an autoencoder to obtain a latent representation of the data, followed by a Gaussian mixture model to identify potentially clinically differentiated subgroups among patients receiving residential treatment.

2.1 Study Publication

The main study publication is:

L. Gutiérrez Castro et al. “Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders”. In: *Artificial Intelligence for Mental Health. ICAIMH 2026*. Vol. 3062. Communications in Computer and Information Science. Springer, 2026. DOI: 10.1007/978-3-032-30396-7_12.

3 Study Objectives

3.1 Primary Objective

To learn a **12-dimensional latent representation** from standardized psychometric and clinical variables using an autoencoder model and to **evaluate the stability of that representation** through repeated training procedures.

3.2 Secondary Objectives

1. To compare the reconstruction performance of the autoencoder with that of **principal component analysis (PCA)** applied to the same variables and with the same dimensionality.
2. To **characterize the clinical meaning** of the latent dimensions through correlations with the original variables and structured clinical labeling.
3. To **explore patient subgroups** using Gaussian mixture models (GMM) applied to the latent space.
4. To evaluate the **stability and interpretability** of the identified subgroups.

3.3 Exploratory Objectives

The following analyses were documented as exploratory analyses and were not registered as outcome measures:

- Psychometric structure and properties of the administered instruments, including reliability assessed by ω, item-level analyses, and exploratory factor structure.
- Network structure of the clinical variables, estimated using EBICglasso regularization, including centrality and stability analyses.
- Methodological comparison of clustering solutions, including hierarchical, partition-based, and fuzzy approaches.
- Identification of functional profiles and their relationship with cumulative clinical burden.
- Exploratory multicriteria decision analysis oriented toward clinical support, without clinical use of the resulting analyses.

4 Study Design

This was an **observational, cross-sectional, non-interventional, single-center study** conducted in a residential substance use treatment facility. Each participant underwent a **single evaluation session** with a *Target Follow-Up Duration* of 1 day.

No intervention was assigned. The study consisted exclusively of the administration of standardized psychometric instruments and the collection of demographic and clinical variables.

The study was publicly registered **before the dissemination of the study results**.

5 Population and Eligibility Criteria

5.1 Description

Adults with a diagnosis of substance use disorder (SUD) according to DSM-5 who were admitted to a residential detoxification and rehabilitation center, with consecutive recruitment between February 2024 and December 2025.

Note on subsequent database updating. Following the recruitment period described in this document, the institutional database continued to incorporate and document clinical information from additional participants. A subsequent publication from the study uses an expanded dataset of 166 participants. This difference reflects the progressive updating of the database and the use of an expanded dataset for subsequent analyses. The sample size

and recruitment period reported in this document correspond to the specific methodological cutoff of this protocol.

5.2 Inclusion Criteria

Participants were required to meet all of the following criteria:

1. Diagnosis of substance use disorder according to DSM-5, confirmed by a psychiatrist or clinical psychologist.
2. Age $\geq$ 18 years.
3. Admission to the residential treatment center at the time of assessment.
4. Ability to complete the psychometric questionnaires independently.
5. Written informed consent.

5.3 Exclusion Criteria

Participants were excluded if any of the following criteria were met:

1. Active psychotic disorder, such as schizophrenia or delusional disorder, that was not pharmacologically stabilized.
2. Severe cognitive impairment, including dementia or severe traumatic brain injury, that prevented adequate comprehension of the questionnaire items.
3. Language barriers or illiteracy that prevented self-administration of the psychometric scales.
4. Scheduled discharge from the residential center within 7 days following the assessment date.

5.4 Sampling

The study used **non-probability consecutive sampling**. Healthy volunteers were not included. The maximum recorded age was 60 years. Participants of all sexes were eligible.

5.5 Sample Size

A total of 155 participants were enrolled, corresponding to the actual enrollment and the planned sample size. The sample size was determined by the availability of eligible cases during the recruitment period and by the feasibility of administering the psychometric battery. **No formal statistical power calculation** was performed because the study objectives were exploratory and non-confirmatory. The limitations associated with the sample size are described in section 18.

5.6 Population and Sampling Considerations

The study population consisted exclusively of adults with substance use disorder receiving residential treatment. Accordingly, the sample represents the clinically accessible population during the recruitment period at the participating center and was not intended to constitute a probability sample of the general population of individuals with substance use disorder.

Recruitment was consecutive among cases meeting the established eligibility criteria. Inclusion was determined according to the protocol criteria, with no inclusion of healthy volunteers or individuals outside the target clinical population.

The sample size of 155 participants corresponds to the number of participants actually enrolled during the specified study period and matches the planned sample size. No formal statistical power calculation was performed because the objectives were exploratory and non-confirmatory.

Accordingly, the precision of the estimates and the ability to detect associations or differences should be interpreted in light of the limitations inherent to the size and composition of the sample.

The characteristics of this population, including its origin from a single residential treatment setting and the use of non-probability consecutive sampling, should be considered when interpreting the transferability of the findings to other clinical populations, institutions, or care settings.

6 Recruitment and Participant Flow

Table 4: Participant recruitment and flow.

Stage	n	Observations
Assessed for eligibility	**158**	Consecutive records from the study center
Excluded	**3**	Due to psychotic symptoms
Included and assessed	**155**	Single cross-sectional assessment session
Complete data for the 21 model variables	**155**	Dataset used for the published model analysis
Analyzed (primary analysis set)	**155**	See SAP §4

The complete participant flow, including the reason for exclusion at each stage, is documented in the *Clinical Study Report* in the participant disposition section and is reported in the results module of the public registry.

Note on subsequent database updating. Following the methodological cutoff corresponding to this document, the database continued to be updated through the incorporation and documentation of additional participants. A subsequent publication uses an expanded dataset of **166 participants**. This difference corresponds to a subsequent database update and does not modify the participant flow or sample size of the primary analysis set declared in this document.
The additional participants incorporated after the methodological cutoff were used in subsequent analyses and in model updating according to the analytical procedures corresponding to each version of the dataset. Therefore, the sample size of **155 participants** presented in this section specifically identifies the cohort and data cutoff used for the primary analysis described in this document, whereas **166 participants** corresponds to a dataset expanded for subsequent analyses.

7 Instruments and Variables

The instruments and variables used in the study are described according to the versions and scoring criteria established for the protocol. For each instrument, the derived variables, scoring range, and, when applicable, item recoding or reverse-scoring procedures are specified.

Instrument scoring was performed using procedures specified and documented in advance. Transformations, scoring rules, range checks, and consistency-control procedures form part of the methodological and reproducibility documentation associated with the study.

The computational implementation, verification procedures, and complementary technical documentation are available through the public study documentation, subject to the scope of publication and applicable data-protection restrictions.

7.1 Psychometric Instruments

The study battery comprised six standardized psychometric instruments. Their main characteristics are summarized below.

BIS-11 *Barratt Impulsiveness Scale.* Measures impulsivity across attentional, motor, and non-planning domains [12]. The instrument comprises 30 items rated on a 1–4 response scale, with a theoretical total range of 30–120.

STAXI-2 *State-Trait Anger Expression Inventory-2.* Assesses state and trait anger and its expression, including dimensions relevant to emotion regulation [13]. The standard instrument contains 57 items rated on a 1–4 scale. The reduced form administered in the present study contained **49 items**; subscale ranges depend on the specific scoring structure.

BADS *Behavioral Activation for Depression Scale.* Assesses behavioral activation and behavioral avoidance [15]. The instrument comprises 25 items rated on a 0–6 scale, with a theoretical total range of 0–150.

BSL-23 *Borderline Symptom List-23.* Assesses severity of borderline symptomatology [17]. The core instrument contains 23 items rated on a 0–4 scale, with a theoretical total range of 0–92. An additional **11-item behavioral block** was also administered, with a theoretical range of 0–44. The integrated score therefore has a theoretical maximum of 136.

GAD-7 *Generalized Anxiety Disorder-7.* Assesses symptoms of generalized anxiety [14]. The instrument contains 7 items rated on a 0–3 scale, with a theoretical total range of 0–21.

EROS *Environmental Reward Observation Scale.* Assesses perceived environmental reinforcement and reward deficit [16]. The instrument contains 10 items rated on a 1–4 scale, with a theoretical total range of 10–40.

The item sets for each subscale, exact scoring formulas, and comparisons with the corresponding instrument manuals are described in the supporting methodological documentation.

7.2 Scoring Rules and Normative Verification by Instrument

7.2.1 BIS-11

- **Official structure:** 30 items, 1–4 response scale, and **11 reverse-scored items** (1, 7, 8, 9, 10, 12, 13, 15, 20, 29, and 30; transformed as `5 - value`), with a theoretical total score of 30–120.
- **Official second-order factors** (Patton, Stanford & Barratt, 1995; impulsivity.org): *Attentional* = Attention + Cognitive Instability; *Motor* = Motor + Perseverance; *Non-planning* = Self-Control + Cognitive Complexity.

- **Correction applied in the analytical pipeline:** the historical capture summed all 30 items without reversing the 11 reverse-scored items and grouped them into three ad hoc blocks (`Cognitive Impulsivity`, `Motor Impulsivity`, and `Non-planning Impulsivity`) that do not correspond to any official factor grouping. The normative recalculation applies the official reverse-scoring procedure and the three-factor second-order structure. The discrepancies are retained in the documentation to audit the impact of capture noise on the autoencoder latent representations.
- **`Attention` subscale:** the local formula (`SUM(P6,P9,P11,P20,P28)`) does not correspond to the standard first-order structure; the normative recalculation follows the official five-item structure.

7.2.2 BSL-23 and Behavioral Supplement

- **Core manual:** Bohus et al. (2009). Twenty-three items, 0–4 response scale, total score 0–92; operational severity = `Total/23`.
- **Behavioral supplement:** 11 items assessing ordinal frequency from 0–4 (*None, once, 2–3 times, 4–6 times, daily or more*), intended for risk triage within the therapeutic community. This is a local extension block and is orthogonal to the emotional-distress thresholds derived from the 23 core items.
- **Integrated total:** `Total + Behavioral` (23 + 11 = 34 items). The normalization denominator is corrected to the actual theoretical maximum of **136** (`34 x 4`), correcting the previous saturation bias that divided by 144.
- **`GENERAL`:** self-rated percentage of subjective well-being on an analogue 0%–100% scale; its negative correlation with total distress is consistent with the expected nomological validity pattern.

7.2.3 GAD-7

- **Official structure:** 7 items, 0–3 response scale, total score 0–21; severity cut points at 5, 10, and 15 (Spitzer et al., 2006). The `Interpretation` field follows these thresholds.
- **Z-score:** parameterized in the historical capture using an **initial pilot window of `n = 10`**, with $\mu = 7.4000$ and $\sigma = 5.8538$, compared with the full cohort (`n = 166`; $\mu = 8.0843$, $\sigma = 5.9027$).
- **Status:** declared pilot-window artifact. This is a constant affine linear transformation (`z_pilot = a x + b`) that can be absorbed by the first layer of the autoencoder through its affine transformation and bias during gradient descent, without introducing nonlinear distortion or altering the topology of the latent covariance structure. The production pipeline performs global standardization across the cohort.

7.2.4 STAXI-2 (Reduced Administered Form)

- **Normative manual:** Spielberger (1999). Standard form: 57 items, Likert scale 1–4; expression/control scales contain **8 items** each; IEI = `(AX-O + AX-I) - (AC-O + AC-I) + 36`.
- **Field-administered form: 49 items** (8 items suppressed), with 6 items per expression/control subscale. The IEI formula is applied isometrically to the reduced subscales.
- **Forensic audit of ordinal bounds: 29 cells outside the 1–4 range** were detected, concentrated in two records: one record contained 27 items captured with value `5` (scale displacement) and another contained one item with value `11` (keyboard-entry artifact). These values are neutralized during preprocessing through domain-bounding and winsorization at the 1st/99th percentiles before Z-score standardization; they do not enter the latent

space as raw values and do not contaminate model training.

- **Comparability boundary:** percentiles from this reduced form should not be directly equated with normative reference values from the general population manual for the 57-item form without explicitly acknowledging the structural truncation bias introduced by the 49-item form.

7.2.5 BADS

- **Normative manual and Spanish adaptation:** Kanter et al. (2007); Barraca, Pérez-Álvarez & Lozano Bleda (2011).
- **Structure:** 25 items, Likert scale 0–6, with four subscales: `Activation` (7 items), `Avoidance/Rumination` (8), `Work Impairment` (5), and `Social Impairment` (5). The local formulas match the normative structure item by item.
- **Integrity: 0 out-of-range cells** (strict compliance with the 0–6 domain). This is the cleanest psychometric block in the capture dataset and serves as an empirical reference anchor for cross-instrument comparison.

7.2.6 EROS

- **Normative manual and Spanish adaptation:** Armento & Hopko (2007); Barraca & Pérez-Álvarez (2010).
- **Structure:** 10 items, Likert scale 1–4, total score 10–40. Reverse-scoring is applied at the capture layer for each observation; `PD = SUM(P1:P10)`.
- **Standardization:** `Z = (PD - 26.23) / 6.495`, anchored to the normative statistics of the Spanish adaptation (Barraca & Pérez-Álvarez, 2010); `T-score = 50 + 10 Z`.
- **Neutralization of atypical bounds:** 19 cells with value `0` (outside the 1–4 range) are masked or bounded according to the normative limits through winsorization or thresholding before standardization.
- **Documentary correction:** the historical attribution associating the instrument with the *Pleasant Events Schedule* (MacPhillamy & Lewinsohn) is corrected. The appropriate source is Armento & Hopko (2007), with Spanish validation by Barraca & Pérez-Álvarez (2010).

7.3 OCI-R (Descriptive Marker, Not Included in the Autoencoder)

- **Normative manual:** Foa et al. (2002). Eighteen items, 0–4 response scale, total score 0–72, with six orthogonal three-item subscales.
- **Empirical evidence of overlap:** Huppert et al. (2007), *Behaviour Research and Therapy*, 45(10):2499–2510 (PMID: 16814981).
- **Local subscale partition:** the six local subscales (`Washing`, `Checking`, `Ordering`, `Obsessions`, `Hoarding`, and `Neutralization`) partition the 18 items in a bijective manner, exhaustively and without overlap, but **do not correspond** to the official factorial assignment reported by Foa et al. (2002). This is formally declared as a version-mapping artifact.
- **Scope boundary:** the OCI-R operates exclusively as a descriptive clinical characterization marker and **is not included among the autoencoder input variables**. Any divergence in the local subscale partition is mathematically inert and has no gradient propagation into the model.

7.4 Baseline Clinical, Sociodemographic, and Substance-Use History Variables

- **Nature:** these variables derive from the initial exploratory clinical interview conducted during admission to the therapeutic community. They constitute descriptive anamnesis metadata and do not represent latent scores, normed ordinal scales, or psychometric isomorphisms.
- **Fields:** `Age`, `Sex`, `Time_using`, `Previous_treatments`, `Reason_for_discharge`, and binary/categorical vectors for substances of impact (7 dummy indicators of presence/absence).
- **Conditional rescue protocol (Variables 20–24):** when spreadsheet asymmetries occur across worksheets, the pipeline first queries the `BADS` matrix; if the cell is null, it then conditionally evaluates `EROS`. The remaining 11 missing values are imputed using the cohort median.
- **Normative directive:** substance-use vectors are strictly prohibited from being interpreted as indices of toxicological severity, withdrawal burden, or consumption frequency. Their role is limited to cohort stratification and exogenous descriptive control. Under the restricted white-box architecture, they do not participate in the latent representation manifold of Phase 1.
- **Audited exception (`compuesto_bis`, Variable 7):** a substring-filtering rule (`"bis"` within `"cannabis"`) injected a toxicological dummy variable into the composite field. The phenomenon is explicitly audited and documented without silently altering the experimental traces.

7.5 Scoring Integrity and Data Provenance

1. **Centralized normative recalculation:** official scoring formulas are implemented through the centralized analysis workflow; discrepancies in the historical capture are retained in the documentation to audit their impact on the autoencoder latent representations.
2. **Preprocessing filters:** domain bounding and 1st/99th-percentile winsorization are applied before Z-score standardization. This neutralizes the 29 anomalous STAXI-2 cells and the 19 EROS cells with value `0`, preventing abnormal variance spikes and preserving the stability of the latent representation.
3. **Traceability:** item sets, formulas, and granular audit traces are maintained in the study's supporting methodological documentation. Aggregated numerical evidence is reported in the publication, including the reconstruction fidelity by construct and the latent structure, clustering, and characterization analyses.
4. **Consistency with the registry:** the published analytical `n` is **155**; internal data-cleaning figures, including 166 records in the expanded database, 141 records with names, and 64 records across the seven worksheets, are documented as audit traces and not as the study denominator. Any recalculation of the BIS-11 total score or the GAD-7 z-score should be explained in the corresponding methodological documentation and reflected in the results module.

7.6 Demographic and Clinical Variables

Age, sex, primary substance of use, years of substance use, number of previous treatment episodes, discharge status, and **substance-impact columns**. The latter were derived from the **initial exploratory interview** conducted at the center and did not originate from a standardized instrument; this is explicitly declared in all publications.

7.7 Model Input Variables

The published model used **21 standardized variables**, resulting from the data-reduction procedure described in SAP §6, including removal of variable pairs with Pearson correlation greater than 0.85. The complete variable dictionary, including variable origin and transformation, is maintained in the study's supporting methodological documentation.

8 Study Procedures

The study procedures were conducted sequentially, from diagnostic confirmation and eligibility assessment through data capture, verification, and processing for analysis. Participant assessment consisted of a single cross-sectional session. Data processing and analysis were subsequently performed on the analytical dataset according to the Statistical Analysis Plan. Research activities were kept separate from clinical decision-making, and the machine learning model had no intervention or automated clinical decision-making function.

1. **Diagnostic confirmation** by a psychiatrist or clinical psychologist, according to DSM-5 criteria.
2. **Verification of eligibility criteria** and written informed consent before inclusion in the study.
3. **Single cross-sectional assessment session.** A self-administered battery of six psychometric instruments was administered, and the demographic and clinical variables defined for the study were recorded. Assessment occurred at a single time point for each participant during the period established for the analysis.
4. **Data capture and traceability.** Responses were recorded in the center's study registry, with one sheet for each instrument and using a pseudonymized case identifier. The identifier used for analysis did not directly contain the participant's personal identifying information.
5. **Quality control and integrity verification.** Checks of completeness and response ranges were performed, together with audits of data provenance and consistency. The integrity of source files and the analytical dataset was verified using hash functions to detect unauthorized modifications or discrepancies between versions.
6. **Construction of the analytical dataset.** Data were processed according to the procedures established in the Statistical Analysis Plan, including extraction, score calculation, data cleaning, imputation, and standardization, in accordance with the corresponding methodological specifications (SAP §5–§6).
7. **Statistical and machine learning analysis.** Data were analyzed according to the SAP using an autoencoder to obtain a latent representation, comparison with principal component analysis, and clustering using Gaussian mixture models. Additional analyses specified in the SAP were also performed.
8. **Absence of clinical intervention derived from the model.** The machine learning model was used exclusively for research and analytical purposes. No score, latent representation, classification, or clustering generated by the model was used to make automated clinical decisions, modify treatment, determine eligibility for interventions, or replace the clinical judgment of treating personnel.
9. **Separation between research and clinical care.** The analytical procedures and model-generated results did not constitute a therapeutic intervention and were not part of the routine clinical decision-making process during participant assessment.

9 Outcomes and Operational Definitions

The following definitions correspond to the outcome measures registered for study NCT07590154 and operationally describe the variables used to evaluate the analytical objectives of the study. The outcome set was derived from the cross-sectional assessment session and the subsequent processing of the clinical and psychometric variables defined in the protocol.

The **primary outcome measure** corresponds to the latent representation obtained through the autoencoder. The additional outcome measures describe cluster membership and the performance, stability, and generalizability of the machine learning procedure. These latter measures constitute analytical model metrics and should not be interpreted as clinical outcomes or as direct evidence of therapeutic benefit.

Table 5: Registered outcomes and operational definitions.

#	Outcome	Operational Definition	Unit
1	**Primary: scores on the 12 latent dimensions**	Twelve continuous dimensions extracted from the bottleneck layer of a symmetric autoencoder trained on 21 standardized clinical variables. Final scores are obtained by averaging the predictions of the ensemble comprising the five best-performing runs.	z; mean 0 and SD 1 in the training sample
2	**GMM cluster membership**	Categorical assignment to clusters from a Gaussian mixture model with full covariance matrix, fitted to the 12-dimensional latent space. The number of clusters is determined using the Bayesian information criterion and clinical interpretability criteria. This measure is exploratory and does not imply that the identified clusters constitute discrete clinical subtypes.	Nominal
3	**Reconstruction pseudo-R^2**	$1 - (MSE_{\text{model}}/MSE_{\text{null}})$, where MSE_{null} corresponds to the error obtained by predicting the mean of each variable. Calculated for the ensemble and for each of the 30 independent runs.	Proportion, 0–1
4	**Reconstruction MSE**	Mean squared difference between the 21 original standardized variables and their corresponding reconstructions generated by the autoencoder.	MSE, dimensionless
5	**Reconstruction MSE CV**	Coefficient of variation calculated as $CV = SD/mean$ of the reconstruction MSE across the 30 independent runs with different random seeds. This measure evaluates model performance stability in the presence of stochastic variation during training.	Percentage
6	**Cross-validated R^2**	Mean and standard deviation of the R^2 obtained using 5-fold cross-validation repeated 3 times, using the same model architecture and hyperparameters.	Proportion, 0–1
7	**Variance explained by 12 principal components**	Total proportion of variance explained by the first 12 principal components obtained from PCA applied to the same 21 standardized clinical variables. This measure serves as a comparative reference for the nonlinear representation obtained with the autoencoder.	Proportion, 0–1

9.1 Temporal Framework of Outcome Measures

The clinical and psychometric variables used to generate the outcome measures were obtained during the **single cross-sectional assessment session**. Measures derived from the autoencoder, GMM, reconstruction analysis, and validation procedures were calculated subsequently from the

corresponding analytical dataset, according to the procedures defined in the SAP.

9.2 Nature of the Outcome Measures

The outcome measures describe the performance and representations obtained through the analytical procedures specified in the study. Identification of clusters using GMM is exploratory and is used to characterize the structure of the latent space. None of these measures was used as a criterion to modify treatment, assign interventions, or make automated clinical decisions.

10 Data Management and Quality Control

- **Pseudonymization:** the participant's name is not included in the analytical dataset. The clinical registry retains direct identifiers under the custody of the study center; the analytical dataset and any shared datasets use a case identifier derived through a salted hash function.
- **Provenance verification:** each link in the data chain, from the clinical registry to the analytical matrix and subsequently to the model artifacts, is documented using a **SHA-256 fingerprint** and verified by recomputation. The hash inventory accompanies this protocol as supporting documentation (see section 20).
- **Data cleaning:** the internal data and code audit was conducted on the complete registry. **All identified findings were remediated**; the analytical dataset documented in this protocol corresponds to the cleaned and verified version.
- **Reproducibility:** the variable dictionary, scoring specification, and verification scripts are published with the study documentation, allowing an independent reviewer to reproduce the preprocessing procedures and verify the reported scores, subject to the applicable data-access and confidentiality restrictions.
- **Backup:** encrypted backup copies are maintained at the study center. Shared datasets are limited to the materials and scope specified in section 16 and the applicable data-sharing plan.

11 Statistical Analysis

The statistical analysis is specified in detail in the **Statistical Analysis Plan (SAP), version 1.0**, which forms part of the supporting documentation for the registry and is aligned with the audited analysis code. The present protocol summarizes the principal procedures used for data preparation, model training and evaluation, cluster identification, and interpretation of the resulting representations.

11.1 Data Preparation

The variables included in the analysis underwent a standardized data preparation procedure. Missing values were imputed using the median of the corresponding variable. To reduce redundancy among predictors, pairwise correlations between variables were evaluated and pairs with a correlation greater than **0.85** were removed, in accordance with the criteria established in the SAP. The retained variables were standardized before being entered into the models.

11.2 Autoencoder Representation

A **symmetric autoencoder** with a **12-dimensional latent bottleneck layer** was trained. Training consisted of **30 independent runs** using different random seeds to evaluate variability associated with the stochastic nature of the optimization process.

The performance of each run was evaluated according to the criteria defined in the SAP. To obtain the final model representation, an **ensemble of the five best-performing runs** was constructed. The five runs were selected according to their validation R^2 values and combined using a weighted average. Stability of the procedure was examined using reconstruction metrics and 5-fold cross-validation repeated 3 times.

11.3 Comparison with Principal Component Analysis

As a comparative reference, **principal component analysis (PCA)** was performed on the same standardized variables used for the autoencoder. The first **12 principal components** were considered so that the dimensionality of the linear representation was comparable with the 12 dimensions obtained from the autoencoder. The comparison was based on the proportion of variance explained and the performance metrics defined in the SAP.

11.4 Clustering Using Gaussian Mixture Models

The structure of the latent space was analyzed using **Gaussian mixture models (GMM)** with a full covariance matrix. Solutions with $k = 1$ **to** 8 clusters were evaluated. Selection of the final solution considered the **Bayesian information criterion (BIC)**, silhouette measure, and solution stability assessed through **bootstrap** procedures.

Cluster membership was interpreted as a probabilistic classification within the latent space. Given the exploratory nature of this procedure, the identified clusters were not automatically considered discrete clinical subtypes.

11.5 Interpretation of Latent Dimensions

The latent dimensions were interpreted by examining their correlations with the original clinical and psychometric variables. Based on these patterns, a **structured clinical labeling** procedure was applied according to previously defined criteria. Where applicable, the consistency of the interpretive assignments was evaluated through **inter-rater agreement**.

11.6 Inferential Nature of the Analysis

The purpose of the analysis was to characterize the multivariate structure of the clinical variables and to explore latent representations and clustering patterns within the study population. The resulting latent dimensions and clusters were not used as substitutes for clinical diagnosis or as automated therapeutic decision-making tools. The estimates and structures identified should be interpreted in the context of the characteristics of the study sample and the cross-sectional design.

12 Ethical Considerations

The study was conducted according to principles of respect for the dignity, autonomy, privacy, and well-being of participants. Ethical procedures were established in accordance with the **Declaration of Helsinki**, Mexican health legislation applicable to research involving human

participants, and the state regulations applicable to the location where the study was conducted. The **Regulations of the General Health Law on Health Research** establish, among other principles, the protection of participants' rights and well-being, the requirement for informed consent, and the protection of privacy. The **NOM-012-SSA3-2012** establishes criteria for the conduct of health research projects involving human participants.

12.1 Ethical and Regulatory Principles

1. **Declaration of Helsinki.** The study was conducted in accordance with the ethical principles applicable to research involving human participants established by the World Medical Association. The Declaration of Helsinki emphasizes the protection of participants' dignity, rights, safety, and well-being, as well as the importance of independent ethical review and transparent reporting of research results.
2. **Federal health framework.** The research was conducted in accordance with the **General Health Law**, the **Regulations of the General Health Law on Health Research**, and **NOM-012-SSA3-2012**, as applicable to the observational and cross-sectional design of the study. The Regulations require that respect for human dignity and protection of participants' rights and well-being prevail in research involving human participants and provide for ethical and scientific review of research protocols.
3. **Jalisco health framework.** The study was conducted in the State of Jalisco and complied with the applicable state health legislation. The **Health Law of the State of Jalisco** includes health research among activities directed toward the study of biological and psychological processes, prevention and control of priority health problems, and development of methods applicable to health care. In addition, the **Mental Health and Addictions Health Law of the State of Jalisco** establishes provisions applicable to mental health and addiction services in both public and private sectors, including protection of dignity, confidentiality, and personal data of service users.

12.2 Institutional Ethics Review

The study protocol underwent the applicable institutional ethical review process, consistent with the information declared in the public registry under the *Oversight* section. The institutional committee, approval number, and approval date are identified in the administrative study documentation and are reproduced in the corresponding appendix (see section 20).

Institutional ethics information
Institutional Ethics Committee: Under The Tree Ethics Committee
Approval number: UTT-CEI-2024_007-ML
Approval date: 14 March 2024

The information above should remain consistent with the institutional ethics documentation and with the information reported in the public study registry.

12.3 Informed Consent

Written informed consent was obtained before any study-related procedures were performed. Participants were provided with sufficient information regarding the nature and objectives of the study, assessment procedures, foreseeable discomforts or risks, potential benefits, the voluntary nature of participation, confidentiality of information, and the right to withdraw from participation in accordance with the conditions established for the study.

The informed consent process was conducted in accordance with the principles established in Articles 20–22 of the **Regulations of the General Health Law on Health Research**, which specify the essential elements of informed consent and its written documentation.

12.4 Privacy and Confidentiality

1. **Confidentiality and privacy.** Research data were managed using pseudonymized identifiers, and the analytical dataset did not directly contain personal identifiers. Access to identifiable information was restricted to authorized personnel according to their institutional responsibilities.
   Privacy protection was implemented in accordance with the **Regulations of the General Health Law on Health Research**, which require protection of participant privacy in research involving human participants and limit identification to circumstances in which it is necessary and authorized.
2. **Protection of personal data.** The clinical and psychometric information used in the study includes health-related data and therefore requires enhanced protection. Where the data controller is a private individual or legal entity, processing is subject to the **Federal Law on Protection of Personal Data Held by Private Parties**, including its provisions concerning sensitive personal data, security measures, information duties, and the exercise of applicable data-subject rights.
   Because the study spans a period during which the federal framework governing personal data protection changed, the handling of information was conducted in accordance with the legislation applicable during each relevant stage of the study and subsequent data management.
3. **Privacy notice and purpose limitation.** Data processing was limited to the research purposes authorized and documented for the study. Information used for analysis, publication, and reproducibility procedures was subjected to de-identification or dissociation measures consistent with applicable privacy-protection requirements.

12.5 Risks and Protection Measures

1. **Foreseeable risks.** The assessment procedures were non-invasive and did not involve an experimental therapeutic intervention. The principal foreseeable risks included mild emotional discomfort, tiredness, or fatigue associated with responding to psychometric instruments, as well as risks related to the privacy and confidentiality of participant information.
2. **Support for discomfort during assessment.** When a participant experienced emotional discomfort or fatigue during the assessment, psychological support was available from personnel at the study center, consistent with the available clinical resources. Participation could be interrupted according to the participant's wishes, without consequences for the clinical care received.
3. **Separation between research and clinical care.** Participation in the study did not condition access to residential treatment and did not modify the treatment provided by the center. Research procedures remained distinct from routine clinical care and rehabilitation activities.

12.6 Use of Results and Machine Learning Model

1. **No clinical intervention derived from the model.** Outputs of the machine learning model were used exclusively for research purposes. No score, latent representation, classifi-

cation, or cluster generated by the model was used to make automated clinical decisions, modify treatment, determine eligibility for an intervention, or replace the judgment of clinical personnel.

2. **Use for knowledge generation.** Results derived from the modeling procedures were considered research findings and hypothesis-generating evidence. Identification of latent dimensions or clusters was not interpreted by itself as a diagnosis, a validated clinical subtype, or a clinical decision-making instrument.
3. **Responsible dissemination.** Presentation, publication, and potential sharing of study results are conducted in aggregated or de-identified form, avoiding disclosure of information that could directly identify participants. Technical documentation and materials intended for reproducibility are distributed according to their applicable access level and restrictions arising from personal data protection requirements.

13 Risks and Benefits

The study procedures consisted of a cross-sectional psychometric assessment and subsequent processing of the data for research purposes. No medications were administered, no invasive procedures were performed, and outputs from the machine learning model were not used to modify treatment or make automated clinical decisions.

Table 6: Risks, mitigation measures, and potential benefits.

Category	Description
Risks	Foreseeable risks were primarily related to participation in the psychometric assessment. Some questionnaire items could cause **mild emotional discomfort, uneasiness, tiredness, or fatigue** when addressing topics related to substance use, emotion regulation, anger, anxiety, borderline symptoms, or other psychological experiences. There was also a potential risk related to the **privacy and confidentiality of clinical and psychometric information**. This risk was mitigated through pseudonymization, access controls, and restricted handling of identifiable information. No invasive physical procedures were performed and no experimental treatment was administered.
Mitigation measures	Participation was voluntary and occurred after informed consent had been obtained. Participants could interrupt or discontinue the assessment in accordance with the conditions established for the study. Psychological support was available from center personnel in the event of emotional discomfort or fatigue during the assessment. Data used for analysis were managed using pseudonymized identifiers and measures designed to preserve confidentiality.
Direct benefits	**No direct benefits to participants were expected as a result of participation in the study.** The research assessment did not constitute a therapeutic intervention, and model outputs were not used to modify individual treatment.

Continued on next page

Category	Description
	Continued from previous page
Indirect benefits	The study may contribute to knowledge regarding the multivariate structure of clinical and psychometric variables in individuals with SUD and to the development of machine learning methods for functional characterization. In the longer term, this knowledge may contribute to the development and evaluation of treatment-personalization strategies, although such potential benefits are not guaranteed for individual participants.
Risk-benefit relationship	The study was designed as observational and cross-sectional research, without experimental therapeutic intervention or invasive procedures. Foreseeable risks were primarily limited to emotional discomfort or fatigue during assessment and to protection of data confidentiality. Formal assessment of the risk-benefit relationship and, where applicable, classification of the study's risk level are matters for the competent ethics review body and its corresponding determination.

14 Data Monitoring and Follow-up

Because of the observational and cross-sectional nature of the study, monitoring procedures focused primarily on the **quality, integrity, traceability, and protection of data**, as well as the identification and documentation of any unanticipated event related to the research procedures. No independent data monitoring committee was established and no treatment-related stopping decisions based on efficacy or safety were required, because the study did not include an experimental therapeutic intervention.

14.1 Data Monitoring

1. **Data Monitoring Committee.** No independent Data Monitoring Committee was established for this study. Study oversight was conducted through the institutional and internal quality-control procedures defined for the research, consistent with its observational and cross-sectional design.
2. **Responsibility of the Principal Investigator.** The Principal Investigator was responsible for supervising data quality, implementation of the established procedures, and consistency between source records and the dataset used for analysis.

14.2 Data Verification and Quality Control

1. **Internal quality control.** Checks of completeness, response ranges, variable consistency, and correspondence between source records and the analytical dataset were performed. These activities were intended to identify capture errors, out-of-range values, discrepancies, and potential unauthorized modifications of data.
2. **Provenance audit.** Variable traceability was verified from the initial record through incorporation into the analytical dataset. Discrepancies identified during review were documented and addressed according to the study's established quality-control procedures.

3. **Integrity verification using hashes.** The integrity of registry files and the analytical dataset was verified using hash functions, allowing correspondence between versions to be confirmed and subsequent file modifications to be detected. Specific integrity and audit procedures are documented in the methodological and technical study documentation (see section 10).

14.3 Participant Follow-up

1. **Post-assessment follow-up.** The protocol included a **single cross-sectional assessment session**. No longitudinal clinical follow-up visits or additional assessments related to the machine learning model were scheduled. Therefore, participant follow-up was limited to the period corresponding to the assessment and the routine care provided by the center.
2. **Continuity of clinical care.** Participation in the study did not replace or modify routine residential treatment. Research procedures remained distinct from ordinary clinical decisions and activities at the center.

14.4 Adverse Events and Unanticipated Events

1. **Adverse event reporting.** Adverse events, discomfort, or unanticipated effects occurring during participation and potentially related to research procedures were recorded, particularly those associated with questionnaire administration.
2. **Expected nature of events.** Given the non-invasive nature of the study procedures, foreseeable events were primarily related to emotional discomfort, uneasiness, or fatigue during the assessment. Any event requiring additional attention was managed in accordance with the clinical resources available at the center.
3. **Documentation and assessment.** Events and unanticipated occurrences were documented according to the applicable institutional procedures, including, when relevant, their nature, time of onset, relationship to the research procedure, and measures taken. Detailed information concerning recorded events and their resolution is provided in the *Clinical Study Report*, §8.
4. **Participant protection.** Participant identification in event documentation and follow-up records was performed using the study identifier, avoiding unnecessary use of direct personal identifiers. This approach is consistent with applicable Mexican provisions concerning privacy protection and documentation of research involving human participants.

15 Protocol Amendments

No substantive protocol modifications were identified that changed the primary objectives, the observational and cross-sectional design, eligibility criteria, assessment procedures, or outcome measures of the primary analysis.

During preparation and consolidation of the study documentation, **editorial and methodological wording revisions** were made to clarify operational definitions, improve procedural traceability, and strengthen the description of analytical and quality-control methods. These revisions did not retrospectively modify the procedures applied to participants or the criteria used for the primary analysis.

Subsequently, the study database continued to be updated through the incorporation and documentation of additional participants. The subsequent publication using **166 participants** corresponds to an expanded dataset relative to the methodological cutoff of **155 participants** established for the primary analysis described in this document. This subsequent update is

documented as an extension of the dataset for subsequent analyses and does not modify the participant flow or sample size of the primary analysis presented in this Study Document.

15.1 Modification History

1. **Document revision and methodological strengthening.** Revisions were made to wording, terminology, and procedural documentation without substantive modification of the study design, eligibility criteria, assessment procedures, or outcome measures of the primary analysis.
2. **Subsequent database update.** Additional participants were incorporated for subsequent analyses, resulting in an expanded dataset of 166 participants used in a subsequent publication. The primary dataset documented in this Study Document remains composed of 155 participants.

16 Publication, Data Sharing, and Intellectual Property

Study findings are disseminated through academic publications, scientific presentations, and data-sharing mechanisms that are compatible with participant protection and the access conditions established for the research data.

16.1 Publication

1. **Scientific publication.** The results of the analysis were published in the following peer-reviewed book chapter:

   > Gutiérrez Castro, L., Mares Ramírez, D. L., Romero López, V., & Calixto Botello, F. J. (2026). *Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders*. In M. G. Orozco-del-Castillo, J. A. Recio-Garcia, & S. Kapetanakis (Eds.), *Artificial Intelligence for Mental Health. ICAIMH 2026. Communications in Computer and Information Science*, Vol. 3062, pp. 168–180. Springer, Cham.

2. **Scientific dissemination.** Study findings may also be disseminated through conference presentations, academic meetings, scientific reports, and other forms of scientific communication consistent with confidentiality obligations and applicable data protection requirements.

16.2 Data Sharing

1. **Availability of individual-level data.** Following publication, access to de-identified individual-level data from the primary dataset of **155 participants** may be considered, provided that disclosure is compatible with informed consent, applicable ethical and legal obligations, and the study's data protection requirements.
2. **Data potentially available upon request.** The dataset potentially available for sharing may include the **21 standardized clinical variables** used in the analysis, provided that the shared form does not permit participant identification. Individual 12-dimensional latent representations are not automatically considered part of the publicly distributed materials. Where applicable, access to these representations may be subject to specific review of the request and to additional data-protection and scientific-governance restrictions.

3. **Methodological materials.** The study protocol, Statistical Analysis Plan, and analysis code may be made available to qualified researchers to facilitate methodological assessment, replication of the analytical procedures, and secondary analyses, subject to the applicable access conditions.

16.3 Access Criteria and Procedure

1. **Access request.** Requests for access to individual-level data should be directed to the corresponding author or the person designated to manage the study data. Requests should describe the scientific purpose, variables requested, methodological plan, and proposed measures for protecting the information.
2. **Request evaluation.** Access may be granted when the proposal provides an appropriate scientific and methodological justification and is compatible with applicable ethical, legal, and confidentiality restrictions. Authorization does not constitute an automatic right of access and may require execution of a specific data-use agreement.
3. **Access period.** Availability for individual-level data requests will follow the study's data-sharing policy, beginning after publication and subject to data-governance requirements, privacy protections, and material availability.

16.4 Intellectual Property and Restricted Materials

1. **Intellectual property.** Trained model weights, unpublished internal parameters, production code, and clinical applications developed from the study findings remain subject to the intellectual property, licensing, and governance conditions established by their respective rights holders.
2. **Separation between reproducibility and artifact distribution.** Availability of study results, methodological documentation, and analysis code does not imply unrestricted distribution of all internal study artifacts. Restricted materials remain subject to the applicable conditions governing access, data protection, confidentiality, and intellectual property.

17 Timeline

The study was conducted sequentially from the beginning of recruitment through completion of study activities and subsequent dissemination of the results. The following dates correspond to the principal operational and reporting milestones of the study.

Table 7: Study timeline.

Milestone	Date
Start of recruitment	February 2024
Recruitment period	February 2024 – December 2025
Primary completion date (actual)	18 February 2026
Study completion date (actual)	22 April 2026
Publication of results	29 June 2026
Submission of results to the public registry	**25 September 2026 (planned)**
Opening of the data-sharing period	**29 March 2027**, corresponding to 9 months after publication

The publication date corresponds to the results chapter published in *Communications in Computer and Information Science* (CCIS 3062). The chapter was published online on 29 June 2026.

Submission of results to ClinicalTrials.gov is planned for **25 September 2026**. This date corresponds to the planned registry update and falls before the general one-year deadline calculated from the primary completion date, which for this study falls in February 2027. When applicable, ClinicalTrials.gov provides for a general results submission period of one year following the primary completion date.

The period for sharing individual-level data begins **9 months after publication**, subject to the ethical, legal, confidentiality, and data-governance conditions described in section 16.

18 Anticipated Methodological Limitations

The following limitations should be considered when interpreting the study results. Some arise directly from the study design and sample characteristics, whereas others are related to the exploratory nature of unsupervised learning and to the clinical interpretation of the resulting representations. These limitations restrict the scope of the conclusions and support the need for replication and external validation.

1. **Moderate sample size and single-center setting.** The analysis included a sample of $n = 155$ participants from a single residential treatment center. Consecutive sampling allows characterization of the clinically accessible population during the study period; however, its non-probability nature and the single institutional setting limit generalizability to other populations, institutions, treatment contexts, or clinical profiles.
2. **Cross-sectional design.** Clinical and psychometric variables were assessed at a single time point. Therefore, the associations identified among variables, latent dimensions, and clusters do not establish causal relationships or determine the temporal direction of the observed phenomena. The design also does not directly permit assessment of the longitudinal stability of the latent representations or their ability to anticipate clinical changes or relapse.
3. **Absence of an independent external test set.** Although internal validation and stability assessment procedures were used, the study did not include an independent external cohort to evaluate model performance outside the population used for model development. Consequently, the observed performance should not be interpreted as sufficient evidence of generalization to new samples or institutions.
4. **Dependence on preprocessing and representation decisions.** Variable selection, treatment of missing values, standardization, and redundancy reduction using the specified correlation criterion are part of the analytical specification and may influence the structure learned by the model. Accordingly, the latent representation should be interpreted as dependent on the methodological decisions adopted for this dataset and not as a unique or universal representation of the psychopathology associated with SUD.
5. **Stochastic nature of autoencoder training.** Neural network training may produce different solutions across runs due to initialization and other stochastic components of the optimization process. The use of multiple runs, cross-validation, and an ensemble reduces this source of variability but does not completely eliminate the possibility of variation across training runs or datasets.
6. **Interpretation of latent dimensions.** The dimensions obtained by the autoencoder are mathematical representations derived from multivariate patterns and do not necessarily

correspond to directly observable clinical constructs. Their interpretation through correlations with the original variables and structured clinical labeling introduces an interpretive component. Although the labeling was performed by **three clinicians with documented agreement**, this interpretation should not be regarded as an independent validation of new psychopathological constructs.

7. **Limited separation between clusters.** GMM solutions may exhibit overlap between groups and membership probabilities that do not correspond to clinically discrete categories. Accordingly, cluster identification should be interpreted as a probabilistic and heuristic characterization of the latent space rather than as evidence of established diagnostic subtypes.
8. **Dependence on clustering-selection criteria.** Selection of the GMM solution depends on statistical and interpretive criteria, including BIC, silhouette, and bootstrap stability. These criteria may yield different solutions when the data, variables, or modeling conditions change. Therefore, the number and configuration of clusters should be considered specific to the analyzed dataset.
9. **Limited comparison with linear methods.** Comparison with PCA provides a reference for evaluating the 12-dimensional representation but does not constitute an exhaustive comparison with all alternative dimensionality-reduction or unsupervised learning methods. Relative performance between the autoencoder and PCA should therefore be interpreted within the methodological specifications used in this study.
10. **Absence of clinical utility validation.** The study evaluates the ability of the procedure to represent and characterize the multivariate structure of the data but does not, by itself, demonstrate clinical utility, therapeutic benefit, or improvement in patient outcomes. The latent dimensions and clusters require further studies evaluating their stability, clinical relevance, and prospective utility.
11. **Potential influence of the care setting and study population.** Clinical and psychometric characteristics of a residential SUD population may differ from those observed in other levels of care or among individuals with SUD who are not receiving residential treatment. The identified latent structure may therefore be influenced in part by the characteristics of the care setting and the population available for the study.
12. **Exploratory and hypothesis-generating nature.** Overall, the findings should be considered **hypothesis-generating**. Identification of latent dimensions and clusters provides exploratory evidence regarding the structure of the data studied but does not, by itself, establish new clinical subtypes, causal relationships, or predictive utility. Confirmation of these findings requires independent replication, external validation, and, where appropriate, longitudinal and prospective evaluation.

19 References

[**0**] Gutiérrez Castro, L.; Mares Ramírez, D. L.; Romero López, V.; Calixto Botello, F. J. *Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders.* In: *Artificial Intelligence for Mental Health* (ICAIMH 2026), *Communications in Computer and Information Science*, Vol. 3062, pp. 168–180. Springer. **DOI:** 10.1007/978-3-032-30396-7_12.

[**1**] Stellern, J.; Xiao, K. B.; Grennell, E.; Sanches, M.; Gowin, J. L.; Sloan, M. E. Emotion regulation in substance use disorders: a systematic review and meta-analysis. *Addiction* **118**(1), 30–47 (2023). **DOI:** 10.1111/add.16001.

[2] Kozak, K.; Lucatch, A. M.; Lowe, D. J.; Balodis, I. M.; MacKillop, J.; George, T. P. The neurobiology of impulsivity and substance use disorders: Implications for treatment. *Annals of the New York Academy of Sciences* **1451**(1), 71–91 (2019). **DOI:** 10.1111/nyas.13977.

[3] National Institute of Mental Health. RDoC Domains and Constructs. U.S. Department of Health and Human Services, National Institutes of Health. **URL:** https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs.

[4] Iyortsuun, N. K.; Kim, S. H.; Jhon, M.; Yang, H. J.; Pant, S. A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare* **11**(3), 285 (2023). **DOI:** 10.3390/healthcare11030285.

[5] Hinton, G. E.; Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006). **DOI:** 10.1126/science.1127647.

[6] Bzdok, D.; Meyer-Lindenberg, A. Machine learning for precision psychiatry: opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **3**(3), 223–230 (2018). **DOI:** 10.1016/j.bpsc.2017.11.007.

[7] Mak, K. K.; Lee, K.; Park, C. Applications of machine learning in addiction studies: a systematic review. *Psychiatry Research* **275**, 53–60 (2019). **DOI:** 10.1016/j.psychres.2019.03.001.

[8] Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J. T. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics* **19**(6), 1236–1246 (2018). **DOI:** 10.1093/bib/bbx044.

[9] Fraley, C.; Raftery, A. E. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* **97**(458), 611–631 (2002). **DOI:** 10.1198/016214502760047131.

[10] Kashihara, J.; Takebayashi, Y.; Kunisato, Y.; Ito, M. Classifying patients with depressive and anxiety disorders according to symptom network structures: a Gaussian graphical mixture model-based clustering. *PLoS ONE* **16**(9), e0256902 (2021). **DOI:** 10.1371/journal.pone.0256902.

[11] Marquand, A. F.; Wolfers, T.; Mennes, M.; Buitelaar, J.; Beckmann, C. F. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **1**(5), 433–447 (2016). **DOI:** 10.1016/j.bpsc.2016.04.002.

[12] Patton, J. H.; Stanford, M. S.; Barratt, E. S. Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology* **51**(6), 768–774 (1995). **DOI:** 10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1.

[13] Spielberger, C. D. *State-Trait Anger Expression Inventory-2: STAXI-2*. Psychological Assessment Resources, Odessa, FL (1999).

[14] Spitzer, R. L.; Kroenke, K.; Williams, J. B.; Löwe, B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine* **166**(10), 1092–1097 (2006). **DOI:** 10.1001/archinte.166.10.1092.

[15] Kanter, J. W.; Mulick, P. S.; Busch, A. M.; Berlin, K. S.; Martell, C. R. The Behavioral Activation for Depression Scale (BADS): psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment* **29**(3), 191–202 (2007). **DOI:** 10.1007/s10862-006-9038-5.

[16] Armento, M. E.; Hopko, D. R. The Environmental Reward Observation Scale (EROS): development, validity, and reliability. *Behavior Therapy* **38**(2), 107–119 (2007). **DOI:** 10.1016/j.beth.2006.05.003.

[17] Bohus, M.; et al. The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. *Psychopathology* **42**(1), 32–39 (2009). **DOI:** 10.1159/000173701.

[18] Abadi, M.; et al. TensorFlow: a system for large-scale machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283 (2016). **DOI:** 10.48550/arXiv.1603.04467.

[19] Klambauer, G.; Unterthiner, T.; Mayr, A.; Hochreiter, S. Self-normalizing neural networks. In: *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, pp. 971–980 (2017). **DOI:** 10.48550/arXiv.1706.02515.

[20] Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* **15**(1), 1929–1958 (2014). **URL:** https://jmlr.org/papers/v15/srivastava14a.html.

[21] Kingma, D. P.; Ba, J. Adam: a method for stochastic optimization. In: *3rd International Conference on Learning Representations (ICLR 2015)* (2015). **URL:** https://arxiv.org/abs/1412.6980.

20 Appendices

20.1 Appendix A — Supporting Study Documentation

The study's supplementary documentation includes the following materials:

1. **Statistical Analysis Plan (SAP), version 1.0.** Document specifying the statistical and machine learning procedures used for data analysis.
2. **Informed Consent Form (ICF).** Institutional form used to document informed consent from study participants.
3. **Clinical Study Report (CSR), version 1.0.** Consolidated report of study conduct, participant disposition, results, and associated study documentation.
4. **Supplementary methodological documentation.** Includes detailed instrument scoring specifications, the variable definitions, methodological verification procedures, and documentation of data provenance and integrity. These materials constitute complementary technical documentation, and their availability is subject to the access, data-protection, and governance conditions described in section 16.

20.2 Appendix B — Publication Reference

The principal results of the study were published as:

Gutiérrez Castro, L., Mares Ramírez, D. L., Romero López, V., & Calixto Botello, F. J. (2026). *Unsupervised Deep Representation Learning for Clinical Stratification in Substance Use Disorders*. In *Artificial Intelligence for Mental Health. ICAIMH 2026. Communications in Computer and Information Science*, Vol. 3062. Springer.

20.3 Appendix C — Administrative Information

Information concerning ethical approval, study responsibilities, document versions, and approval dates is maintained in the corresponding institutional study records and should remain consistent with the information reported in the public study registry.

Institutional Ethics Committee: Under The Tree Ethics Committee
Approval number: UTT-CEI-2024_007-ML
Approval date: 14 March 2024