Cross-Scale Modeling of Healthcare Norms and Patient Features Dynamics with Interpretable Machine Learning

Chao Li[1][0009−0000−1233−3281], Dutao Zhang[1], Fei Ren[2][0000−0001−7308−686X], and Sergey Kovalchuk[1][0000−0001−8828−4615]

[1] ITMO University, Saint Petersburg, Russia
316325@niuitmo.ru, kovalchuk@itmo.ru
[2] Independent researcher, China

Abstract. This study proposes an interpretable machine learning framework to model bidirectional dynamic interactions between macroscopic norms and microscopic features in clinical data. Leveraging real-world medical records from a specialized chest hospital (containing unstructured text, complex categorical variables, temporal indicators, and non-random missing patterns), we perform numerical processing through Latent Semantic Analysis and dimensionality reduction via Non-negative Matrix Factorization. Macroscopic therapeutic norms are identified using HDBSCAN clustering, while SHAP-XGBoost integration selects critical microscopic features, including multidrug-resistant tuberculosis diagnosis and liver function biomarkers. We integrate symbolic regression with the Peter-Clark Momentary Conditional Independencecausal discovery method based on partial correlation, constructing cross-scale functional relationships with temporally rigorous constraints. Specifically, PySR derives nonlinear mapping equations, while partial correlation-based conditional independence tests establish time-lagged dynamic dependency networks. Guided by the Dynamic Maximum Entropy across Scales (DyMES) principle, multi-scale perturbation experiments reveal bidirectional mechanisms. Within our dataset and framework, DyMES reveals dynamic constraints' interplay driving statistical equilibrium between macroscopic clinical norms and microscopic patient characteristics through nonlinear coordination and threshold-triggered time-encoded mechanisms. Persistent constraint interactions induce novel steady states formation with dynamically preserved system memory.

Keywords: Medical Norms · Interpretable Machine Learning · Dynamic Maxent across Entwined Scales · Symbolic Regression · Distributed healthcare

1 Introduction

Medical norms, predominantly designed for human practitioners, are encoded in unstructured natural language with implicit references to clinical expertise,

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

posing significant challenges for autonomous agents to dynamically adapt to evolving norms (e.g., treatment guidelines for drug-resistant tuberculosis) in distributed healthcare systems. Enabling autonomous agent systems to learn and perceive the norms among healthcare professionals is both intriguing and essential for their integration into real-world distributed healthcare environments [7, 13]. The heterogeneous phenomena of medical norm propagation and adoption within autonomous multi-agent systems are intrinsically linked to the emergence of collective behaviors and the formation of organizational structures. These processes constitute fundamental manifestations of organized complexity [1, 16]. One of the key challenges in current complex system modeling lies in the weak interpretability of emergent behaviors and insufficient formalized descriptions. Integrating machine learning into agent-based modeling is a highly promising research direction. Feature-based explanation approaches provide novel perspectives for understanding the complex emergent behaviors of multi-agent systems and linking micro- and macro-level characteristics [11, 8].

Current research on norm propagation in medical autonomous multi-agent systems has achieved progress in formal modeling and validation with real-world clinical datasets [9]. However, two critical limitations persist. First, the absence of formalized cross-scale dynamic coupling mechanisms, particularly the lack of quantitative methodologies for macro-level norm and micro-level patient feature co-evolution. Second, existing models fail to effectively characterize the association between norm dynamic adaptation processes and system multi-scale characteristics [3, 15]. For instance, the inability to quantify how microscopic feature variations (e.g., liver function test results) trigger macroscopic norm adjustments (e.g., medication plan revisions), and how such macroscopic adjustments subsequently influence treatment cycle timelines across diverse patients. This constitutes a core manifestation of organized complexity in healthcare systems. The central challenge lies in constructing a formal framework for dynamic constraints between microscopic individual features and macroscopic therapeutic norms.

Building upon the two limitations explored above, we extended the Dynamic Maxent across Entwined Scales (DyMES) theory [4] in complex systems to study medical norms in autonomous multi-agent systems. DyMES is a dynamic theory combining Top–Down information-theoretic inference with Bottom–Up state-variable-dependent mechanisms. In this framework, state variables influence microscale dynamics while being computed as averages over probability distributions of microvariables. This integration enables simultaneous prediction of time-evolving state variables and microvariable distributions. Central to DyMES is the notion of transition functions, which govern microvariable dynamics. Scale entwinement, and in particular, downward causation, is captured by explicit dependence of transition functions on state variables as well as on microvariables [4].

To address these gaps, we propose a three-stage computational framework synergizing interpretable machine learning with DyMES theory. First, XAI techniques disentangle macro-micro correlations from noisy EHR data. Second, symbolic regression distills these associations into cross-scale transition functions.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

Third, we integrate PCMCI-based causal discovery with partial correlation conditional independence tests to introduce temporally modulated functions for the derived transitions, optimizing their parameters through grid search based on these metrics. Finally, we construct a DyMES model with the optimized dynamic transition functions, conducting multiscale perturbation experiments on strictly monotonic temporal sequences to simulate bidirectional macro-norm-micro-feature interactions.

This study makes three core contributions: first, it establishes a DyMES framework that integrates maximum entropy principles with interpretable machine learning to formalize bidirectional cross-scale interactions; second, it develops a unified methodology combining symbolic regression-derived equations with PCMCI-validated bidirectional feature causality; finally, it introduces the first computational dynamical model for co-evolution between institutional medical norms and personalized patient characteristics.

2 Cross-Scale Dynamic Modeling Framework

We propose a general extensible framework comprising two categories of components. The first category corresponds to the colored sections in Figure 1, specifically a Two-Stage Computational Architecture responsible for all XAI-related operations. The second category (white sections in Figure 1) extends the DyMES model to healthcare datasets through simulation components, aiming to rigorously interpret macro-micro correlations within the dataset.

2.1 Two-Stage Computational Architecture

Fisrt stage. The initial stage can be abstracted as: decoupling macro-micro correlations in datasets through diverse XAI tools based on their inherent characteristics and structural composition, where machine learning methods and feature processing approaches are selectively employed according to data properties and sparsity levels.

The fundamental principle of this first stage involves identifying crucial features from noisy data, uncovering strong dependencies between significant feature vectors, and subsequently distinguishing macro/micro features through integration with domain expertise and clinical knowledge.

For processing hybrid medical datasets containing clinical norms, the primary methodology involves unifying heterogeneous features into computable encodings, removing overly sparse and insignificant feature columns, followed by reasonable dimensionality reduction methods to prevent matrix oversizing. Our observations indicate these datasets typically exhibit semi-structured formats — organized as structured tables with explicit headers corresponding to the categories documented in the 'Main Module' column of Table 1 (e.g., Basic Information, Clinical Process, Diagnostic Testing Modules). In this architecture, each patient sample comprehensively populates all categories defined under the 'Main Module' column of Table 1, forming a complete longitudinal record.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

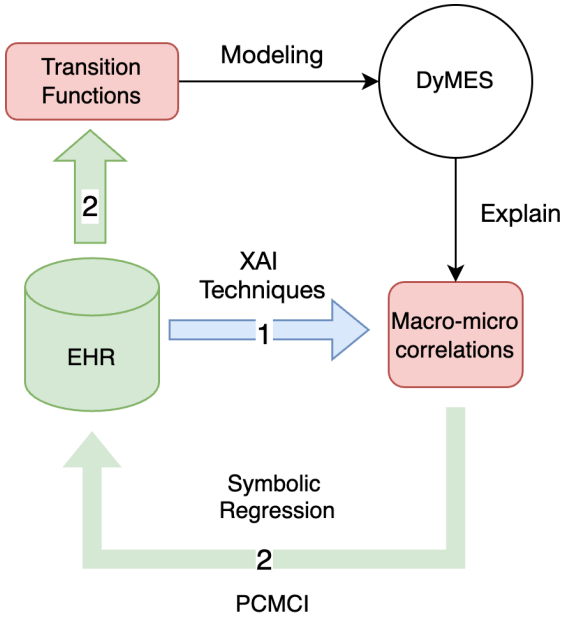

Fig. 1. Dynamic cross-scale norm interactions

The identification of critical feature vectors within the processed feature matrix is primarily achieved through interpretable clustering approaches. Recommended clustering methods for handling hybrid feature matrices typically include HDBSCAN, Fuzzy C-Means Clustering, and Autoencoder KMeans. Subsequently, feature importance analysis is conducted based on the clustering results. Recommended static analysis methods applicable here include cluster persistence scores, density analysis of clusters, cluster membership probabilities, and cluster hierarchy trees. A more efficient dynamic approach involves training supervised classification models using cluster labels, followed by analyzing feature contributions to cluster label prediction through SHAP and LIME techniques. Based on the hybrid content formats of datasets, methods including association rule mining (Apriori/FP-Growth), dependency and correlation analysis, and rule induction (decision trees/RIPPER) can also serve as recommended alternative approaches for analyzing clustering results.

For encoded matrices derived from hybrid medical datasets containing clinical norms, analysis of clustering results typically yields a finite set of correlation combinations. Ultimately, through integration with medical domain knowledge and clinical expert validation, we can identify clinically significant macro-micro correlations within these combinatorial patterns.

Second stage. Macro-micro correlations based on feature importance cannot be directly transformed into computational models. Therefore, the second-stage work involves mathematical formula mining through actual data of differ-

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

ent feature columns corresponding to macro-micro correlations in the dataset. The core methodology here combines symbolic regression and PCMCI. The essence of symbolic regression lies in exploring vast function spaces using evolutionary algorithms or heuristic searches to identify mathematical expressions that optimally fit the data [17]. Symbolic regression achieves balance between "unknown candidate function forms in the library" and "required structural constraints for continuous dynamical systems." In multivariate, multiscale time series, causal structures often exhibit complexity. PCMCI (Peter-Clark Momentary Conditional Independence) is a statistical method that discovers statistically significant cross-scale causal mechanisms in medical temporal data through partial correlation conditional independence tests [14].

By integrating the two methodologies from Stage 2 — performing function form searching across datasets and conducting multi-parameter optimization of identified functions on test sets — we derive the formalized functional dependency between micro and macro features as expressed in equation 1. To streamline exposition, the complete mathematical definition of equation 1 is methodologically consolidated with the implementation framework for extending the DyMES to medical dataset simulations, thereby establishing an integrated analytical paradigm.

2.2 DyMES model on medical datasets

We present key mathematical conventions based on the core DyMES framework. Detailed theoretical derivations are provided in the work of John Harte et al [4].

First define m macroscale variables $X = (X_1, X_2, \cdots, X_m)$ with at most m corresponding microscale variables $\mathbf{x} = (x_1, x_2, \cdots, x_m)$. Here macroscale represent norms themselves while microscale variables manifest as salient features of individual patient samples. Both vector types X and $\mathbf{x}$ derive from previously mined macro-micro correlations within the dataset's feature vectors.

$R(\mathbf{x})$ represents the joint probability distribution of microscale variables. To determine $R(\mathbf{x})$, we maximize the Shannon information entropy of $R(\mathbf{x})$ under constraints imposed by X and dX/dt. We express these constraints as $F = (h_1(X), \cdots, h_m(X), \frac{dX_1}{dt}, \cdots, \frac{dX_m}{dt})$, where $h_\mu(X)$ are functions of macroscale variables. The average values of these functions over $R(\mathbf{x})$ yield the constraint conditions, denoted by $f_\mu(\mathbf{x}, X)$. For $\mu = 1, \cdots, m$, the functions f_μ depend solely on x_μ. In more complex cases when $\mu = m+1, \cdots, 2m$, f_μ may be functions of multiple microscale variables. Scale entwinement arises when f_μ serving as transfer functions can depend on both macroscale variables X and microscale variables $\mathbf{x}$ for $\mu = m+1, \cdots, 2m$. Therefore, we formulate all constraints as:

$$F_\mu = \sum_{\mathbf{x}} f_\mu(\mathbf{x}, X) R(\mathbf{x}|X) \tag{1}$$

where $\mu = 1, 2, \cdots, 2m$, the summation indicates integration over each microscale variable x_i, and explicitly denotes the conditional dependence of $R(\mathbf{x})$ on X.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

Note that the transition function f_μ is methodologically extracted from the dataset through our Stage 1 and 2 XAI techniques.
By maximizing the Shannon information entropy of R: $H = -\sum_{\mathbf{x}} R\log(R)$, we obtain [5, 6, 12]:

$$R(\mathbf{x}|X) = \frac{e^{-\sum_\mu \lambda_\mu f_\mu(\mathbf{x},X)}}{Z} \tag{2}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{2m})$ are Lagrange multipliers obtained by solving the constraint conditions [5, 2]. Z is the normalization constant ensuring total probability sums to 1.

Equations 1 and 2 establish the foundational definitions for our DyMES framework. A core assumption of DyMES theory concerns the dynamic constraint updating process. When X and $\frac{dX}{dt}$ are known at time t, the Lagrange multipliers λ can be determined through maximum entropy conditions at time t.
We omit the rigorous derivation process from John Harte et al. [4] and directly cite the key computational equations:

$$\sum_{v=1}^{2m} \mathrm{Cov}(f_{m+i}, f_v)\frac{d\lambda_v}{dt} = 0 \tag{3}$$

where index i ranges from 1 to m ,and $\mathrm{Cov}(A, B) = \langle AB\rangle - \langle A\rangle\langle B\rangle$ denotes covariance between A and B. Equation (3) provides m relationships among the $2m$ time derivatives of Lagrange multipliers.

$$\frac{dX_i}{dt} + \sum_{\mu=1}^{2m}\left(\mathrm{Cov}(f_i, f_\mu)\frac{d\lambda_\mu}{dt}\right) + \left(\mathrm{Cov}\left(f_i, \frac{df_\mu}{dt}\right)\lambda_\mu\right) = 0 \tag{4}$$

Equations (3) and (4) can be efficiently solved through matrix inversion to determine the time derivatives of Lagrange multipliers, which are then iteratively updated. Equations (2), (3), and (4) formulate the theoretical foundation of DyMES. These equations characterize the intertwined dynamic relationships between macrostate variables and Lagrange multipliers within the system [4].

3 Experiments and simulations

3.1 Dataset processing

We conducted experiments using the dataset (collected in a specialized chest hospital), comprising longitudinal clinical data from the multidrug-resistant tuberculosis (MDR-TB) diagnosis and treatment database established and maintained by our research team. All enrolled patients underwent monthly follow-up assessments in strict accordance with therapeutic protocols developed by a multidisciplinary therapy group [10].

The dataset exhibits three primary characteristics: high sparsity with non-random missingness and heterogeneous medical data types. It comprises 31 major feature categories (see Table 1 Submodule Components) containing 1,245

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

Table 1. Main features in the dataset

Main Module	Submodule Components
Basic Information Module	Patient Identification Demographic Characteristics Clinical Baseline
Clinical Process Module	Clinical Examination Records Initial Diagnosis Documentation Follow-up Information Transfer Records
Diagnostic Testing Module	Hain GenoType MTBDRplus GeneXpert MTB/RIF Mycobacterial Speciation Chest Radiography Sputum Smear Microscopy Sputum Culture Liver Function Tests Conventional Drug Susceptibility (Selected Key Features)
Assessment & Monitoring Module	Evaluation Metrics Visit Assessment Protocols Adverse Events (AE/SAE Records) Therapeutic Outcome Documentation
Treatment Management Module	Therapeutic Regimen Specifications Treatment Protocol Documentation
Research Management Module	Case Enrollment Forms Longitudinal Follow-up Records Serial Number Identification

feature columns, where 23 categories demonstrate > 0.8 sparsity. The time span is from April 11, 2018 to December 14, 2023. These 1,245 columns incorporate temporal (follow-up dates, report dates, etc.); categorical (sputum smear results, conventional drug susceptibility testing, medication regimen codes, etc.); numerical (ALT levels from hepatic panels, serum creatinine values from renal profiles, etc.); binary (sputum culture submission flags, cavitation presence in chest imaging, etc.); and natural language data types (radiographic findings descriptions, etc.).

We take the zero-missing "follow-up date" column as the temporal axis. The datetime values are normalized to [0,1] with day granularity, followed by timestamp micro-adjustments for same-day samples to ensure strictly increasing time series aligned with dataset span; For categorical and binary feature columns, missing value indicator columns are appended before one-hot encoding, with subsequent NMF dimensionality reduction applied to high-dimensional features; Numerical columns with sparsity threshold <0.8 are filtered and retained; Textual description fields undergo TF-IDF vectorization extracting unigrams and bigrams as base features, accompanied by binary indicators for text missingness. Truncated SVD (i.e., LSA) reduces TF-IDF matrix dimensionality based

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

on singular value decay curves. Semantic features are concatenated with missing indicators, forming final structured encoded features where each column has its own CSV file.

Pre/post-processing metadata including data types, encoding schemes, and notes are recorded in JSON files. All feature CSVs are merged into a 20872 (samples) × 387 (features) matrix, followed by Gower distance matrix computation.

3.2 Mining Macro-micro Correlations

For this high-dimensional dataset characterized by elevated sparsity, non-random missingness, and heterogeneous data types, we evaluated and implemented the four clustering methods detailed in Table 2. The experimental results demonstrate that HDBSCAN achieves optimal performance, attaining the highest Silhouette Score (0.5275) and lowest Davies-Bouldin Index (0.8192) among all evaluated approaches. While the hybrid Autoencoder+K-Means method exhibits potential competitiveness, its practical implementation faces challenges in architectural optimization of the deep neural network, which incurs significant engineering overhead and compromises computational efficiency.

Table 2. Clustering method performance Comparison

Clustering Method	**Silhouette Score**	**Davies-Bouldin Index**
HDBSCAN	**0.5275**	**0.8192**
Fuzzy C-Means	0.3489	0.9717
KMeans	0.3179	1.0144
Autoencoder + K-Means	0.4663	0.8754

The HDBSCAN clustering results show: Number of clusters = 55. Number of noise points = 7,980. These noise points represent specific cases that are not the current focus due to the high matrix dimensionality and sparsity. The final corrected valid samples shape is (12892, 12892). Initial static analysis reveals: the largest cluster contains 837 samples, the smallest cluster has 235 samples. The maximum persistence score is 0.6304, with 4 clusters exceeding 0.1 persistence score threshold.

To mine macro-micro correlations from the clustering model and results, we compared multiple methods.

Regarding core objectives: association rule mining primarily identifies frequent co-occurrence patterns among features (e.g., "feature A and feature B frequently co-occur"), dependency/correlation analysis focuses on detecting statistical relationships between features (using metrics like Pearson correlation coefficients and mutual information), rule induction aims to generate human-readable "if-then" rules (e.g., "age >60 AND complications > 3 → high-risk cluster"), while SHAP analysis explains model decision logic for cluster assignments. Comparative evaluation reveals rule induction and SHAP methods demonstrate superior performance in medical data analysis (see table 3).

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

Particularly, decision tree or RIPPER-based rule induction produces intuitive "if-then" rules that prove invaluable for clinical visualization and cross-domain expert collaboration. However, experiments on the considered dataset confront challenges from high-dimensional heterogeneous data (containing numerical, encoded, missing indicators, and NMF/SVD dimensionality-reduced features) - requiring noise control through feature selection, clinical binning, and pruning optimization. The feature type diversity and high dimensionality may lead to verbose rules with reduced interpretability, especially when features aren't rigorously refined. Single decision trees or RIPPER algorithms might generate complex logical structures with excessive branching, necessitating domain knowledge-guided secondary optimization.

Table 3. Comparative analysis of interpretation methods

Metric	**Association Rules**	**Dependency Analysis**	**Rule Induction**	**SHAP**
High-dim Support	Low	Medium	Low	High
Pattern Efficiency	Low	Medium	Medium	High
Interpretability	Medium	Low	High	High
TB Applicability	Low	Medium	High	High
Clinical Operability	Medium	Low	High	High

Here we selected XGBoost - a tree-based model demonstrating superior performance on tabular data. We trained the XGBoost model using 55 cluster labels as classification targets, then conducted global feature importance analysis on 387 features determining each cluster label, and generated their respective summary plots (bee swarm plots).

As shown in Figure 2, this SHAP (SHapley Additive exPlanations) feature contribution diagram displays: The Y-axis lists semantically mapped feature names using clinically interpretable descriptions, sorted in descending order of global feature importance with the most discriminative key features positioned at the top. The X-axis represents the distribution range of SHAP values, which physically signifies the directional impact of features on sample assignment to specific clusters: Data points distributed on the right side (SHAP values > 0) indicate positive driving effects that enhance model confidence in assigning samples to corresponding clusters; points clustered on the left side (SHAP values < 0) reflect inhibitory effects on cluster membership. The color gradient (red-blue spectrum) encodes the magnitude of original feature values: Red spectrum indicates high feature values (e.g., abnormally elevated biomarker levels), while blue spectrum denotes relatively low-value states (e.g., physiological parameters at lower reference limits).

Figure 2 displays Cluster 30 with the highest persistence score. We selected the TherapyStatus-Feature1 column, corresponding to the TherapyStatus variable containing eight distinct categories in the dataset, as a macroscale feature.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

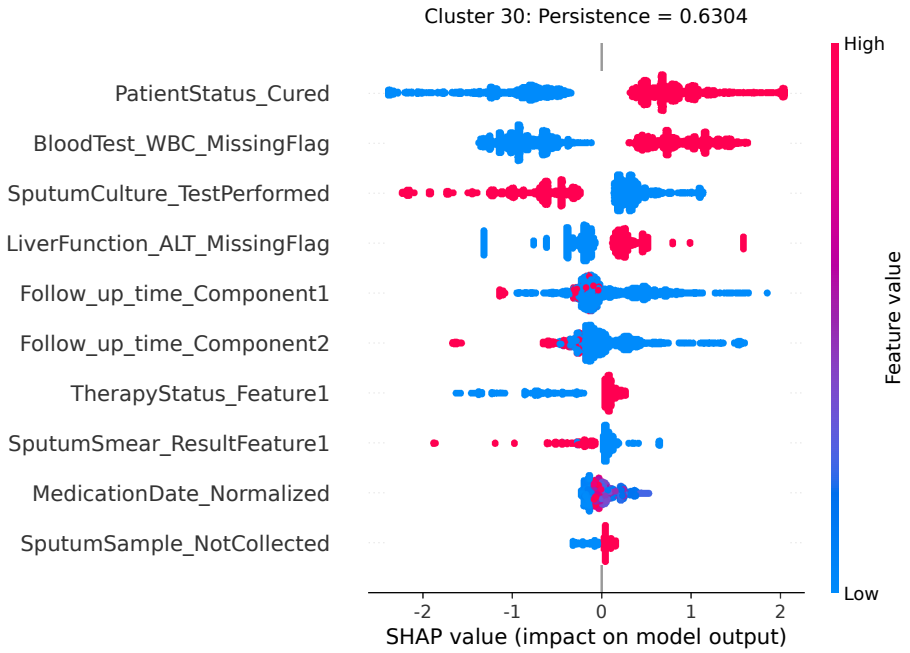

Fig. 2. Global feature importance analysis for Cluster 30

These eight categories are: continuation of existing regimen, no treatment initiated, establishment/modification of treatment regimen, transfer-out, adverse drug reaction, voluntary discontinuation by patient, other, and exclusion of multidrug-resistant tuberculosis (MDR-TB) diagnosis.

The remaining features in the diagram predominantly represent patient-specific characteristics. To identify micro-level features exhibiting strong correlations with the TherapyStatus column, we retrained an XGBoost model using the eight macro-level states as classification targets and computed the predictive contribution of the remaining 386 features. The four most significant features are listed in Table 4.

Table 4. Top-4 micro-Level features contributing to macro-Level TherapyStatus classification

Feature	**SHAP Value**
x_0 =MxDataExt_MxName_NMF_2 (Follow up time-Component2)	0.8150031
x_1 =TbDiagnosis-Multidrug-Resistant Tuberculosis (MDR-TB)	0.6279506
x_2 =LiverFunc.Result.Dbil_DoubleValue (Direct Bilirubin)	0.19392538
x_3 =LiverFunc.Result.Alb_DoubleValue (Albumin)	0.18071306

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

3.3 Transition function search and validation

The macroscale features encoded by the eight TherapyStatus categories were reduced to a single feature vector y via NMF dimensionality reduction. The four features in Table 4 were sequentially designated as x_0, x_1, x_2, x_3 in descending order of importance.

Subsequent symbolic regression was implemented by utilizing the preprocessed globally monotonically increasing matrix as the search space. We employed PySR to execute the evolutionary algorithm with: 1,000 evolutionary rounds, 5 parallel populations, population size of 500 individuals (fundamental evolutionary units), symbolic parameter controlling maximum expression tree nodes $n_{\max} = 20$, and early stopping criteria Terminate if loss $<$ 1e$-$12 or complexity $C > 25$. The search results were autonomously logged with an evaluation rate of 2.710×10^3 expressions/second. Here the parameters of the evolutionary algorithm were configured solely based on computational resource availability, and comparative analysis across multiple independent experimental trials confirmed their negligible impact on search outcomes.

Subsequently, we conducted causal lag analysis on treatment state evolution patterns using the PCMCI (Peter-Clark Momentary Conditional Independence) method. The dynamic impacts of key variables exhibited the following characteristics: x_0 demonstrated negative regulatory effects at lag-1 (-0.0936) and lag-2 (-0.0310); x_1 showed positive driving effects at lag-1 (0.1066) and lag-2 (0.0646); x_2 revealed a positive association at lag-1 (0.0486).

We therefore introduced quadratic and exponential temporal modulation terms for cross-validation. The quadratic modulation form is expressed as $y_{\text{mod}} = y_{\text{base}} \times (1 + at + ct^2)$, where the linear term coefficient $a \in [0.01, 0.15]$ and quadratic term coefficient $c \in [0.001, 0.02]$. Through grid search on the training set (80% samples) with MSE as evaluation metric, we obtained optimal parameters $a = 0.01$, $c = 0.001$, achieving validation MSE 4.6×10^{-5} ($R^2 = 0.98$).

For exponential modulation $y_{\text{mod}} = y_{\text{base}} \times e^{bt}$, the growth rate $b \in [0.01, 0.1]$ was constrained to 2$\times$ the 0.05-level effect of x_2. Using 10-point uniform sampling, we determined optimal parameter $b = 0.1$, yielding validation MSE 4.8×10^{-4} ($R^2 = 0.79$). These results demonstrate effective capture of time-varying treatment state characteristics through our modulation functions.
The final transition functions we obtained are as follows:

$$Y(t) = \left[1.0001 - \left(x_0 + x_1^2\right)^{3.1569 \times 10^{-5}}\right] \times 0.13185 \cdot \left(1 + 0.01t + 0.001t^2\right) \quad (5)$$

$$Y(t) = (0.091956 - x_0)^{\exp(x_2)} \cdot e^{0.1t} \quad (6)$$

Interestingly, the fitting results of these transition functions on the dataset demonstrate that they embody computational formulations bridging macro-micro relationships. This connection manifests mathematically as expressions that remain computable for machines/models yet counterintuitive for human experts.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

4 DyMES simulation on the dataset

We simulate the model using all 20,872 strictly increasing samples from the dataset by incorporating macro and micro feature columns contained in Equations 5 and 6. The DyMES framework is rigorously constructed following Equations 1-4. Our algorithmic innovation introduces precomputed acceleration matrices and covariance matrix-approximated Jacobians to significantly accelerate solving Equations 3 and 4, enabling whole-sample modeling without subsampling.

The precomputed acceleration matrix accelerates candidate function set $f_\nu(x, X, t)$ evaluations across all microstates $x \in x_{\text{array}}$ and macrovariables X. Conventional methods [4] require recomputing f_ν per iteration, yielding $\mathcal{O}(N_{\text{iter}} \cdot N_x \cdot N_\nu)$ complexity (N_x = microstate count, N_ν = constraint count). We preconstruct matrix $f_{\text{matrix}} \in \mathbb{R}^{N_x \times N_\nu}$ with elements: $f_{\text{matrix}}[i, j] = f_j(x_i, X, t)$. Matrix reuse strategy: Reusing f_{matrix} in probability distribution $R(x|X)$, constraint equations, and Jacobian computations reduces complexity to $\mathcal{O}(N_x \cdot N_\nu)$.

Traditional finite difference Jacobian calculation costs $\mathcal{O}(N_\nu^2 \cdot N_x \cdot N_\nu)$ with step-size sensitivity. Through leveraging f_{matrix} and distribution R via $\texttt{np.cov}(f_{\text{matrix}}^\top, \texttt{aweights} = R, \texttt{bias} = \text{True})$, we achieve secondary optimization: Eliminating extra function evaluations reduces complexity to $\mathcal{O}(N_\nu^2 \cdot N_x)$. Since we derived from Equation 1:

$$J_{\mu\nu} = \frac{\partial \left(F_\mu - \mathbb{E}_R[f_\mu]\right)}{\partial \lambda_\nu} = -\text{Cov}_R\left(f_\mu, f_\nu\right) \tag{7}$$

where $\mathbb{E}_R[f_\mu]$ denotes the expectation with respect to the probability distribution R.

Figure 3 demonstrates the fitting results across 100 time steps, where each grid unit on the horizontal axis encompasses 10 time steps. Here, λ_1 corresponds to the mean constraint of x_0 ($h_1(X) = \mathbb{E}[x_0]$), representing the follow-up time Component 2, while λ_2 corresponds to the mean constraint of x_1 ($h_2(X) = \mathbb{E}[x_1]$), associated with the MDR-TB diagnosis status. λ_3 and λ_4 encode the dynamic constraints governed by Equations 5 and 6, respectively. The experiment reveals that λ_2 remained constant, reflecting the stability of x_1's statistical distribution, which implies that the MDR-TB detection status maintains statistical equilibrium during microstate evolution, with the system preserving structural integrity through conserved mean values. The slight decline of λ_3 in later stages indicates temporal accumulation effects in Equation 5's dynamic constraint (quadratic $0.001t^2$ term), requiring prolonged time modulation signal integration to trigger constraint adjustments.

The coupled dynamics of λ_1 and λ_4 – manifested through their exponential decay phase in the first 70 steps – arise from the nonlinear interaction between x_0's mean constraint (λ_1) and Equation 6's dynamic constraint (λ_4) mediated by $\exp(x_2)$. The synergistic decay emerges from the coupling between x_0 and hepatic function indicators (x_2). The abrupt transition in later stages reveals a critical

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

threshold (at $t \approx 70\Delta t$) where the time modulation factor $e^{0.1t}$ in Equation 6 dominates the dynamical phase transition. The non-zero terminal states of λ_3/λ_4 signify the system's evolution toward a novel steady state incorporating time modulation terms, where persistent dynamic constraints maintain a "dynamic memory" encoded jointly by Equation 5's quadratic temporal term and Equation 6's exponential temporal driver.

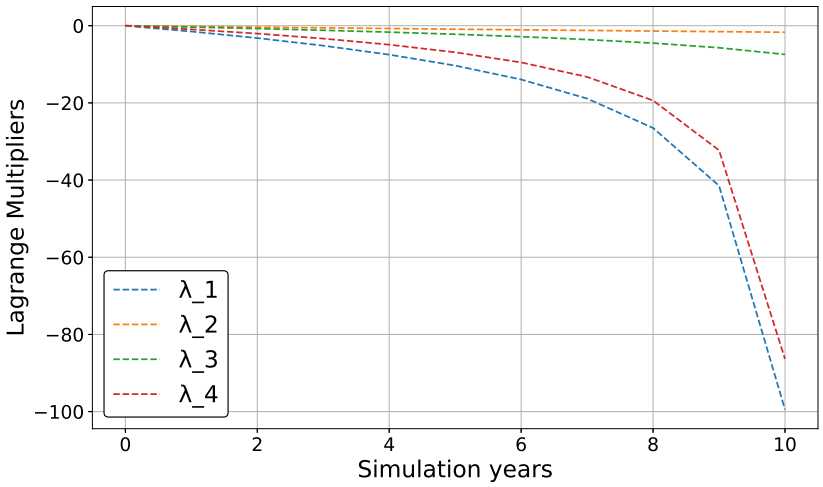

Fig. 3. Dynamic cross-scale norm interactions

5 Discussion and Future Work

In most complex systems, causal relationships prove challenging to disentangle; DyMES may provide a quantitative methodology for determining both the directionality and magnitude of causal links [4]. Distinct from conventional top-down approaches [18], DyMES hybridizes mechanistic principles with Maximum Entropy (Maxent) theory, establishing an inferential framework that bridges fine-scale phenomena with coarse-grained outcomes, thereby enabling prediction of microscopic distributions from macroscopic knowledge. This methodology demonstrates capability in forecasting both the temporal evolution of state variables and probability distributions over microvariables.

These properties hold significant implications for investigating dynamic interactions between microscopic clinical practices/behaviors and macroscopic medical norms in distributed healthcare systems. Particularly, it facilitates modeling the dissemination and shared understanding of medical norms within autonomous multi-agent systems [9]. While conventional reinforcement learning paradigms employ reward-based mechanisms (e.g., reinforcement learning) to

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

characterize and approximate agent behaviors at deeper levels, their utility remains limited for directly analyzing the complex scientific properties and inherent patterns within raw medical information datasets.

The application of DyMES, as a general mathematical framework in complexity science, to construct computable models for medical norm systems presents three principal challenges: (1) Transformation of information from unstructured multi-type non-random missing datasets into modelable microscopic features; (2) Formal definition and dynamic modeling of evolving clinical norms; (3) Reliable extraction of authentic transition function relationships between these elements from empirical data. Our current work systematically addresses these three fundamental issues.

Through implementation of an Explainable AI (XAI) framework, we propose a comprehensive three-phase mathematical modeling approach and empirically validate the effectiveness of extracted transition functions using clinical datasets. This investigation establishes a computational foundation for subsequent analyses of dynamic norm properties in healthcare environments.

Future research directions focus on three primary objectives: (1) Formalization and extraction of comprehensive composite microscopic features coupled with dynamic medical norms; (2) Investigation of bidirectional dynamic norm interactions under DyMES conditions in autonomous multi-agent systems; (3) Examination of micro-level agent practice (feature evolution) impacts on macroscopic norms and reciprocal constraint mechanisms. These explorations are anticipated to drive synergistic evolution of medical normative systems across theoretical and practical domains.

Acknowledgments. The research was supported by the Russian Science Foundation, agreement No. 24-11-00272, https://rscf.ru/project/24-11-00272/.

References

1. Anand, M., Gonzalez, A., Guichard, F., Kolasa, J., Parrott, L.: Ecological systems as complex systems: Challenges for an emerging science. Diversity **2**(3), 395–410 (March 2010). https://doi.org/10.3390/d2030395, http://dx.doi.org/10.3390/d2030395
2. Arfken, G.B.: Mathematical methods for physicists (1967), https://api.semanticscholar.org/CorpusID:122141371
3. Boccara, N.: Modeling Complex Systems. Springer New York (2010). https://doi.org/10.1007/978-1-4419-6562-2, http://dx.doi.org/10.1007/978-1-4419-6562-2
4. Harte, J., Brush, M., Umemura, K., Muralikrishnan, P., Newman, E.A.: Dynamical theory of complex systems with two-way micro–macro causation. Proceedings of the National Academy of Sciences of the United States of America **121** (2024), https://api.semanticscholar.org/CorpusID:274566522
5. Jaynes, E.T.: Information theory and statistical mechanics. Physical Review **106**, 620–630 (1957), https://api.semanticscholar.org/CorpusID:17870175
6. Jaynes, E.T.: On the rationale of maximum-entropy methods. Proceedings of the IEEE **70**, 939–952 (1982), https://api.semanticscholar.org/CorpusID:42335268

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9

7. Kovalchuk, S.V., Funkner, A.A., Metsker, O.G., Yakovlev, A.N.: Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification. Journal of biomedical informatics **82**, 128–142 (2017), https://api.semanticscholar.org/CorpusID:14429841
8. Kovalchuk, S.V., Metsker, O.G., Funkner, A.A., Kisliakovskii, I.O., Nikitin, N.O., Kalyuzhnaya, A.V., Vaganov, D.A., Bochenina, K.O.: A conceptual approach to complex model management with generalized modelling patterns and evolutionary identification. Complexity **2018**, 1–15 (Nov 2018). https://doi.org/10.1155/2018/5870987, http://dx.doi.org/10.1155/2018/5870987
9. Li, C., Petruchik, O., Grishanina, E., Kovalchuk, S.: Multi-agent norm perception and induction in distributed healthcare (2024), https://arxiv.org/abs/2412.18454
10. Ma, J., Zeng, L., Ren, F., yun Dang, L., Luo, H., Wu, Y., Yang, X.J., Li, R., Yang, H., Xu, Y.: Treatment outcomes and risk factors of multidrug-resistant tuberculosis patients in xi'an china, a retrospective cohort study. Infection and Drug Resistance **15**, 4947 – 4957 (2022), https://api.semanticscholar.org/CorpusID:251947734
11. Olsen, M., Kuhn, D.R., Raunak, M.: Explaining the impact of parameter combinations in agent-based models. Journal of Computational Science **81**, 102342 (Sep 2024). https://doi.org/10.1016/j.jocs.2024.102342, http://dx.doi.org/10.1016/j.jocs.2024.102342
12. Pressé, S., Ghosh, K., Lee, J., Dill, K.A.: Principles of maximum entropy and maximum caliber in statistical physics. Reviews of Modern Physics **85**, 1115–1141 (2013), https://api.semanticscholar.org/CorpusID:16150191
13. Rajpurkar, P., Chen, E., Banerjee, O., Topol, E.J.: Ai in health and medicine. Nature Medicine **28**, 31 – 38 (2022), https://api.semanticscholar.org/CorpusID:246098480
14. Runge, J.: Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. ArXiv **abs/2003.03685** (2020), https://api.semanticscholar.org/CorpusID:212633741
15. Sayama, H.: Introduction to the modeling and analysis of Complex Systems. Open Suny Textbooks (Aug 2015)
16. Stallings, W.: Gerald m. weinberg. an introduction to general systems thinking. new york: Wiley, 1975, 279 pp. Behavioral Science **21**(4), 289–290 (July 1976). https://doi.org/10.1002/bs.3830210409, http://dx.doi.org/10.1002/bs.3830210409
17. Udrescu, S.M., Tegmark, M.: Ai feynman: A physics-inspired method for symbolic regression. Science Advances **6** (2019), https://api.semanticscholar.org/CorpusID:167217655
18. Zhu, S.C., Zhang, R., Tu, Z.: Integrating bottom-up/top-down for object recognition by data driven markov chain monte carlo. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662). vol. 1, pp. 738–745. IEEE (2000)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97557-8_9