Dynamic Temporal Modeling of Abdominal Aortic Aneurysm Morphology with Z–SINDy

Joseph A. Pugar[1,**], Junsung Kim[1,**], Nhung Nguyen[1], C. J. Lee[2], Hence Verhagen[3], Ross Milner[1], Andrei A. Klishin[4,*], and Luka Pocivavsek[1,*]

[1]Department of Surgery, University of Chicago, Chicago, IL, USA
[2]Endeavor Health, Evanston, IL, USA
[3]Department of Vascular Surgery, Erasmus University Medical Center, Rotterdam, the Netherlands
[4]Department of Mechanical Engineering, University of Hawai‘i at Mānoa, HI, USA
[*]Authors to whom correspondence should be addressed.
[**]Co-first authors.

Abstract. Abdominal aortic aneurysm (AAA) is characterized by a localized enlargement of the aorta with risk of rupture. Endovascular aneurysm repair (EVAR) alters sac morphology in ways not captured by size alone and typically observed only at sparse follow-up times. We present a physically interpretable, noise-aware framework that models postoperative remodeling dynamics in a low-dimensional state space defined by normalized sac surface area ($\widetilde{A}$) and the normalized fluctuation in integrated Gaussian curvature ($\widetilde{\delta K}$). Using Sparse Identification of Nonlinear Dynamics (Z–SINDy), we infer ordinary differential equations governing the temporal evolution of $(\widetilde{A}, \widetilde{\delta K})$ for clinically defined cohorts of regressing and stable sacs. The learned models yield class-specific flow fields and fixed points that summarize long-term behavior: regressing sacs converge toward a low-size/low-shape attractor, whereas stable sacs maintain near-constant size with persistently elevated shape. Embedding these dynamics in a Bayesian decision framework enables complementary *static* (coordinate-based) and *dynamic* (derivative-based) classifiers. Across cohorts, the dynamic classifier separates outcomes earlier and with higher confidence, reflecting that rates of change become discriminative before static distributions diverge. Stress tests injecting spatial noise and imposing realistic, irregular follow-up cadences quantify performance degradation and sensitivity to measurement fidelity and scheduling. Together, these results provide a principled route from interpretable geometric features to individualized, probabilistic forecasts of AAA remodeling post-EVAR and offer actionable guidance for surveillance design.

Introduction

Abdominal aortic aneurysms (AAA) are localized dilations of the aorta that can develop into a balloon-like *sac* with substantial rupture risk if left untreated. Endovascular aneurysm repair (EVAR) is the dominant minimally invasive intervention used to exclude the sac from arterial flow, promote thrombus formation, and ideally induce sac regression over time (Fig. 1) [1, 2, 3]. Despite widespread adoption, predicting long-term outcomes post-intervention remains challenging [4, 5, 6]. Clinicians primarily rely on size assessments at *discrete time points* (typically cross-sectional diameter or sac volume) to characterize status [7, 8, 9]. In addition, pathological shape, captured by the fluctuation in integrated Gaussian curvature δK, contains information about disease progression [10].

However, instantaneous anatomical descriptors are often insufficient to infer the *underlying trajectory* of disease progression. Differentiating a stable post-EVAR sac from one undergoing regression or expansion requires dynamic characterization of morphology over time. The importance of dynamics is reflected in supplemental guidance that characterizes regression by annual rates of maximal diameter change [11], with thresholds such as $\geq$5–10 mm at 1 year used in practice [4, 12]. Failure to regress at 1 year and sac expansion are associated with higher long-term mortality independent of endoleaks or reinterventions [4, 5], and a "stable" sac carries worse prognosis than a regressing sac [6, 13]. Early sac shrinkage (e.g., $\geq$10 mm at 1 year) is linked to freedom from fatal adverse events [12], whereas delayed and persistent type II endoleaks correlate with sac growth and adverse late outcomes [14, 15]. The ability to model and anticipate trajectories could inform surveillance frequency, enable early identification of treatment failure, and guide personalized intervention strategies. Substantial regression to $\leq$40 mm has been associated with very low risk and may permit extended follow-up intervals in selected patients [16].

The dynamics of AAA remodeling are shaped by biomechanical, anatomical, and patient-specific factors. Real-world datasets are limited in size and irregularly sampled, yet they encode latent patterns of morphological change. Rather than reducing outcomes to discrete labels, it is valuable to view post-EVAR behavior as a continuum governed by rates and directions of change in biologically meaningful features. Clinical imaging is sparse: patients may contribute as few as two usable time points (typically one pre- or perioperative scan and one postoperative scan) and at most about five, separated by months or years. This sparsity—compounded by variation in imaging frequency, segmentation algorithms, and device type—demands models that are robust to noise and heterogeneity and capable of extracting population-level trends from underdetermined individual trajectories. Heterogeneity is further amplified by device-specific effects and repair complexity: most sac regression occurs within the first two years [17]; regression patterns vary across endografts [12]; and complex repairs (fenestrated/branched) exhibit distinct sac dynamics requiring intensified surveillance when growth occurs [18, 19].

Traditional approaches to temporal modeling in clinical science—spanning biostatistics, time-series machine learning [20, 21], and black-box deep learning [22, 21]—are often inadequate here. Linear models underfit nonstationary biological processes, while flexible nonlinear methods risk overfitting noise and spurious fluctuations. These frameworks also rarely accommodate uncertainty from variable sampling intervals or missing data [23, 20, 24]. Black-box models may extract unimportant patterns while overlooking core mechanistic relationships [22, 21]. Effective clinical models must balance complexity and interpretability, capturing coarse-grained system dynamics without overfitting high-dimensional noise. This calls for frameworks that incorporate physical structure, account for data corruption,

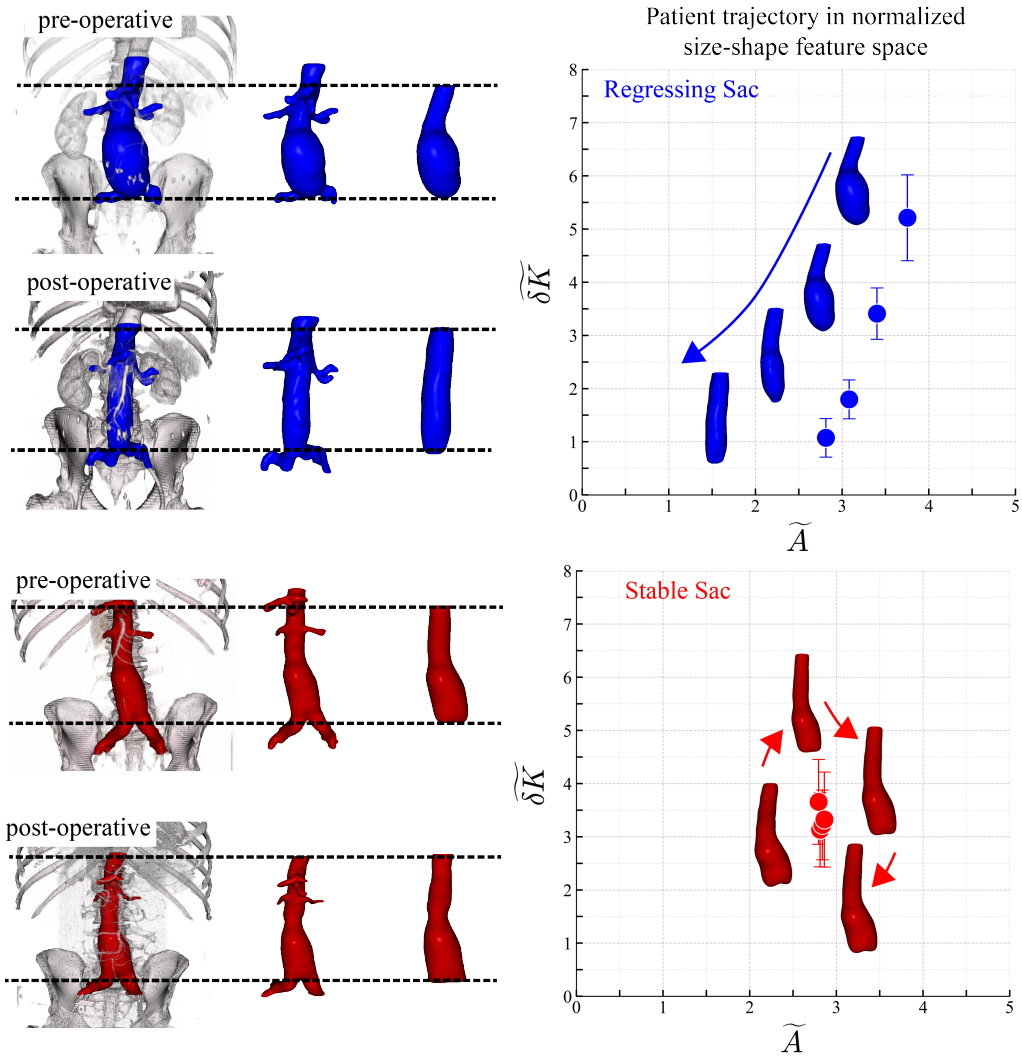


Figure 1: Defining aortic anatomy and mapping trajectories into size–shape space. **Left:** 3D segmentations of AAA scans before and after EVAR. Dashed horizontal lines mark the superior and inferior cut planes used to standardize the analyzed sac segment across time points. **Right:** Each scan is projected into a normalized size–shape space ($\widetilde{A}$, $\widetilde{\delta K}$), where $\widetilde{A}$ denotes normalized surface area and $\widetilde{\delta K}$ the normalized fluctuation in integrated Gaussian curvature. Points are serial scans connected chronologically; arrows indicate temporal progression. Top row: regressing sac (blue; $> 10\%$/year decrease in surface area). Bottom row: stable sac (red; $< 10\%$/year change). Together, the panels show the workflow—from standardized segmentation via consistent cut planes to quantitative embedding—used to characterize postoperative sac behavior in a common, interpretable feature space.

and remain interpretable to clinicians—particularly when predicting forward trajectories under real-world constraints such as follow-up attrition and irregular scan timing. In parallel, morphovolumetric metrics can be more sensitive than diameter alone for detecting clinically meaningful sac change, underscoring the need to target features that better reflect underlying biomechanics [7].

In this work, we develop a physically interpretable, noise-aware framework for modeling post-EVAR AAA remodeling using a sparse identification approach. Sparse Identification of Nonlinear Dynamics (SINDy) [25, 26] has successfully identified governing equations from synthetic data and reduced-order models, but has seen more limited success with noisy experimental data [27, 28, 29]. We use Z–SINDy [30], a version equipped with uncertainty quantification rooted in statistical mechanics. We infer sparse, interpretable ordinary differential equations (ODEs) governing the temporal evolution of sac surface area A and sac shape δK, features that compactly encode aneurysm morphology.

We curated a dataset of 74 patients and 164 CT scans across two clinical cohorts in different countries. Our framework integrates clinical CT time-series data with finite-element–based synthetic upsampling (FEA) [31] to mitigate sparsity and enhance signal. We demonstrate that Z–SINDy reproducibly identifies dynamic signatures associated with sac regression versus stabilization, provides interpretable steady-state fixed points with clinical meaning, and enables individualized probabilistic forecasts of anatomical evolution. We further examine how sampling frequency, noise, and attrition affect model fidelity, and we propose a principled way to assess confidence in outcome prediction using a Bayesian classification model—thereby aligning physical interpretability with clinical utility.

Results

Data curation

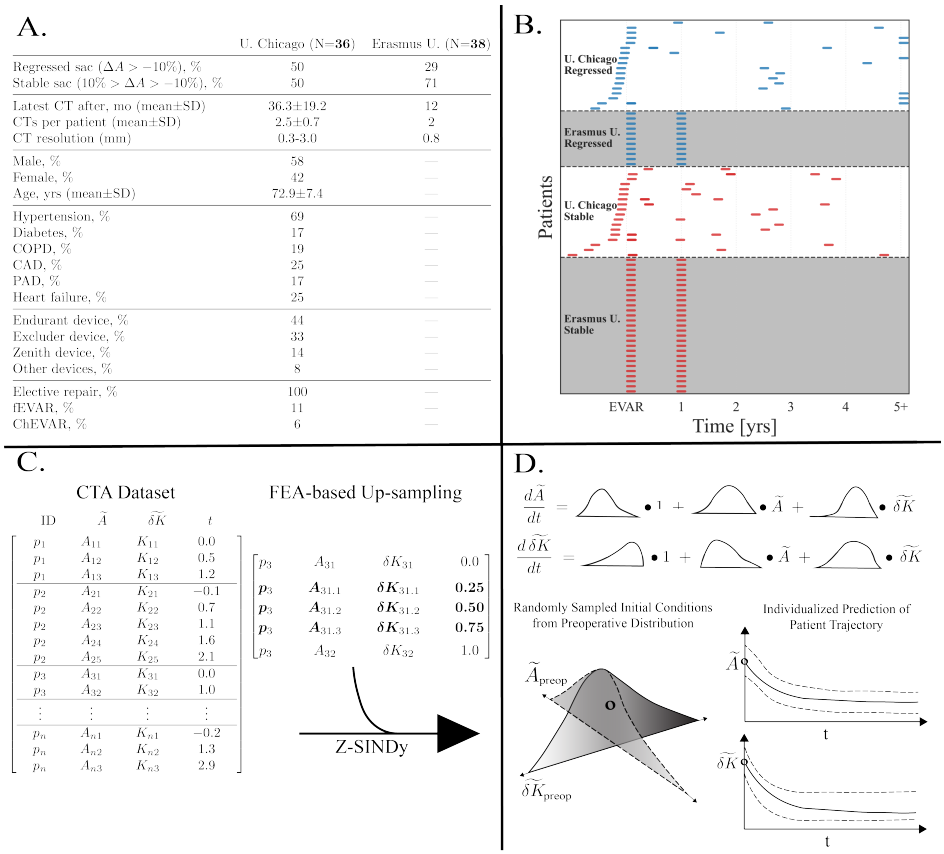
We curated a dataset of 74 patients and 164 CT scans across two clinical cohorts from different institutions: the University of Chicago (USA) and Erasmus University Medical Center (the Netherlands), with basic demographic statistics in Fig. 2A. Owing to differences in healthcare systems, the acquisition cadence differs substantially: in the Erasmus cohort each patient typically has one preoperative CT scan immediately before EVAR and one follow-up exactly 12 months later, whereas in the University of Chicago cohort the preoperative scan may precede EVAR by several months and follow-ups are irregular in both spacing and number (Fig. 2B). The scan-based data points were upsampled with the FEA procedure to increase data density (Fig. 2C).

Modeling post-EVAR dynamics with Z–SINDy

We first asked whether coarse-grained dynamics of postoperative sac remodeling can be captured by low-order, interpretable models in the normalized size–shape space ($\widetilde{A}$, $\widetilde{\delta K}$):

$$\frac{d\widetilde{A}}{dt} = f(\widetilde{A}, \widetilde{\delta K}; \boldsymbol{\Xi}), \qquad \frac{d\widetilde{\delta K}}{dt} = g(\widetilde{A}, \widetilde{\delta K}; \boldsymbol{\Xi}), \tag{1}$$

where $\boldsymbol{\Xi}$ are model coefficients inferred from the training data. We fitted Z–SINDy models separately to subjects clinically labeled *regressing* and *stable* to obtain posterior distributions of coefficients (Fig. 2D). We then resampled both coefficients and initial conditions



	U. Chicago (N=**36**)	Erasmus U. (N=**38**)
Regressed sac ($\Delta A > -10\%$), %	50	29
Stable sac ($10\% > \Delta A > -10\%$), %	50	71
Latest CT after, mo (mean±SD)	36.3±19.2	12
CTs per patient (mean±SD)	2.5±0.7	2
CT resolution (mm)	0.3-3.0	0.8
Male, %	58	—
Female, %	42	—
Age, yrs (mean±SD)	72.9±7.4	—
Hypertension, %	69	—
Diabetes, %	17	—
COPD, %	19	—
CAD, %	25	—
PAD, %	17	—
Heart failure, %	25	—
Endurant device, %	44	—
Excluder device, %	33	—
Zenith device, %	14	—
Other devices, %	8	—
Elective repair, %	100	—
fEVAR, %	11	—
ChEVAR, %	6	—

Figure 2: Dataset, sampling, and Z–SINDy pipeline. **A** Demographic and clinical metadata for the two study cohorts (University of Chicago and Erasmus University) used in all analyses. **B** Barcode plot summarizing data-acquisition cadence: each row is a patient; horizontal position is time post-EVAR; ticks indicate CT acquisitions. This view makes temporal sparsity explicit and shows where derivative information for SINDy can be estimated. **C** Input data matrix for modeling, consisting of discrete observations of normalized size–shape variables ($\widetilde{A}$, $\widetilde{\delta K}$, t) per patient. Patients contribute a variable number of rows reflecting the sparsity seen in panel B. To reduce sparsity, the true observations are augmented with FEA-based surrogates so that each patient has samples on a 3-month grid between their first and last scan. **D** Schematic of the Z–SINDy methodology. The matrix in panel C is converted into a sparse system of ODEs with coefficient distributions. The ODEs are integrated forward from initial conditions sampled from the preoperative distribution ($\widetilde{A}_{\text{preop}}$, $\widetilde{\delta K}_{\text{preop}}$), yielding probabilistic trajectories for the regressing and stable dynamical models.

and integrated the equations forward, yielding ensembles of trajectories with uncertainty bands around the mean solutions (Fig. 3). For the regressing cohort (blue), both $\widetilde{A}(t)$ and $\widetilde{\delta K}(t)$ decline monotonically toward a low-size/low-shape fixed point; for the stable cohort (red), $\widetilde{A}(t)$ remains approximately constant while $\widetilde{\delta K}(t)$ remains elevated relative to the regressing group. Fixed-point analysis of the mean ODEs gives $\left(\widetilde{A}^*, \widetilde{\delta K}^*\right) \approx (2.1,\ 0.2)$ for the regressing model and $\approx (3.8,\ 8.1)$ for the stable model (Fig. 3, dashed lines), providing compact, clinically interpretable summaries of long-term behavior for each class.

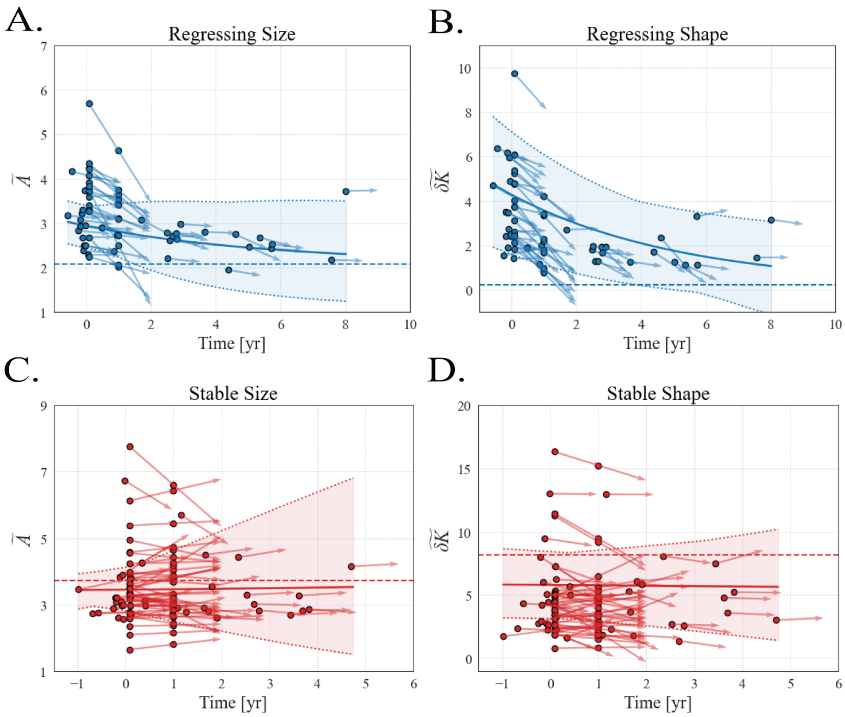


Figure 3: Z–SINDy trajectories for regressing and stable sacs in normalized size–shape space. Panels show the evolution of normalized surface area $\widetilde{A}(t)$ and normalized fluctuation in integrated Gaussian curvature $\widetilde{\delta K}(t)$ over time t (years since EVAR). **A**–**B** (blue): regressing model; **C**–**D** (red): stable model. Points are observed scan values; faint lines are individual Z–SINDy trajectory realizations obtained by integrating the learned ODEs from population-level initial conditions while sampling coefficient distributions. Bold lines are mean trajectories; shaded regions indicate ensemble spread induced by coefficient uncertainty. Dashed horizontal lines mark steady-state (fixed-point) solutions computed from the mean ODEs. Mean fixed points in $(\widetilde{A},\ \widetilde{\delta K})$ are approximately (2.1, 0.2) for the regressing model and (3.8, 8.1) for the stable model. Panel assignments: **A** $\widetilde{A}(t)$ (regressing), **B** $\widetilde{\delta K}(t)$ (regressing), **C** $\widetilde{A}(t)$ (stable), **D** $\widetilde{\delta K}(t)$ (stable).

To relate model-level dynamics to cohort behavior, we rendered population flow fields in the $(\widetilde{A},\ \widetilde{\delta K})$ plane (Fig. 4). The *Raw* panels plot finite-difference patient derivatives overlaid on the convex hull of observed states, showing individual progression vectors: regressing sacs flow down and left, while stable sacs circulate within a higher-$\widetilde{\delta K}$ region with comparatively small changes in $\widetilde{A}$. *Interpolated* derivatives, projected onto a convex-hull grid, concentrate streamlines and motion magnitude, yielding average behavior for each cohort. The *Model*

panels, generated by evaluating the mean Z–SINDy vector fields, mirror the empirical trends and demonstrate the effectiveness of simple parametric models. Stars mark fixed points predicted by each class-specific ODE; local eigenvectors (black line glyphs) align with the principal directions of flow. The regressing model has a stable fixed point (both eigenvalues with negative real parts), whereas the stable model exhibits a saddle (eigenvalues with opposite signs). Characteristic time scales estimated from the eigenvalues via $\tau = 1/|\mathrm{Re}\,\lambda|$ indicate substantial remodeling for regressing sacs ($\sim$5.0 years) and much slower dynamics for stable sacs ($\sim$16.3 and 51.5 years).

Together, Figs. 3–4 show that (i) the two outcome groups are governed by distinct dynamics in a two-variable description and (ii) Z–SINDy reproduces cohort flow structure observed in the empirical trajectories while providing a unified system of equations for each outcome class.

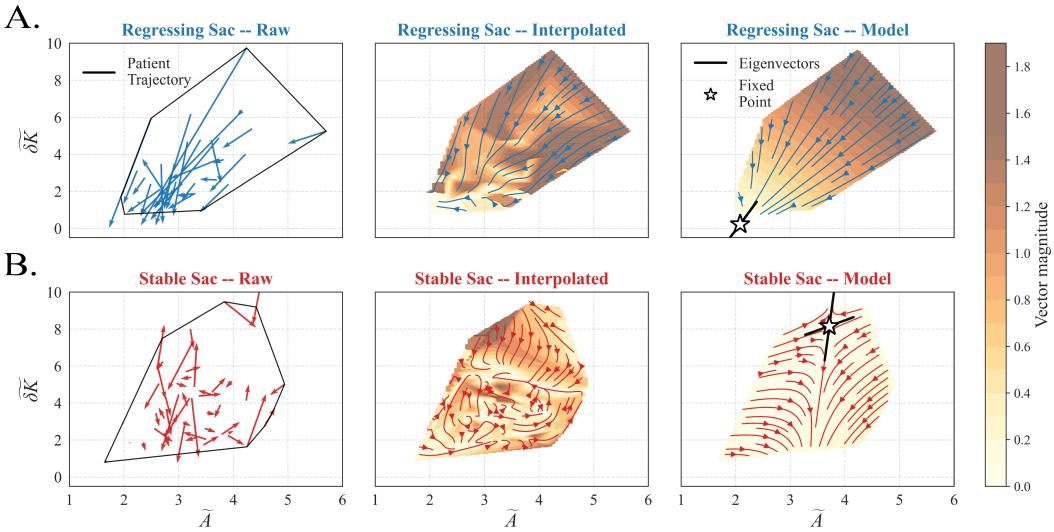


Figure 4: Cohort-level flow fields in normalized size–shape space. Panels illustrate empirical and modeled dynamics of sac remodeling in the $(\widetilde{A}, \widetilde{\delta K})$ plane for regressing (**A**, top row, blue) and stable (**B**, bottom row, red) post-EVAR sacs. *Raw:* finite-difference derivatives from longitudinal trajectories (black outlines = convex hull of observed states; arrows = per-patient vectors). *Interpolated:* smoothed flow fields obtained by gridding/interpolating pooled derivatives with background shading proportional to vector magnitude. *Model:* mean Z–SINDy vector fields with fixed points (stars) and associated eigenvectors (black lines) indicating local stability structure. Eigenvalue real parts set the local time scales; for the regressing model the real parts are similar, yielding comparable decay rates along both eigen-directions.

Bayesian classification and model stress tests

Having established class-specific dynamics, we next asked whether the Z–SINDy models could support individualized, probabilistic classification of postoperative outcome and how early such decisions can be made under simulated real-world sampling. Figure 5A shows ensembles of 10,000 simulated trajectories generated from the two Z–SINDy models. These ensembles reflect plausible temporal evolutions under coefficient uncertainty for regressing (blue) and stable (red) sacs and already illustrate divergence over time. Figure 5B re-

formats the same ensembles into one- and two-dimensional views across reference times. Kernel-density estimates of $\widetilde{A}$ and $\widetilde{\delta K}$ marginals (left and middle) and their joint densities (right) summarize distributional evolution. Immediately after repair the classes overlap; by one to two years their marginals and joint densities diverge, with regressing sacs shifting toward lower $\widetilde{\delta K}$ while stable sacs maintain higher values. Importantly, the middle column of Fig. 5B and the inputs to Fig. 5C use the same underlying data, providing a direct comparison between raw ensemble separation and classifier decision structure. Figure 5C applies the *static* Bayesian classifier, which evaluates likelihoods $p(\mathbf{x}_k \,|\, c, t_k)$ based only on coordinates. By contrast, Fig. 5D shows the *dynamic* Bayesian classifier, which evaluates derivative likelihoods $p(\dot{\mathbf{x}}_k \,|\, c)$ under the class-specific ODEs, utilizing directional flow rather than instantaneous position. Across time points, the dynamic approach yields higher confidence earlier: by one to two years post-EVAR, more patients exit the "uncertain" band (less than 90% confidence) in Fig. 5D than in Fig. 5C, demonstrating that derivative information becomes discriminative before coordinate distributions fully separate.

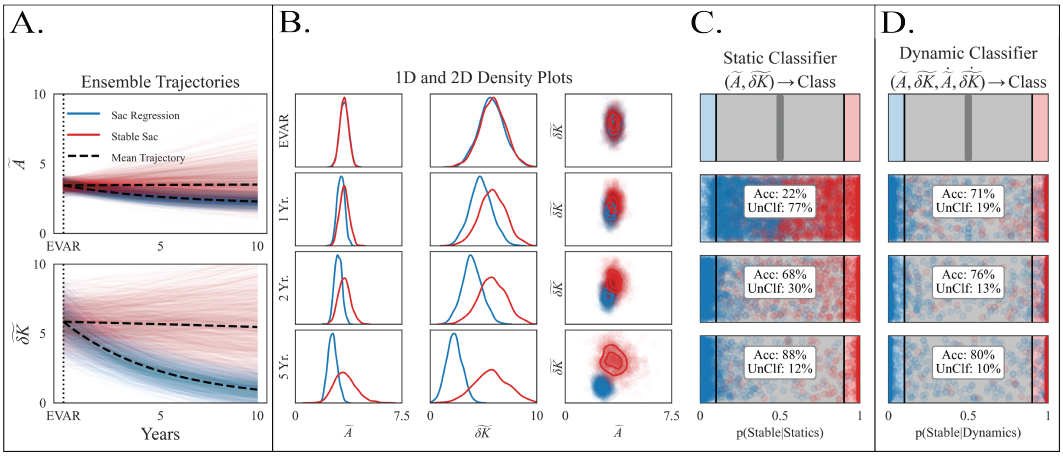


Figure 5: From Z–SINDy ensembles to static and dynamic Bayesian classification. **A** Ensembles of 10,000 simulated trajectories from the regressing (blue) and stable (red) Z–SINDy models for $\widetilde{A}(t)$ (top) and $\widetilde{\delta K}(t)$ (bottom). **B** Distributional views of the same ensembles: 1D marginals (left and middle) and joint densities (right) at EVAR, 1 yr, 2 yr, and 5 yr. The middle column and the inputs to panel C use identical point clouds. **C** Posterior probabilities from the static Bayesian classifier, pooling $p(\mathbf{x}_k \mid c, t_k)$ over observed states. **D** Posterior probabilities from the dynamic Bayesian classifier using empirical derivatives $p(\dot{\mathbf{x}}_k \,|\, c)$. The dynamic classifier exits the "uncertain" region earlier than the static classifier, reflecting the added value of directional information.

Finally, we quantified robustness to clinically realistic spatial and temporal noise (Fig. 6). We injected controlled spatial noise into ($\widetilde{A}$, $\widetilde{\delta K}$) and coarsened sampling from exactly annual (*Optimal 1.0*) to jittered (*Ideal*, 0.5–1.5 yr) and sparse/irregular (*Reality*, 1–3 yr) intervals, re-estimating derivatives and reapplying the dynamic classifier. Three consistent trends emerged: (i) Accuracy improves with postoperative time across all conditions, reflecting growing separation of class dynamics and accumulation of information from additional scans; EVAR corresponds to the non-informative 50% prior. (ii) Spatial noise degrades performance monotonically; when perturbations reach a quarter of the feature scale

("Noise 0.25"), performance approaches the 50% prior. (iii) Temporal irregularity modestly reduces accuracy relative to exactly annual follow-up but does not eliminate monotonic gains with time; the ∼5–10% drops under non-ideal sampling indicate sensitivity to clinic scheduling. Together, these stress tests show that the Z–SINDy + Bayes pipeline degrades gracefully under measurement noise and attrition while maintaining its early advantage over coordinate-only classification.

Across two cohorts and under heterogeneous sampling, a two-variable Z–SINDy representation yields class-specific ODEs with interpretable fixed points and flow structure (Figs. 3–4). Embedded in a Bayesian framework, the models support patient-level posterior inference that (i) leverages dynamical information to outperform static, coordinate-only approaches early in follow-up and (ii) remains robust to spatial/temporal corruptions typical of clinical data (Figs. 5–6). These results establish a path from physically grounded dynamics to practical decision support for EVAR surveillance.

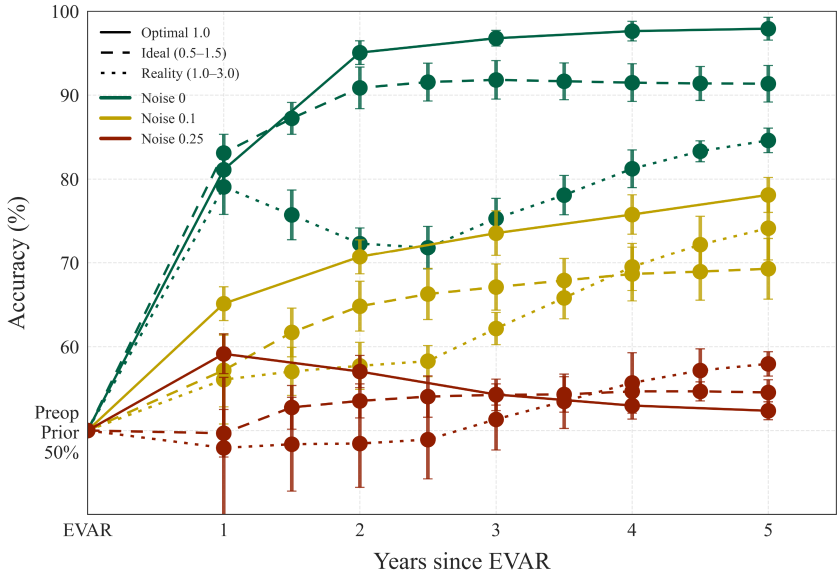


Figure 6: Robustness of the dynamic Bayesian classifier to spatial and temporal noise. Mean classification accuracy versus time since EVAR under six simulated data conditions. Colors encode spatial noise added to the normalized features prior to classification (green: none; yellow: ± 0.1; red: ± 0.25 applied to both $\widetilde{A}$ and $\widetilde{\delta K}$). Line styles encode temporal sampling: solid = exactly annual (*Optimal 1.0*); dashed = jittered intervals randomly sampled from $[0.5, 1.5]$ yr (*Ideal*); dotted = sparse/irregular intervals from $[1.0, 3.0]$ yr (*Reality*). The baseline at EVAR corresponds to the non-informative 50% prior. Error bars represent variability across 50 realizations. Accuracy improves with postoperative time but degrades with spatial perturbations and more clinically realistic temporal irregularity.

Discussion

This work leverages a physically interpretable, anatomically grounded approach to model postoperative AAA remodeling and uses the fitted models for clinically actionable classification. Motivated by the limitations of size-only assessments and by sparse, irregular

follow-up, we framed sac regression or stabilization as dynamics in a low-dimensional, interpretable state space described by normalized size ($\widetilde{A}$) and shape ($\widetilde{\delta K}$). Two main findings emerged: (i) outcome-specific Z–SINDy models capture cohort dynamics and yield meaningful fixed points (Figs. 3–4); and (ii) these dynamics enable probabilistic classification that exploits directional information rather than position alone in size–shape space, providing earlier and more accurate decisions that can be stress-tested for robustness (Figs. 5–6).

Dynamical forecasting

A central challenge for clinically useful time-series models is balancing accuracy with interpretability. By restricting the library to linear terms, Z–SINDy identifies the simplest models consistent with the data while preserving a direct physiological mapping between features and coefficients [32]. For regressing sacs, trajectories trend toward a stable fixed point near $(\widetilde{A}, \widetilde{\delta K}) \approx [1, 1]$, i.e., the definition of a nonpathological aorta in the normalized space; the characteristic time scale ($\sim$5 years) matches typical follow-up imaging. For stable sacs, the mean field exhibits a saddle and much longer time scales; the model predicts near-constant $\widetilde{A}$ with persistently higher $\widetilde{\delta K}$ on clinical time horizons. Importantly, Fig. 4 closes the loop between model and data: raw finite-difference vectors, interpolated flow fields, and Z–SINDy vector fields show consistent directions across the observed convex hull, indicating that the learned ODEs reflect underlying cohort behavior.

Bayesian classifiers

Translating learned dynamics into clinical utility requires principled decision-making under measurement uncertainty. Our first practical contribution is a pair of Bayesian classifiers—*static* (coordinate-based) and *dynamic* (derivative-based)—that take an individual patient's trajectory and return the posterior probability of regressing or stable outcome. Because derivatives encode *where a patient is heading* rather than *where they are now*, the dynamic classifier becomes discriminative sooner: by 1 year it achieves 71% accuracy with only 19% of patients remaining uncertain, whereas the static classifier yields 22% and 77%, respectively (Fig. 5). This is crucial in practice, where only 2–3 scans may be available per patient and early risk stratification can guide surveillance and reintervention strategy. The static classifier remains useful, albeit on longer postoperative timelines.

Stress testing and study design

Our second contribution is explicit quantification of how spatial measurement error (e.g., CT resolution and segmentation variability) [33, 34, 35] and temporal irregularity (loss to follow-up, variable scheduling) [36, 37, 38, 39] influence performance (Fig. 6). Three conclusions follow. (i) Accuracy improves monotonically with postoperative time across conditions, reflecting accumulating information and growing separation between outcomes. (ii) Spatial noise degrades performance in a predictable fashion; when perturbations reach roughly a quarter of the feature scale, performance approaches the 50% prior. (iii) Temporal irregularity induces meaningful losses relative to exactly annual imaging, emphasizing the value of timely scans for derivative estimation. These stress tests provide a quantitative lens for planning studies, enabling investigators to trade off sample size, cadence, and anticipated noise to meet prespecified accuracy targets.

The proposed framework suggests an anatomically defined, rate-based approach to post-EVAR surveillance. The size–shape mapping is straightforward to compute from standard CT-derived segmentations, and Z–SINDy has an adaptable open-source implementation [40]. Posterior-predictive simulations generate reference distributions for both outcomes, supporting downstream analyses and visualization. The framework naturally accommodates new data as additional scans become available, enabling adaptive surveillance: patients whose dynamic posteriors favor the regressing model might maintain standard follow-up, whereas those with persistently high $\widetilde{\delta K}$ could be flagged for closer evaluation or adjunctive treatment. More broadly, summarizing cohort behavior with fixed points (and eigenvectors) aids communication by distilling complex remodeling into intuitive descriptors ("which fixed point are you approaching, and how fast?").

Future work

Our formulation rests on simplifying geometric, dynamical, and clinical assumptions that guide future extensions. *Geometry:* we deliberately restricted the feature space to $\widetilde{A}$ and $\widetilde{\delta K}$ to highlight size–shape coupling; richer variables (e.g., thrombus morphology, graft type, hemodynamics) may refine subgroup structure. The fluctuation in integrated Gaussian curvature δK is empirically informative [10] but remains a heuristic; principled shape metrics from differential geometry could be incorporated [41]. *Dynamics:* we used a linear library to balance fit and interpretability; given current data volume/noise, nonlinear extensions offer limited gains, but future datasets may warrant higher-order terms. Measurement noise chiefly affects empirical derivatives ($\dot{\widetilde{A}}, \dot{\widetilde{\delta K}}$); adaptive denoising and derivative-regularization methods could further stabilize estimates in clinical settings [42, 43, 44, 45]. *Clinical scope:* here we modeled two classes within AAA. Extending to expanding sacs (Type I endoleak or pre-rupture), thoracic aneurysms, post-TEVAR remodeling, and other morphology-driven conditions—paired with stress-test paradigms to inform surveillance design and quality standards—is a natural trajectory. Heterogeneity across devices, imaging practices, and segmentation quality was embraced rather than harmonized, underscoring the need for prospective, multi-institutional validation.

By reframing AAA remodeling as inference of simple, physically meaningful dynamics, we bridge mechanistic understanding and clinical decision-making. Z–SINDy provides compact models that match cohort behavior and yield interpretable flow structure; Bayesian classifiers convert those models into early, patient-specific probabilities; and stress tests reveal when and how performance degrades under real-world constraints. Together, these elements support translational progress toward reliable, transparent, and resource-aware modeling that can inform how, when, and for whom postoperative surveillance should adapt.

Methods

Clinical data and morphological feature definition

We curated longitudinal computed tomography (CT) imaging for patients treated with endovascular aneurysm repair (EVAR). Each patient contributed a small number of usable time points (typically 2–5), separated by months to years. To place all subjects in a common, low-dimensional morphological space, we extracted two features from each segmented sac surface: (i) sac surface area A, and (ii) a shape descriptor defined as the fluctuation in

integrated Gaussian curvature, δK, computed over the sac surface. To improve the anatomical descriptors, we applied scale averaging and normalized both features by cohort means to obtain dimensionless variables $\widetilde{A}$ and $\widetilde{\delta K}$, following our previously established methods on the role of $\widetilde{\delta K}$ in disease-state prediction [46]. The patient state at time t is thus $\mathbf{x}(t) = [\widetilde{A}(t),\, \widetilde{\delta K}(t)]^\top$.

FEA-based upsampling and polynomial resampling

To address irregular clinical follow-up and inherent sparsity, we upsampled trajectories using a two-step procedure. Using finite-element analysis (FEA) simulations, we first constructed dense surrogate trajectories for A and δK that respect sac mechanics (see prior work for details [31]). These sequences furnish physically plausible intermediate states between clinical scans without extrapolating beyond the postoperative ground truth $\mathbf{x}(t_f)$. We then performed uniform resampling by fitting, for each subject, a third-order polynomial to each feature as a function of time and evaluating the fit at $\Delta t = 0.25$-year increments from the subject's first to last scan. If this evaluation exceeded a fixed per-subject point budget, we instead generated exactly $N = 20$ evenly spaced samples between the observed endpoints. This soft cap limits oversampling of any single trajectory while preserving short-term changes. The resulting dataset is sufficiently upsampled and evenly spaced for derivative estimation and dynamical modeling, while mitigating variance inherent to the native clinical data (e.g., infrequent sampling and patient attrition).

Derivative estimation

We estimated time derivatives using pairwise finite differences between successive observations within subject. To avoid spuriously large derivatives from nearly coincident scans, we required a minimum time separation of 3 months. Let $(t_i,\, \widetilde{A}_i,\, \widetilde{\delta K}_i)_{i=1}^{N}$ denote the ordered time points and corresponding normalized features for a subject. To estimate instantaneous rates of change $(\dot{\widetilde{A}}, \dot{\widetilde{\delta K}})$, we applied second-order central differences to the compressed time series:

$$\dot{\widetilde{A}}_i \;=\; \frac{\widetilde{A}_{i+1} - \widetilde{A}_{i-1}}{t_{i+1} - t_{i-1}}, \quad \dot{\widetilde{\delta K}}_i \;=\; \frac{\widetilde{\delta K}_{i+1} - \widetilde{\delta K}_{i-1}}{t_{i+1} - t_{i-1}}, \quad \text{for } i = 2, \ldots, N-1. \tag{2}$$

We also applied light, physically motivated outlier rejection on $\mathrm{d}\,\widetilde{\delta K}/\mathrm{d}t$ rates to remove biologically implausible spikes introduced by algorithmic error (segmentation, curvature estimation, or timestamps). Specifically, if

$$\frac{\mathrm{d}\,\widetilde{\delta K}}{\mathrm{d}t} \;>\; 6\ \mathrm{yr}^{-1} \quad \text{or} \quad <\; -3\ \mathrm{yr}^{-1},$$

the offending scans were removed from the dataset, and the remaining data were then used to fit the dynamical model.

Sparse identification of dynamics (Z–SINDy)

We modeled the temporal evolution of the morphological state $\mathbf{x}(t) = [\widetilde{A}(t),\, \widetilde{\delta K}(t)]^\top$ using a sparse identification approach with uncertainty quantification (Z–SINDy) [30]. We seek

to explain the dynamics in the form

$$\dot{\mathbf{x}}(t) \approx \boldsymbol{\Xi}^\top \boldsymbol{\Theta}(\mathbf{x}(t)), \quad \boldsymbol{\Xi} \in \mathbb{R}^{N\times 2}, \tag{3}$$

where $\boldsymbol{\Theta}$ is a library of N (possibly nonlinear) functions of the coordinates and $\boldsymbol{\Xi}$ is a sparse matrix of coefficients (i.e., most entries are zero). We use Bayesian inference to find the coefficients, borrowing the setup from Z–SINDy. The overall Bayesian formula is

$$p(\boldsymbol{\Xi} \mid \{\dot{\mathbf{x}}\}) = \frac{p(\{\dot{\mathbf{x}}\} \mid \boldsymbol{\Xi})\, p(\boldsymbol{\Xi})}{p(\{\dot{\mathbf{x}}\})}, \tag{4}$$

where $\{\dot{\mathbf{x}}\}$ is the set of available derivative data; computations require defining the likelihood and prior in the numerator. For the likelihood we assume Gaussian observation noise around the model prediction:

$$p(\{\dot{\mathbf{x}}\} \mid \boldsymbol{\Xi}) \propto \prod_l \exp\left(-\frac{1}{2\rho^2} \sum_k \left(\dot{x}_{l,k} - \boldsymbol{\Xi}_l^\top \boldsymbol{\Theta}(\mathbf{x}_k) \right)^2 \right), \tag{5}$$

where noise is assumed independent across dimension l and data point k, $\boldsymbol{\Xi}_l$ is the lth column of $\boldsymbol{\Xi}$, and the resolution hyperparameter ρ controls the noise magnitude (estimated below). For the prior we penalize models by sparsity (number of active terms) rather than coefficient magnitude, using a Bernoulli–Gaussian prior:

$$p(\boldsymbol{\Xi}_l) \propto \prod_i \left(\delta(\Xi_{l,i}) + w(\Xi_{l,i})\, e^{-\lambda n} \right), \tag{6}$$

where $w(\cdot)$ is a wide Gaussian (flat in the limit), λ is a sparsity penalty, and n is the number of data points used for fitting.

Posterior probability Given the above likelihood and prior, the posterior distribution becomes a Gaussian mixture:

$$p(\boldsymbol{\Xi}_l \mid \{\dot{\mathbf{x}}\}) = \frac{\sum_\gamma p(\boldsymbol{\Xi}_l \mid \gamma, \{\dot{\mathbf{x}}\})\, \mathcal{Z}_\gamma}{\sum_\gamma \mathcal{Z}_\gamma}, \tag{7}$$

where γ indexes a particular set of active terms and $\mathcal{Z}_\gamma$ is the statistical weight driven by fit and sparsity [30]. In practice the mixture is dominated by a single set. In the main text we restrict the function library to an affine form (constant and linear terms):

$$\boldsymbol{\Theta}(\mathbf{x}) = [\,1,\ \widetilde{A},\ \widetilde{\delta K}\,]^\top, \tag{8}$$

and explore alternate libraries in the SI. Within a fixed active set, the parameter distribution is multivariate Gaussian with mean and covariance computed from empirical covariances on the training data:

$$\boldsymbol{\Xi}_l \mid \{\dot{\mathbf{x}}\} \sim \mathcal{N}(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}), \tag{9}$$

$$\boldsymbol{\mu} = \mathbf{C}^{-1}\mathbf{V}, \tag{10}$$

$$\boldsymbol{\Sigma} = \rho^2 \mathbf{C}^{-1}, \tag{11}$$

$$\mathbf{C} \equiv \sum_k \boldsymbol{\Theta}(\mathbf{x}_k)\, \boldsymbol{\Theta}(\mathbf{x}_k)^\top, \tag{12}$$

$$\mathbf{V} \equiv \sum_k \boldsymbol{\Theta}(\mathbf{x}_k)\, \dot{\mathbf{x}}_k^\top, \tag{13}$$

where $\boldsymbol{\mu} \in \mathbb{R}^{N\times 2}$ and $\boldsymbol{\mu}_l$ denotes the lth column.

Resolution matching and hyperparameter tuning The posterior covariance in Eqn. 11 scales with the resolution parameter ρ, which we determine from the data. Given the *mean* fitted model, we estimate ρ from the variance of the model-fit residual:

$$\rho \;\approx\; \sqrt{\frac{1}{d\,n}\sum_{k,l}\big(\dot{x}_{k,l} - \boldsymbol{\mu}_l^\top \boldsymbol{\Theta}(\mathbf{x}_k)\big)^2}, \tag{14}$$

which is then used to compute $\boldsymbol{\Sigma}$.

Uncertainty-aware simulation and fixed-point analysis

We quantified forecast uncertainty by generating ensembles via sampling model coefficients from the posterior $\boldsymbol{\Xi}^{(m)} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and initial conditions from a Gaussian around a preoperative baseline. Each sampled model was integrated forward to obtain trajectories $\mathbf{x}^{(m)}(t)$. We visualized uncertainty with shaded bands about the mean trajectory for each class (*regressing* and *stable*) for $\mathbf{x}(t)$ and for $\dot{\mathbf{x}}(t) = \boldsymbol{\Xi}^{(m)\top}\boldsymbol{\Theta}(\mathbf{x}(t))$. Fixed points $\mathbf{x}^*$ were computed from the mean model by solving $\dot{\mathbf{x}}(\mathbf{x}) = 0$ and are plotted as horizontal lines in the respective panels.

Population flow-field visualization in morphological space

To compare empirical patient trajectories with model-predicted dynamics, we visualized cohort-level flow fields in the $(\widetilde{A}, \widetilde{\delta K})$ plane. For the original CT-derived data (labeled *Raw*), we computed the convex hull of all observed states and overlaid finite-difference estimates of $\big(\mathrm{d}\,\widetilde{A}/\mathrm{d}t,\ \mathrm{d}\,\widetilde{\delta K}/\mathrm{d}t\big)$ as arrows anchored at the corresponding scan locations. Next, these derivative estimates were pooled across patients and interpolated onto a regular grid spanning the convex hull, with the background magnitude $\|\dot{\mathbf{x}}\|$ rendered to indicate local dynamic speed (labeled *Interpolated*). Streamlines were seeded using a kernel-density estimate over the empirical point cloud so that streamline density reflects regions with data support. Lastly, we evaluated the mean Z–SINDy vector field for each class, rendering streamlines and shading by vector magnitude (labeled *Model*). Fixed points from the mean ODEs are marked with stars, and associated eigenvectors are shown as black lines to indicate local stability structure.

Dynamic Bayesian classification

We constructed a *dynamic* Bayesian classifier using the fitted dynamical models for two outcome classes. For a patient with observations at times $\{t_k\}$, we used the empirical derivatives $\{\dot{\mathbf{x}}_k\}$ and evaluated, for each class $c \in \{\text{regress}, \text{stable}\}$,

$$z_c \;=\; p(\{\dot{\mathbf{x}}_k\} \mid c) \;=\; \prod_{k,\ell} \frac{1}{(2\pi\rho_c^2)^{1/2}}\, \exp\!\left(-\frac{1}{2\rho_c^2}\,\big(\dot{x}_{k,\ell} - \boldsymbol{\Xi}_{c,\ell}^\top \boldsymbol{\Theta}(\mathbf{x}_k)\big)^2\right), \tag{15}$$

where $\boldsymbol{\Xi}_c$ is the class-specific coefficient matrix and the class-specific resolution ρ_c is estimated from each model's residuals. Here $d = 2$ is the dimension of dynamics and n is the number of derivative data points; the total normalization is $(2\pi\rho_c^2)^{dn/2}$. The posterior class probability is

$$p(c \mid \{\dot{\mathbf{x}}_k\}) \;=\; \frac{z_c\, p(c)}{\sum_{c'} z_{c'}\, p(c')}, \tag{16}$$

using $p(c) = 1/2$ as the prior. Likelihoods can be computed at any postoperative time window.

Static Bayesian classification

As a complement, we defined a *static* classifier that operates directly on observed states. For each class model, we simulated an ensemble of trajectories from distributions of initial conditions and coefficient samples, then estimated a time-indexed likelihood $p(\mathbf{x} \mid c, t)$ over simulated states at time t. For a patient with observations $\{\mathbf{x}_k, t_k\}$, we computed the joint likelihood and the posterior:

$$z_c = \prod_k p(\mathbf{x}_k \mid c, t_k), \tag{17}$$

$$p(c \mid \{\mathbf{x}_k, t_k\}) = \frac{z_c \, p(c)}{\sum_{c'} z_{c'} \, p(c')}. \tag{18}$$

Robustness to clinical sparsity: noise and sampling inhomogeneity stress tests

To emulate real-world data degradation, we applied controlled corruptions to the resampled trajectories before classification. First, we injected spatial Gaussian noise with a standard deviation set to a fixed percentage of each feature's magnitude (e.g., 10% and 25% of $\widetilde{A}$ and $\widetilde{\delta K}$), independently at each time point. Second, we used either regular temporal sampling (e.g., annual intervals) or random time increments within bounds (e.g., 0.5–1.5 years), thereby reducing effective time resolution to stress the dynamic classifier. After corruption, we re-estimated derivatives where applicable, reapplied the dynamic classifier, and quantified shifts in posterior class probabilities, directly assessing sensitivity to segmentation/measurement noise and to attrition typical of clinical datasets.

Author contributions

Conceptualization: L.P., A.K., J.P., J.K. Data Curation: J.K., J.P. Formal Analysis: J.P. Funding Acquisition: L.P. Investigation: J.P. Methodology: L.P., A.K. Project Administration: L.P. Resources: L.P. Supervision: L.P., A.K. Validation: J.P. Writing - Original Draft: J.P. Writing - Review & Editing: L.P., A.K., J.P.

References

[1] R. M. Greenhalgh et al. "Endovascular versus open repair of abdominal aortic aneurysm". In: *New England Journal of Medicine* 362 (2010), pp. 1863–1871.

[2] J. L. De Bruin et al. "Long-Term Outcome of Open or Endovascular Repair of Abdominal Aortic Aneurysm". In: *New England Journal of Medicine* 362 (2010), pp. 1881–1889.

[3] N. Pandey et al. "Surveillance Imaging Following Endovascular Aneurysm Repair". In: *Seminars in Interventional Radiology* 32 (2015), pp. 239–247.

[4] T. F. X. O'Donnell et al. "Aneurysm sac failure to regress after endovascular aneurysm repair is associated with lower long-term survival". In: *Journal of Vascular Surgery* 69 (2019), pp. 414–422.

[5] S. E. Deery et al. "Aneurysm sac expansion is independently associated with late mortality in patients treated with endovascular aneurysm repair". In: *Journal of Vascular Surgery* 67 (2018), pp. 157–164.

[6] G. J. Boer et al. "A stable aneurysm sac after endovascular aneurysm repair as a predictor for mortality: An in-depth analysis". In: *Journal of Vascular Surgery* 76 (2022), pp. 445–453.

[7] M. Franchin et al. "A morphovolumetric analysis of aneurysm sac evolution after elective endovascular abdominal aortic repair". In: *Journal of Vascular Surgery* 74 (2021), pp. 1423–1432.

[8] T. Smith and K. B. Quencer. "Best Practice Guidelines: Imaging Surveillance After Endovascular Aneurysm Repair". In: *American Journal of Roentgenology* 214 (2020), pp. 1220–1230.

[9] F. Zaiem et al. "A Systematic Review of Surveillance After Endovascular Aortic Repair: Imaging Modality and Frequency". In: *Journal of Vascular Surgery* 67 (2018), 320–331.e37.

[10] K. Khabaz et al. "The geometric evolution of aortic dissections: Predicting surgical success using fluctuations in integrated Gaussian curvature". In: *PLOS Computational Biology* 20 (2024), e1011815.

[11] E. L. Chaikof et al. "The Society for Vascular Surgery Practice Guidelines on the Care of Patients with an Abdominal Aortic Aneurysm". In: *Journal of Vascular Surgery* 67 (2018), 2–77.e2.

[12] S. Ikeda et al. "One-year sac regression is associated with freedom from fatal adverse events after endovascular aneurysm repair". In: *Journal of Vascular Surgery* 77 (2023), 136–142.e2.

[13] G. Tinelli et al. "The Sac Evolution Imaging Follow-up After Endovascular Aneurysm Repair". In: *Journal of Vascular Surgery* 80 (2024), pp. 510–520.

[14] W. Zhou et al. "Outcome and clinical significance of delayed endoleaks after endovascular aneurysm repair". In: *Journal of Vascular Surgery* 59 (2014), pp. 915–920.

[15] J. E. Jones et al. "Persistent type 2 endoleak after endovascular repair of abdominal aortic aneurysm is associated with adverse late outcomes". In: *Journal of Vascular Surgery* 46 (2007), pp. 1–8.

[16] E. Andraska et al. "Longer follow-up intervals following EVAR are safe and appropriate after marked aneurysm sac regression". In: *Journal of Vascular Surgery* 76 (2022), pp. 454–460.

[17] P. Jetty et al. "Variability in aneurysm sac regression after endovascular aneurysm repair based on a comprehensive registry of patients in Eastern Ontario". In: *Journal of Vascular Surgery* 70 (2019), pp. 1469–1478.

[18] V. Rastogi et al. "Aneurysm Sac Dynamics and its Prognostic Significance Following Fenestrated and Branched Endovascular Aortic Aneurysm Repair". In: *European Journal of Vascular and Endovascular Surgery* 67 (2024), pp. 728–736.

[19] H. L. Marecki et al. "Characterization and management of type II and complex endoleaks after fenestrated/branched endovascular aneurysm repair". In: *Journal of Vascular Surgery* 78 (2023), pp. 29–37.

[20] Y. Rubanova, R. T. Q. Chen, and D. Duvenaud. "Latent ODEs for Irregularly-Sampled Time Series". In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2019.

[21] F. Doshi-Velez and B. Kim. "Towards A Rigorous Science of Interpretable Machine Learning". In: *arXiv:1702.08608* (2017).

[22] C. Rudin. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead". In: *Nature Machine Intelligence* 1 (2019), pp. 206–215.

[23] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[24] P. Kidger et al. "Neural Controlled Differential Equations for Irregular Time Series". In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2020.

[25] S. L. Brunton, J. L. Proctor, and J. N. Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems". In: *Proceedings of the National Academy of Sciences* 113 (2016), pp. 3932–3937.

[26] S. H. Rudy et al. "Data-driven discovery of partial differential equations". In: *Science Advances* 3 (2017), e1602614.

[27] B. M. De Silva et al. "Discovery of physics from data: Universal laws and discrepancies". In: *Frontiers in Artificial Intelligence* 3 (2020), p. 25.

[28] M. S. Schmitt et al. "Machine learning interpretable models of cell mechanics from protein images". In: *Cell* 187 (2024), pp. 481–494.

[29] R. Lyons, V. Dukic, and D. M. Bortz. "Learning Structured Population Models from Data with WSINDy". In: *arXiv preprint arXiv:2506.24101* (2025).

[30] A. A. Klishin et al. "Statistical mechanics of dynamical system identification". In: *Physical Review Research* 7 (2025), p. 033181.

[31] K. Khabaz et al. "Temporal geometric mapping defines morphoelastic growth model of type B aortic dissection evolution". In: *Computers in Biology and Medicine* 182 (2024), p. 109194.

[32] J. N. Kutz and S. L. Brunton. "Parsimony as the ultimate regularizer for physics-informed machine learning". In: *Nonlinear Dynamics* 107 (2022), pp. 1801–1817.

[33] P. Jaakkola et al. "Interobserver variability in measuring the dimensions of the abdominal aorta: Comparison of ultrasound and computed tomography". In: *Radiology* 198 (1996), pp. 763–767.

[34] N. S. Cayne et al. "Variability of Maximal Abdominal Aortic Aneurysm Diameter Measurements on Computed Tomography: Implications for Clinical Decision Making". In: *Journal of Vascular Surgery* 40 (2004), pp. 678–683.

[35] J. R. Leach et al. "Abdominal Aortic Aneurysm Measurement at CT and MRI: potential clinical ramifications of non-standard measurment technique and importance of multiplanar reformation". In: *Quant Imaging Med Surg.* 11 (2021), pp. 823–830.

[36] T. Garg, L. C. Baker, and M. W. Mell. "Adherence to Postoperative Surveillance Guidelines After Endovascular Aortic Aneurysm Repair Among Medicare Beneficiaries". In: *JAMA Surgery* 150 (2015), pp. 550–557.

[37] A. Schanzer et al. "Follow-up Compliance After Endovascular Abdominal Aortic Aneurysm Repair in Medicare Beneficiaries". In: *Journal of Vascular Surgery* 61 (2015), 16–22.e1.

[38] G. A. Antoniou et al. "Meta-Analysis of Compliance with Endovascular Aneurysm Repair Surveillance". In: *European Journal of Vascular and Endovascular Surgery* 65 (2023), pp. 875–884.

[39] L. E. Newton, A. Ponukumati, and G. Zwain. "Imaging Surveillance Adherence After Endovascular Aneurysm Repair at US Veterans Affairs Hospitals". In: *JAMA Network Open* 8 (2025).

[40] J. Bakarji and A. A. Klishin. *ZSINDy*. 2024.

[41] T. J. Willmore. *Total Curvature in Riemannian Geometry*. Ellis Horwood, 1982.

[42] R. Chartrand. "Numerical differentiation of noisy, nonsmooth data". In: *International Scholarly Research Notices* 2011 (2011).

[43] A. Cortiella, K.-C. Park, and A. Doostan. "A priori denoising strategies for sparse identification of nonlinear dynamical systems: A comparative study". In: *Journal of Computing and Information Science in Engineering* 23 (2023), p. 011004.

[44] J. M. Hokanson, G. Iaccarino, and A. Doostan. "Simultaneous identification and denoising of dynamical systems". In: *SIAM Journal on Scientific Computing* 45 (2023), A1413–A1437.

[45] L. Fung. "Overcoming error-in-variable problem in data-driven model discovery by orthogonal distance regression". In: *arXiv:2507.23426* (2025).

[46] J. A. Pugar et al. "Thoracic Aortic Shape: A Data-Driven Scale Space Approach". In: *medRxiv* (2024).