Credibility assessment of explicit-dynamics models for neurovascular coil packing density: dual-model, tri-mechanism study under ASME V&V 40

Lena Hart, PhD[a], Nikhil Rao, PhD[b], Simone Calder, MS[c], Miguel Alvarez, PhD[1], Priya Venkataraman, PhD[1], Jordan P. Meyer, MD[1]

[a] *Center for Computational Reliability in Medicine, Metro Institute of Technology, Boston, MA, USA,*
[b] *Department of Biomedical Engineering, North Valley University, Sacramento, CA, USA,*
[c] *Neurointerventional Service, St. Helena Medical Center, Denver, CO, USA,*

Abstract

Purpose: Packing density achieved by detachable neurovascular embolization coils is a key determinant of early flow disruption and long-term durability in saccular intracranial aneurysm treatment. This study performs a risk-informed credibility assessment, under ASME V&V 40, of two explicit-dynamics computational models that estimate intrasaccular coil configuration and density across three mechanisms: (M1) deployment into a compliant sac, (M2) early settling under physiologic pulsatility, and (M3) friction heterogeneity due to nascent thrombus. Methods: Model A resolves coil mechanics with geometrically exact beam elements and frictional coil–coil and coil–wall contact in a distensible aneurysm phantom. Model B is a reduced-order, explicit mass–spring–damper surrogate with penalty contact. We planned and executed bench tests using silicone aneurysm phantoms of known compliance, physiologic hemodynamic actuation, and controlled wall friction heterogeneity. Micro-computed tomography (micro-CT) and 3D fluoroscopy captured coil geometry to derive packing density, regional void fraction, and coil–wall surface engagement. We defined an Evidence Strength Scale (ESS) from 0 to 5 and scored six credibility factors for each model × mechanism pair: model structural representation, code-level numerical checks, correspondence of model outputs to measurements, practical value of the computed metrics to the decision, alignment of the experiments with the intended use, and fidelity of the environmental and procedural setup. Results: For Model A, absolute error in packing density was 2.8% for M1, 7.8% for M2, and 11.6% for M3; for Model B the respective errors were 5.6%, 10.1%, and 14.8%. Energy error on a manufactured coil dynamics benchmark was 0.7% for Model A and 1.9% for Model B at production timesteps. The bench conditions reproduced 37 ± 0.2 °C fluid temperature, 3.2 mPa·s viscosity, and 90/140 mmHg pulsation at 1.2 Hz with traceable calibrations, and wall compliance within 0.6 kPa1 to 1.4 kPa1. Scores on the ESS reflected stronger evidence for Model A in M1–M2 and diminished confidence in M3 for both models due to contact friction heterogeneity. Conclusions: The explicit beam–contact formulation provides decision-useful estimates of packing density immediately after placement and following early compaction, whereas the reduced-order surrogate supports rapid scenario screening with modest accuracy. Heterogeneous friction remains a dominant uncertainty that limits predictions of final voids. The factor-by-factor scores elucidate where additional evidence would most efficiently raise credibility: systematic friction mapping and targeted validation of void fraction dynamics under spatially varying wall adhesion.

Keywords: aneurysm coils, packing density, V&V 40

1. Introduction

Endovascular coil embolization is a widely used treatment for saccular intracranial aneurysms, particularly in locations where open surgical access to the aneurysm neck poses high risk. The immediate hemodynamic intent of coil placement is to attenuate intra-aneurysmal flow and trigger thrombosis, while the long-term structural goal is stable occlusion without migration or compaction. Within this multi-scale problem, the volumetric fraction of the sac occupied by coils—often termed packing density—remains one of the few accessible, quantitative indicators that link procedural execution to both early and mid-term outcomes. Retrospective series have associated higher coil fill with lower recurrence, although the relationship is modulated by neck geometry, sac compliance, and the evolving material state of nascent thrombus.

In practice, interventionalists select coil families, primary diameters, and lengths guided by aneurysm size, neck width, and intra-procedural feedback of coil responses in the sac. Although commercial planning tools provide catalog-based coverage suggestions, explicit physics-based models could add value in three areas that are difficult to infer from tables: (i) how wall distensibility modulates initial fill, (ii) how physiologic pulsatility compacts the coil bundle within minutes, and (iii) how localized variations in friction due to thrombus precursors alter void distribution. Explicit dynamics are particularly suited to capture transient contact-rich interactions among slender coils and a compliant sac, but their credibility must be established for the intended question and use.

ASME V&V 40 provides a framework for risk-informed credibility assessment tailored to medical devices. This framework couples the importance of a decision to patient safety and regulatory impact with the degree of evidence required to trust model predictions in that decision. Our study instantiates the framework for a detachable embolization coil system through a dual-model, tri-mechanism assessment. We developed two explicit-dynamics models with differing levels of resolution and computational cost; then, we executed a targeted bench program to generate measurements for both initial deployment and short-term settling. The overarching aim is to quantify evidence for each model–mechanism pair rather than to declare a universal winner, thereby offering transparent guidance on where and how each model may be relied upon.

Three mechanisms are emphasized based on clinical relevance within the early post-implantation horizon: deployment into a distensible sac, immediate post-deployment settling under pulsatile pressure, and friction heterogeneity from nascent thrombus formation. These mechanisms are not exhaustive of coil behavior—long-term biologic remodeling and parent-vessel hemodynamics lie outside our scope—but they represent the time window and physics that dominate the packing density achieved in the interventional suite and the ensuing minutes.

The present article reports, in full, the computational model formulations, the test articles and actuation conditions, the planning of credibility activities, the results of model–data comparisons for quantities that speak to the clinical question, and the factor-by-factor evidence scores. The results are organized to allow readers to follow both the flow of scientific argument and the practical summaries used for decision-making within a V&V 40 evaluation.

Evidence Strength Scale (ESS) levels used throughout — gradation.

a.
0: No direct evidence; assertions unsupported or extrapolated.

b.
1: Minimal evidence from indirect sources or analogs; no quantitative checks.

c.
2: Limited quantitative checks or partial analog testing; gaps in traceability.

d.
3: Moderate, targeted evidence with traceable tests; some residual gaps.

e.
4: Strong, mechanism-specific evidence with convergent sources; minor gaps.

2. Question of Interest and Context of Use

The question of interest is: To what extent can explicit-dynamics models estimate the intrasaccular coil configuration and packing density for a detachable embolization coil system immediately after deployment and after a brief period of physiologic pulsatility, including the influence of spatially heterogeneous wall friction? The immediate clinical decision supported by such estimates is the selection of coil sizes and the determination of whether additional coils are required to achieve a target fill during treatment of a saccular intracranial aneurysm.

The context of use is a preprocedural and intraprocedural computational aid that predicts final intrasaccular geometry, global packing density, and regional voids for a specified aneurysm sac shape, neck geometry, and wall distensibility, given a coil product line with known mechanical properties and surface chemistry. Predictions are consumed by engineers during device design and by clinical users when training or rehearsing a case on geometrically similar phantoms. The time horizon is minutes rather than months, and the target decisions tolerate modest error in coil configuration provided the global fill fraction and prominent voids are reliably captured.

The computed metrics—volumetric fill of the sac by the coils, the fraction of volume left unoccupied regionally, and the area of surface engagement between coils and sac—are directly actioned by interventionalists when deciding whether to deploy an additional coil or to reposition the current coil; they are also available from the bench modalities without transformation through unrelated surrogates.

Fitness of computed metrics to the clinical decision — gradation.

a.
Metrics weakly correlated with decisions; require surrogate translation before use.

b.
Metrics partially aligned; qualitative mapping to decisions seldom consistent.

c.
Metrics aligned but not routinely measurable in tests; indirect linkage.

d.
Metrics directly measurable in tests and used qualitatively in decisions.

e.
Metrics directly measurable and quantitatively used in decisions in some settings.

3. Model Risk Analysis

We analyzed model risk with respect to the specific decision and patient safety consequences as envisioned by ASME V&V 40. The decision class is procedural planning and device design iteration, not real-time control, and thus the direct patient risk from an erroneous prediction is indirect and mediated through clinician judgment. However, systematically optimistic estimates of packing density could encourage undersizing or premature termination of coil deployment, potentially elevating the risk of residual flow channels and coil compaction. Conversely, systematic underestimation could prompt unnecessary additional coil placement, increasing cost and procedure time. Accordingly, we assessed the consequences of model inadequacy as moderate and the decision criticality as moderate, which places the credibility target at a level where mechanism-specific evidence is desirable and where evidence gaps should be explicitly supported by sensitivity analysis or conservative margins.

Three sources of uncertainty dominate the risk posture in this context: uncertainties in wall mechanical properties across patient-specific aneurysms, uncertainties in local friction

[redacted]

Figure 1: Schematic of the study design showing the two computational models, the three mechanisms of interest, and the flow of validation data from the bench program into model–data comparison and credibility scoring.

coefficients as thrombus precursors form, and uncertainties in transient contact patterns during the initial seconds of deployment. The bench program prioritized these aspects by spanning a range of wall compliances observed in clinical and autopsy studies, embedding spatially varying surface coatings to mimic nascent thrombus effects, and using high-speed imaging of coil ingress and early settling to elucidate transient patterns.

The bench experiments mirrored the deployment workflow, sac stiffness range, and pulsation amplitude typical of coil embolization in the immediate post-placement interval, successfully reducing the chance that any model–data disagreements stem from using an unrealistic scenario rather than from the models themselves.

Decision consequence tiering for risk-informed credibility — gradation.

a. Minimal consequence: incorrect model has negligible impact on safety or cost.

b. Low consequence: minor impact; decisions easily reversible without harm.

c. Moderate consequence: impact mediated through user judgment; reversible but meaningful.

d. High consequence: direct safety impact; difficult to reverse.

e. Very high consequence: life-critical; direct link to morbidity/mortality.

4. Mechanisms of Interest

M1—Deployment into a compliant sac: In the first seconds of coil deployment, the advancing helix interacts with a distensible sac where local expansion allows the primary coil to seat differently than in a rigid cavity. Wall distensibility varies with aneurysm composition and history; ex vivo studies suggest a compliance range equivalent to 0.6 kPa1 to 1.4 kPa1 for thin-walled saccular outpouchings. The trajectory and seating of the first coil loops define the scaffold for subsequent loops; therefore, the coupling of coil stiffness and wall compliance can alter achievable fill fraction even with identical coil catalogs.

M2—Early post-deployment settling under cyclic pressurization: Within minutes after detachment, physiologic pulsatility can reconfigure the coil bundle. Pressure oscillations of 90/140 mmHg at around 1.2 Hz generate wall motions and sac volume fluctuations that induce sliding and compaction in the coil pack. These rearrangements can alter the volumetric fill, particularly in larger sacs and where stiffer coils were used. Prior studies have observed additional compaction of 3%–8% by volume within minutes to hours; here we focus on the first 5–10 minutes to remain within the experimental window where thrombus growth is minimal.

M3—Contact friction heterogeneity due to nascent thrombus: Within minutes, biochemical precursors and platelet aggregates can begin to alter surface friction locally. In a benchtop analog, spatial variation in surface energy or hydration can be introduced through patterned coatings or hydrogels. Such heterogeneity encourages asymmetric coil redistribution and can cause larger voids to persist or migrate. Because friction laws at the coil–wall and coil–coil interfaces mediate stick–slip behavior, the predictive power of any deployment simulator depends on the realism of these contact properties.

Environmental and procedural settings, including fluid temperature at 37 ± 0.2 °C, a heparinized glycerol–saline mixture with viscosity 3.2 mPa·s, inlet pressure oscillations cycling between 90 and 140 mmHg at 1.2 Hz, and catheter advancement speeds of 0.6–1.2 mm/s, were held within or bracketing clinical values and recorded with traceable calibrations to decouple physics errors from environmental mismatches.

Mechanism coverage completeness — gradation.

a. Mechanisms only vaguely stated; key drivers omitted.

b. Mechanisms identified but not justified with data or literature.

c. Mechanisms justified qualitatively; partial data support.

d. Mechanisms justified with quantitative ranges from literature or pilot tests.

e. Mechanisms justified and spanned by targeted experiments.

5. Computational Models

Two explicit-dynamics models were constructed to estimate final intrasaccular geometry, packing density, and related observables. Model A resolves the coil as a geometrically exact, shear-deformable beam with circular cross-section, discretized using Simo–Reissner beam elements with selective reduced integration to avoid shear locking. The coil's primary helical shape is embedded via an intrinsic curvature field, with secondary looping behavior emerging from contact interactions and global constraints imposed by the sac geometry and catheter path. Time integration uses a central-difference explicit scheme with adaptive timestep control to satisfy a safety factor of 0.7 on the critical timestep determined by the highest mode of the discretized coil. Contact among coils and with the sac is modeled using a penalty method with a Hunt–Crossley normal force law and a regularized Coulomb friction model that transitions from static to kinetic friction via an exponential slip rate law. The sac wall is represented as a thin shell coupled to an interior cavity pressure boundary condition; a Fung-type nonlinear constitutive model captures distensibility measured in the phantoms. Coil–wall friction coefficients were assigned from measured ranges (coil–wall 0.15–0.30 in baseline, with spatial patches 0.05–0.40 in M3) and coil–coil friction coefficients were 0.08–0.15; these values are consistent with stainless steel or platinum alloy coils with hydrophilic or bare surfaces interacting with silicone surfaces treated to control energy.

Model B is a reduced-order surrogate designed for computational speed. The coil is represented as a chain of lumped masses connected by axial, bending, and torsional springs and dashpots that reproduce the global force–displacement characteristics measured for the coil family. The chain aligns along a reference helix whose curvature can be modulated by local bending state, allowing the surrogate to transition between straightened and coiled segments as it is pushed into the sac. Contact is handled with a penalty approach using spherically blended nearest-neighbor distances, with a viscous regularization term to stabilize high-frequency chatter. Friction is modeled with a rate-and-state surrogate calibrated to reproduce macroscale sliding forces in ring-on-plate tests. The sac wall is modeled as a quasi-static compliance using radial springs anchored to the sac surface; pressure-induced expansions are imposed as boundary displacements consistent with the measured compliance curve. Time integration uses an explicit Euler–Cromer update with a velocity cap to control overshoot at coarse timesteps.

Both models read input geometries from triangulated surface meshes of aneurysm phantoms and use catheter centerlines from bench measurements. Coils are initialized by advancing the proximal end into the neck plane at prescribed feed rates. For M2, a pulsatile internal pressure was applied to the sac model for 5 minutes of physical time, and the simulation continued until kinematic energies decayed below 0.5% of initial deployment maxima. For M3, spatially varying friction patches were embedded on the sac wall mesh or state variables in the surrogate, according to the bench pattern: four 30-degree arcs around the equator and a 15-mm-diameter patch on the dome with elevated or reduced friction levels relative to background.

Round-trip energy error on a manufactured pendular coil benchmark, which exercises large rotations, curvature change, and repeated contact–separation cycles, remained below 0.7% for Model A and 1.9% for Model B across the stable timestep range used in production runs, indicating that integration and contact resolution errors were small compared with experimental scatter in density measurements.

Numerical implementation checks — gradation.

a. No code-level checks beyond compilation.

b. Basic unit tests for isolated functions; no dynamics checks.

c. Simple regression tests compare to past runs; non-physical benchmarks.

d. Manufactured-solution or analytic tests for elements; limited contact tests.

e. Manufactured-solution tests and targeted contact benchmarks representative of use.

6. Credibility Plan and Evidence Strength Scale

We constructed a credibility plan anchored in the risk posture derived above and in the specific mechanisms of interest. The plan enumerates six evidence threads for each model × mechanism pair: adequacy of the structural representation to express the governing physics, quantitative checks of the time integration and contact algorithms, concordance of simulated and measured outputs in relevant tests, clinical relevance of the computed metrics to the decision, proximity of the test scenarios to the intended use, and the degree to which environmental and procedural settings match real-world deployment. Each thread leads to measurements, analyses, and numerical expectations that can be assessed independently and combined to yield an evidence strength level.

To quantify the strength of the evidence we used an ordinal scale from 0 to 5, the Evidence Strength Scale (ESS), where levels 0 and 1 represent qualitative or indirect support, 2 and 3 represent partial or mechanism-targeted quantitative support with gaps, and 4 and 5 indicate strong or comprehensive, replicated evidence aligned to each mechanism. Scores were assigned by consensus of three assessors after reviewing the documentation and data. Disagreements were resolved through re-examination of the test artifacts and, when needed, additional sensitivity analyses.

Traceability from the plan to the evidence was maintained with unique identifiers for each test and simulation, cross-referenced in an evidence map. We emphasized convergent lines of argument where possible—for instance, comparing both global fill and regional void fraction in M2 to ensure that matching packing density was not achieved by compensating geometry errors. Where a thread was judged weak or unsupported, we documented the reason explicitly and scored it at the lower end of the scale to avoid the false impression of strength through omission.

The planned outputs include numerically stable time histories of coil insertion force, coil tip kinematics at the neck plane, final intrasaccular geometry after detachment, volumetric fill, regional void fraction partitioned into sac regions (dome, equator, near-neck), and coil–wall contact area distributions. The validation measurements were designed to provide corresponding observables using high-speed 3D fluoroscopy during deployment, micro-CT after detachment and post-pulsation settling, and surface staining to indicate areas of sustained coil–wall contact.

Planning completeness for credibility activities — gradation.

a.
No explicit credibility plan.

b.
Partial plan lists activities without linking to decisions or mechanisms.

c.
Plan covers activities and mechanisms; weak traceability to evidence.

d.
Plan covers activities, mechanisms, and traceability; limited go/no-go criteria.

e.
Plan includes clear go/no-go criteria per thread and cross-references to evidence IDs.

7. Test Article and Test Conditions

We fabricated elastomeric aneurysm phantoms with hemispheroidal and lobulated geometries representing internal carotid and anterior communicating artery aneurysm morphologies. The phantoms were cast in two-part platinum-cured silicone with Shore 00-30 base hardness, with wall thickness controlled between 150 and 350 micrometers using mandrels. Compliance curves were obtained by incrementally pressurizing the sac from 60 to 160 mmHg and measuring volume change via a piston displacement transducer, yielding sac compliance in the range of 0.6 kPa1 to 1.4 kPa1; these values were embedded into the shell constitutive model for Model A and into radial springs for Model B. Neck diameters ranged from 3.0 to 5.5 mm, and sac volumes ranged from 120 to 450 mm3.

A detachable platinum-tungsten coil system was used with coil primary diameters of 2–6 mm and wire diameter of 30–35 micrometers. Coil axial stiffness and bending rigidity were measured on benchtop fixtures: axial spring constants ranged from 0.15 to 0.35 N/mm, and bending stiffnesses (EI) ranged from 0.12 to 0.31 mN·mm2 across the product line. Ring-on-plate friction tests characterized coil–silicone friction with and without hydrophilic surface treatments: baseline kinetic coefficients of 0.18 ± 0.03 for bare silicone and 0.26 ± 0.05 for treated surfaces; coil–coil friction measured 0.11 ± 0.02. To simulate nascent thrombus, we patterned the sac interior with four equatorial arcs and one dome patch treated with polyvinyl alcohol (PVA) hydrogel layers at two hydration states, producing local friction coefficients of 0.05–0.12 (low) and 0.32–0.40 (high). The pattern geometry was digitized and mapped to the sac surface for model inputs.

Deployment was performed through a 0.021-inch inner-diameter microcatheter inserted through a rigid parent-vessel analog. Coil advance rates were 0.6–1.2 mm/s controlled by a linear actuator with encoder feedback. Physiologic pulsation was generated with a programmable piston pump, delivering 90/140 mmHg at 1.2 Hz; waveform fidelity was verified with a calibrated pressure transducer positioned at the sac inlet. The fluid was a glycerol–saline solution adjusted to 3.2 mPa·s at 37 °C to approximate blood viscosity. The test bath was temperature-controlled at 37 ± 0.2 °C, and the sac wall motion under pulsation was recorded with a laser displacement sensor to confirm volume oscillation amplitude. Each condition was repeated in triplicate across three sac geometries, yielding nine runs per mechanism per model input set.

Imaging used two modalities: during deployment, high-speed biplanar fluoroscopy captured coil ingress and looping at 120 frames per second, and after detachment and settling, micro-CT at 15-micrometer isotropic resolution provided coil geometry for volumetric analysis. Packing density was computed as the ratio of coil volume (computed from the wire diameter and path length extracted from segmentation) to sac volume (computed from pre-scan segmentation), reported in absolute percentage points. Regional void fraction was computed by partitioning the sac into dome, equator, and near-neck regions and subtracting the local coil volume fraction from unity. Coil–wall contact area was inferred from a combination of micro-CT proximity analysis (coil nodes within 50 micrometers of the wall surface) and a post-test staining protocol in which a contact-sensitive dye marked sustained coil contact during pulsation.

Environmental and procedural control quality — gradation.

a.
Uncontrolled, undocumented environment.

b.
Partial control with qualitative logs; missing calibrations.

c.
Quantitative control for some variables; incomplete calibrations.

d.
Quantitative control for key variables with traceable calibrations.

e.
Quantitative control for all relevant variables; bracketing of physiologic range.

8. Equivalency of Input Parameters

Mapping bench measurements to model inputs required several translation steps. The compliance curve of the silicone sac, measured under quasi-static pressure increments, was transformed to parameters of the Fung-type shell model in Model A and to radial spring constants in Model B. Coil–wall friction coefficients measured in ring-on-plate tests were assigned to contact pairs in the models, while acknowledging that curvature

Figure 2: Photograph of the aneurysm phantom mounted in the pulsatile flow loop, showing the microcatheter position, pressure transducer, and temperature-controlled bath.

and local normal forces during deployment differ from the ring geometry. The spatial friction pattern produced by PVA coatings was digitized from photographs and registered to the computational sac surfaces via texture coordinates. Coil mechanical properties derived from axial and bending fixtures informed the intrinsic curvature and stiffness matrices in the beam model and the spring constants in the surrogate. Where measurements existed only at discrete points, spline fits provided continuous parameter fields.

$$\text{CI\_\{score\}} = \frac{\sum_{i=1}^{n} w_i \, x_i}{\sum_{i=1}^{n} w_i} \tag{1}$$

$$\text{\textbackslash varepsilon\_\{rel\}} = \frac{\sum_{i=1}^{n} w_i \, x_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

$$\text{U\_\{val\}} = \frac{\sum_{i=1}^{n} w_i \, x_i}{\sum_{i=1}^{n} w_i} \tag{3}$$

9. Model Form Assessment

The adequacy of the structural representation was reviewed against the physics needed to express the three mechanisms. For deployment into a compliant sac (M1), the beam-based Model A can express large rotations, bending–torsion coupling, and dense contact networks among slender bodies and a distensible shell. The regularized Coulomb friction model allows stick–slip transitions as local normal force and tangential velocity change. The mass–spring–damper chain in Model B coarse-grains these effects into lumped stiffness and damping, potentially attenuating the fidelity of local loop formation and twisting, but can still capture bulk compaction and wall engagement through appropriate parameterization. For early settling under pulsatility (M2), both models incorporate sac compliance and cyclic forcing; Model A does so with an explicit shell responding to pressure, while Model B imposes boundary displacements congruent with the compliance curve. For friction heterogeneity (M3), Model A's contact framework supports spatially varying coefficients at the node-to-surface level; Model B maps heterogeneity through state variables associated with chain segments that interact with sac patches, with a risk of smearing abrupt transitions.

Both models neglect fluid–structure interaction beyond prescribed pressure forcing and viscous regularization in contact; circulation within the sac during deployment is not explicitly modeled. This omission is reasonable for the quantities of interest because intrasaccular flow is low-inertia and the dominant energy exchanges are in coil bending and frictional dissipation. Furthermore, both models enforce kinematic compatibility at the catheter neck plane but do not resolve detailed catheter tip stiffness or torque; this was an intentional simplification to retain generality across packs and vendors and to minimize uncontrolled degrees of freedom.

Parametric studies were performed to ensure that the models could express variability seen in experiments. For Model A, reducing coil bending stiffness by 20% increased initial packing density by 2.1 percentage points for compliant sacs and 0.8 for stiffer sacs, consistent with observed trends. For Model B, similar stiffness reduction increased density by 1.4 and 0.6 points, respectively. Introducing local friction patches in M3 produced persistent voids at the equator in both models, but the magnitude and spatial sharpness were larger in Model A. These studies support the hypothesis that the model forms are not overconstrained and can capture observed sensitivities, though sharp friction gradients remain challenging for the lower-order surrogate.

Adequacy of structural representation — gradation.

a.
Form omits governing physics for mechanisms; key effects unexpressible.

b.
Form includes some physics; crucial interactions (e.g., contact) missing or ad hoc.

c.
Form includes governing physics but with simplifications that restrict mechanism expression.

d.
Form expresses mechanisms with validated submodels; some approximations remain.

e.
Form expresses mechanisms with high fidelity and spatial heterogeneity.

10. Numerical Code Verification

We conducted element-level and system-level numerical checks to guard against implementation errors. For Model A, Simo–Reissner beam elements were verified against analytic solutions for cantilever bending and torsion, achieving displacement and rotation errors below 0.3% at 32 elements per coil turn. A manufactured solution test imposed known curvature and torsion fields on a closed loop with an analytic displacement field; the numerical solution matched the manufactured field with L2 norm error under 0.5% at the production discretization. Contact treatment was probed with a sliding ring on a cylindrical post; normal penetration remained under 2% of the wire diameter for penalty stiffness values chosen to maintain timestep stability. Time integration stability was assessed by advancing a single coil through cycles of bending and release; energy fluctuations remained bounded and decayed with viscous regularization as expected from the scheme. The round-trip energy error benchmark exercised multiple contact pairs through ingress, looping, and withdrawal, and produced a 0.7% error at the production timestep of 0.9 microseconds.

For Model B, chain segment stiffnesses were verified by comparing static deflections to those of the beam model under equivalent loading, matching within 3%. The Euler–Cromer integrator was tested on harmonic oscillators and on the manufactured pendular benchmark adapted to the surrogate mass–spring topology; phase error remained bounded, and energy error on the round-trip test was 1.9% at a timestep of 3 microseconds. Contact regularization parameters were tuned to avoid numerical chatter; residual penetration averaged 3% of wire diameter in dense pack regions. Regression tests confirmed invariant results across compiler and hardware for identical inputs. Both codes were profiled for race conditions and memory errors using thread sanitizers, and no defects were found in the tested configurations.

These checks establish that, at the implemented discretizations and timesteps, numerical errors are small relative to the experimental variability in measured packing density (standard deviation 2.2 percentage points across repeats) and regional void fraction (standard deviation 3.5 percentage points across regions). They do not, however, prove correctness of the models; rather, they increase confidence that observed discrepancies with experiments are attributable to model form or inputs rather than coding defects.

Rigor of code-level verification — gradation.

a.
No quantitative code verification.

b.
Only unit tests; no physics-based checks.

c.
Some physics-based checks; no contact or dynamics verification.

d.
Element-level checks and simple dynamics benchmarks.

e.
Element-level checks plus contact/dynamics benchmarks aligned with use.

11. Relevance of the Quantities of Interest

We selected three computed outputs: global packing density, regional void fraction, and coil–wall contact area distribution after deployment and after early settling. Packing density is the most widely cited fill metric and has operational thresholds in many labs (e.g., 20%–30% minimum target for small necks, higher for wide-neck aneurysms). Regional void fraction highlights whether persistent low-fill zones exist at the dome or near the neck where residual flow channels may persist; clinicians often assess this qualitatively in the angio suite. Coil–wall contact area gauges how much of the sac surface is mechanically engaged by the coil pack; although not used as a procedural metric directly, it is a physically grounded precursor to assessing the potential for surface thrombus and to understanding compaction risk.

These outputs are accessible from the bench data without surrogate proxies: micro-CT provides coil segmentations from which coil volume and contact proximity can be computed, and sac volume is known from the phantom design and prescan. Regional partitioning is performed on the same geometric model used for simulation, ensuring consistent spatial definitions. During deployment, biplanar fluoroscopy informs coil ingress patterns that can be qualitatively compared to model trajectories to check plausibility, though we do not rely on deployment-time fluoroscopy for quantitative fill because of overlapping structures and projection ambiguity.

Alternative outputs considered included coil insertion force and coil tip kinematics at the neck plane. While informative for model tuning and plausibility checks, these were not elevated as primary quantities of interest because they are not directly used by clinicians to decide on additional coil deployment and because force measurements are less stable across devices and setups. We therefore retained them as secondary observables supporting qualitative assessment of the deployment model rather than as scored outputs.

Directness of selected outputs — gradation.

a.
Selected outputs are proxies of proxies; low visibility to end decision.

b.
Outputs are proxies with weak relation to decision triggers.

c.
Outputs relate to decision triggers but require significant interpretation.

d.
Outputs map qualitatively to decision triggers; measured in tests.

e.
Outputs map quantitatively to decision triggers; measured with same modalities.

12. Validation Activities and Relevance to the COU

Validation centered on bench trials that bracket clinically observed sac compliances and drive coils into phantoms under a flow loop tuned to physiologic pressure and frequency. For each sac geometry, we conducted three serial deployments: baseline homogeneous friction surface (M1), the same sac exposed to 5 minutes of pulsation (M2), and the same sac treated with friction heterogeneity patterns (M3). This sequence allowed separation of immediate deployment effects from subsequent settling and from friction-driven redistribution.

We used the same coil product line and sizes as in typical anterior circulation aneurysm cases; the catheter path and neck plane were matched to the sac geometry. The timing between detachment and micro-CT scanning was minimized to reduce changes unrelated to the intended mechanism. For M3, the friction patterning was applied after the M2 run to avoid cross-contamination. The sequence was repeated across three sac geometries and three coil series (primary diameters 3, 4, and 5 mm), covering small-to-moderate aneurysms and neck ratios commonly encountered.

We recognize that bench environments cannot fully replicate patient-specific variability, including vessel tortuosity, intraprocedural pharmacology, and thrombus biochemistry. Nonetheless, for the time horizon of minutes and the quantities of interest selected, the targeted bench program captures the dominant physical influences and provides a tractable, controllable, and traceable environment for comparison.

Alignment of validation with intended use — gradation.

a.
Validation scenario unrelated to intended use.

b.
Validation uses distant analogs; mechanisms misaligned.

c.
Validation partially aligned; key use conditions missing.

d.
Validation aligned in most respects; minor deviations noted.

e.
Validation closely aligned with intended use across mechanisms.

13. Output Comparison

Simulated outputs were compared with bench measurements for M1, M2, and M3 across the three sac geometries and coil series. For M1, Model A's predicted packing density matched micro-CT-derived values within an absolute mean error of 2.8 percentage points (95% limits of agreement 4.1 to +4.7 points). Regional void fraction errors averaged 3.9 points across the dome, equator, and near-neck partitions. Model B's M1 packing density error was 5.6 points (8.2 to +8.9), and regional void fraction errors averaged 6.7 points. Qualitative overlays of simulated and fluoroscopic coil trajectories at the neck plane showed consistent looping patterns for both models, with Model A capturing tighter first-loop radii in compliant sacs.

For M2, after 5 minutes of pulsation at 90/140 mmHg and 1.2 Hz, both models predicted compaction relative to the as-deployed state. Model A's packing density error relative to micro-CT after settling was 7.8 points (10.2 to +11.5), and regional void fraction errors increased modestly to 7.2 points, reflecting the sensitivity of contact rearrangements to friction and wall motion. Model B's M2 packing density error was 10.1 points (13.7 to +14.6), and regional void fraction errors were 9.5 points. Simulated coil–wall contact area changes from pre- to post-pulsation were directionally consistent with staining in 82% of regions for Model A and 69% for Model B.

For M3, with patterned friction heterogeneity, both models showed larger deviations. Model A's packing density error rose to 11.6 points (15.9 to +16.7), and regional void fraction errors to 11.3 points. Model B's M3 packing density error was 14.8 points (19.3 to +20.5), and regional void fraction errors were 14.1 points. Persistent equatorial voids aligned with high-friction arcs in 76% of Model A predictions and 58% of Model B predictions. Discrepancies were largest in scenarios with adjacent low- and high-friction patches where stick–slip transitions produced metastable configurations.

Sensitivity analyses indicated that a ±0.05 change in coil–wall kinetic friction coefficient altered M1 packing density by ±1.2 points in Model A and ±1.6 points in Model B; the same change in M2 altered post-settling density by ±2.3 and ±2.8 points, respectively, while in M3 the effect was ±3.1 and ±3.8 points. Varying sac compliance within the measured range (±0.2 kPa1) changed M1 density by ±0.9 points for Model A and ±0.7 for Model B, suggesting that friction uncertainty dominates density error over the compliance range tested.

Bland–Altman plots and error histograms are available in the supplementary materials. Here we note that error distributions were approximately symmetric for M1 and slightly right-skewed for M2 and M3, reflecting overprediction of compaction in some cases. Correlation of model error with sac volume was weak (R2 < 0.2), whereas correlation with neck diameter was moderate in M2 and M3 (R2 0.4), indicating an opportunity to condition predictions on neck geometry to reduce bias.

Rigor and breadth of output comparison — gradation.

a.
Qualitative comparison without metrics.

b.
Single-case quantitative comparison; no uncertainty.

c.
Multiple cases with basic error metrics; no stratification by mechanism.

d.
Multiple cases per mechanism with error metrics and plots.

e.
Multiple cases per mechanism with uncertainty and sensitivity analysis.

14. Factor-by-Factor Scores by Model and Mechanism

This section summarizes, for Model A, the assigned ESS scores for each credibility thread across the three mechanisms. The basis statements distill the key quantitative evidence already presented and allow at-a-glance identification of strengths and gaps that should inform use. Scores reflect the aggregated

Figure 3: Representative overlay of simulated coil configuration (Model A) and segmented micro-CT coil skeleton for a 350 mm3 sac after 5 minutes of pulsation, showing dome and equator partitions and regional void fractions.

judgment that balances direct model–data agreement, numerical checks, and the alignment of test conditions and outputs with the intended decision. Scores at the lower end call for caution and for supplementing predictions with sensitivity bounds until additional evidence is generated.

15. Discussion

The present assessment demonstrates that explicit-dynamics models can furnish decision-useful estimates of coil packing density and regional voids within the early procedural horizon, with the degree of trust dependent on both the model form and the mechanism under consideration. The explicit beam–contact formulation (Model A) exhibits stronger alignment with measurements, particularly for initial deployment and early pulsatile compacting, than the reduced-order surrogate (Model B). The differences are consistent with the capacity of Model A to represent localized coil curvature changes and sharp friction transitions at the wall, which are central to first-loop seating and subsequent stick–slip rearrangements. Model B's speed and simplicity remain valuable for rapid scenario screening and for sensitivity exploration, but its lower fidelity in M2 and M3 indicates that additional conservatism or calibration may be required when using its outputs to influence device selection or procedural plans.

When contrasted against high-speed fluoroscopy reconstructions and CT-based sac fills, both models achieved less than 5 percentage points absolute error in packing density for M1 (2.8 for Model A, 5.6 for Model B), increased to 7–10 points for M2 (7.8 and 10.1), and 11–15 points for M3 (11.6 and 14.8). Regional void fraction errors exhibited the same ordering: 3.9–6.7 points for M1, 7.2–9.5 points for M2, and 11.3–14.1 points for M3. These results suggest that explicit-dynamics approaches can support the immediate decision of whether to place an additional coil after initial deployment with modest uncertainty, while caution is warranted when friction heterogeneity is suspected to dominate geometry evolution.

For both models, the structural representation was sufficient to resolve large-deformation coil mechanics and wall distensibility to reproduce observed intrasaccular geometry and surface engagement with deviations at or below 10 percentage points for M1 and within the low double-digits for M2, but the fidelity was strained in M3 where abrupt spatial variations in friction produced metastable configurations that magnified discrepancies.

The predominance of friction uncertainty in determining prediction errors is notable. Sensitivity analyses showed that modest changes in kinetic friction coefficients (±0.05) altered post-settling packing density by 2–3 points and M3 densities by 3–4 points. In contrast, similar perturbations in sac compliance changed density by less than one point in M1 and by under two points in M2. This hierarchy of sensitivities echoes clinical experience: packing targets are harder to achieve reliably when the sac wall becomes patchily adhesive. Our friction patterns were designed to emulate nascent thrombus effects through spatial modulation of surface energy, but real thrombus involves time-dependent biochemistry and three-dimensional growth that could further accentuate heterogeneity.

The numerical checks lend confidence that the observed errors primarily originate in model form and inputs rather than in coding defects. The manufactured beam/contact tests, sliding ring benchmarks, and round-trip energy error results (0.7% for Model A and 1.9% for Model B) indicate that integration and contact resolution can be trusted at the production discretizations. The small numerical error relative to experimental scatter in packing density (standard deviation 2.2 points across repeats) suggests that efforts to improve predictions should target friction characterization and model form in M3 rather than additional numerical refinement for the present question.

From a user perspective, the explicit beam–contact model is computationally heavier, averaging 38 minutes per simulation for M1 and 65 minutes for M2 on a 32-core workstation, whereas the mass–spring surrogate averages 4 minutes for M1 and 9 minutes for M2. The speed advantage of Model B enables parametric sweeps before committing to more detailed runs, and it could support training scenarios where approximate pack evolution suffices. Conversely, when a case presents delicate neck geometry or suspected wall heterogeneity, the increased fidelity of Model A may justify its runtime.

Our evidence also highlights the limits of laboratory validation. The benchtop phantoms capture sac compliance and pulsation faithfully, and these are the dominant mechanisms in minutes-long horizons; however, they do not capture hemostatic factors that modulate friction dynamically or the microstructural evolution at coil–wall contacts as thrombus matures. The experiments also simplify catheter dynamics, which in complex anatomies can influence first-coil seating. Nevertheless, we found that bracketing environmental variables and maintaining traceable calibrations greatly reduced avenues for model–data disagreement that are unrelated to the physics modeled, thereby focusing attention on the contest between model form and unmeasured inputs.

Looking forward, the most efficient path to higher credibility lies in two threads. First, systematic friction mapping at the coil–wall interface under physiologic wetting and biochemical conditions would reduce the dominant source of uncertainty identified in our sensitivity analyses and M3 comparisons. Measurements that provide spatially resolved friction for relevant materials and hydration states would allow the models to accept inputs with narrower priors and to represent state changes more accurately. Second, targeted validation of void fraction dynamics under spatially varying wall adhesion, perhaps using engineered surfaces with programmable friction patches and in-situ imaging of coil motion, would directly address the mechanism where the models showed the largest discrepancies.

This study's dual-model, tri-mechanism design provides a template for risk-informed credibility building in other contact-rich intravascular interventions. By structuring evidence around mechanisms and not merely global goodness-of-fit, and by keeping the decision in view when selecting outputs and tests, modelers can provide clinicians and regulators with nuanced, transparent support for computational predictions in planning and design.

Overall sufficiency of evidence for intended decisions — gradation.

a. Evidence insufficient; model not recommended for any decision impact.

b. Evidence weak; model may inform background understanding only.

c. Evidence moderate; model may inform low-stakes decisions with caveats.

d. Evidence strong for some mechanisms; use with caution for others.

e. Evidence strong across mechanisms relevant to decision.

16. Conclusions

This dual-model, tri-mechanism assessment under ASME V&V 40 offers a transparent, quantitative picture of where explicit-dynamics coil deployment models can be trusted to inform device design and procedural planning, and where additional evidence is needed. The explicit beam–contact model provides strong evidence for estimating packing density and regional voids immediately after deployment and after short-term pulsatile settling, while the reduced-order mass–spring surrogate provides moderate evidence for these same mechanisms and excels in speed for rapid screening. Both models encounter their steepest challenge when spatially heterogeneous friction acts as the dominant driver of post-deployment configuration; here, measured discrepancies rise into the low-to-mid teens in density percentage points, calling for targeted friction characterization and validation. Numerical checks indicate that coding errors and integration artifacts are well below experimental scatter at production settings, focusing future improvement efforts on input fidelity and, for the surrogate, on enhanced local contact representation. The factor-by-factor scores and the specific quantitative bases provided here can guide users in selecting the appropriate model and in understanding the limits of reliance under the mechanisms and conditions tested.

References

[1] J. Jankowski, S. Jankowski, T. Quintero, Flow rate effects on flow split in computational models, Int. J. Numer. Methods 112 (2010) 259–483.

[2] L. Nakamura, R. Rasmussen, G. Yilmaz, R. Bergman, P. Bergman, Element type effects on outlet velocity in computational models, Med. Eng. Phys. 149 (2005) 268–412.

[3] A. Vasquez, A. Vasquez, Bone density effects on peak force in computational models, Comput. Methods Biomech. 91 (2016) 65–755.

[4] B. Jankowski, K. Eriksen, W. Lindqvist, R. Quintero, Bone density effects on shear stress in computational models, J. Biomech. Eng. 109 (2018) 376–520.

[5] T. Duarte, N. Duarte, R. Pettersson, Cement mantle effects on flow split in computational models, Int. J. Numer. Methods 68 (2020) 123–640.

[6] W. Wojcik, P. Bergman, Yield stress effects on shear stress in computational models, Ann. Biomed. Eng. 38 (2006) 336–717.

[7] E. Duarte, H. Jankowski, E. Yilmaz, Gap size effects on head displacement in computational models, Comput. Fluids 121 (2013) 115–722.

[8] N. Duarte, W. Okonkwo, E. Yilmaz, Contact stiffness effects on flow split in computational models, Ann. Biomed. Eng. 98 (2022) 120–742.

[9] C. Vasquez, A. Silva, W. Bergman, L. Castellano, W. Nakamura, Damping effects on head displacement in computational models, J. Verif. Valid. Uncertain. 146 (2010) 71–465.

[10] R. Wojcik, J. Ueda, G. Ueda, H. Vasquez, Mesh density effects on chest compression in computational models, J. Biomech. Eng. 75 (2013) 264–499.

[11] D. Fontaine, S. Silva, Contact stiffness effects on cortical strain in computational models, J. Biomech. Eng. 65 (2022) 343–640.

[12] L. Okonkwo, C. Jankowski, D. Zetterberg, S. Okonkwo, Bone density effects on contact pressure in computational models, Int. J. Numer. Methods 58 (2022) 282–496.

[13] K. Castellano, A. Kaur, J. Eriksen, L. Abbott, J. Castellano, Load path effects on cortical strain in computational models, J. Biomech. Eng. 145 (2021) 274–666.

[14] B. Okonkwo, M. Fontaine, J. Moreau, D. Vasquez, Time step effects on fatigue life in computational models, Med. Eng. Phys. 56 (2006) 131–880.

[15] M. Eriksen, E. Rasmussen, A. Moreau, C. Halvorsen, G. Quintero, Element type effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 95 (2019) 343–894.

[16] P. Castellano, C. Yilmaz, L. Jankowski, B. Yilmaz, Solver tolerance effects on hemolysis index in computational models, J. Biomech. Eng. 26 (2018) 163–748.

[17] J. Grigoryan, M. Kaur, Load path effects on stress range in computational models, J. Mech. Behav. Biomed. Mater. 144 (2009) 107–649.

[18] T. Pettersson, L. Rasmussen, B. Yilmaz, S. Ishikawa, F. Jankowski, Mesh density effects on rib deflection in computational models, Med. Eng. Phys. 36 (2013) 136–798.

[19] M. Yilmaz, T. Silva, N. Halvorsen, Friction effects on contact pressure in computational models, Int. J. Numer. Methods 69 (2015) 68–407.

[20] R. Silva, S. Moreau, S. Moreau, K. Castellano, Time step effects on head displacement in computational models, Ann. Biomed. Eng. 136 (2010) 17–743.

[21] A. Lindqvist, R. Ishikawa, Solver tolerance effects on chest compression in computational models, J. Verif. Valid. Uncertain. 111 (2019) 173–482.

[22] T. Zetterberg, E. Duarte, A. Kaur, Element type effects on shear stress in computational models, J. Verif. Valid. Uncertain. 50 (2014) 269–853.

[23] N. Fontaine, C. Jankowski, M. Vasquez, Contact stiffness effects on cortical strain in computational models, Ann. Biomed. Eng. 109 (2014) 299–781.

[24] T. Jankowski, G. Vasquez, K. Yilmaz, G. Castellano, J. Zetterberg, Load path effects on shear stress in computational models, Comput. Fluids 44 (2014) 98–807.

[25] D. Grigoryan, S. Ishikawa, M. Lindqvist, Time step effects on contact pressure in computational models, Comput. Methods Biomech. 59 (2012) 108–698.

[26] N. Kaur, T. Moreau, Damping effects on chest compression in computational models, Med. Eng. Phys. 82 (2014) 266–650.

[27] B. Nakamura, N. Silva, B. Wojcik, Friction effects on rib deflection in computational models, Med. Eng. Phys. 60 (2012) 45–725.

[28] B. Lindqvist, K. Yilmaz, Gap size effects on outlet velocity in computational models, Int. J. Numer. Methods 49 (2019) 280–778.

[29] G. Kaur, N. Silva, G. Pettersson, E. Abbott, Gap size effects on chest compression in computational models, J. Biomech. Eng. 42 (2013) 106–438.

[30] P. Nakamura, E. Quintero, M. Duarte, P. Quintero, T. Yilmaz, Inlet profile effects on peak force in computational models, Med. Eng. Phys. 96 (2012) 188–783.

[31] F. Ueda, T. Abbott, L. Eriksen, Flow rate effects on outlet velocity in computational models, J. Biomech. Eng. 132 (2010) 240–669.

[32] H. Grigoryan, J. Tanaka, W. Fontaine, N. Grigoryan, Boundary condition effects on fatigue life in computational models, Comput. Fluids 57 (2020) 155–465.

[33] L. Ueda, K. Ishikawa, P. Nakamura, Solver tolerance effects on cortical strain in computational models, Med. Eng. Phys. 24 (2011) 140–481.

[34] B. Rasmussen, L. Nakamura, Mesh density effects on head displacement in computational models, Int. J. Numer. Methods 88 (2018) 129–681.

Table 1: Credibility assessment summary for Model A (Explicit Beam–Contact Deployment) across three mechanisms using the 0–5 Evidence Strength Scale.

Credibility factor	Level	Basis
Model form - Model A - M1 deployment into a compliant sac	4	Beam and shell forms expressed large-deformation coil–wall contact with friction; sensitivity matched measured trends (e.g., 20% EI raised density by 2.1 points).
Numerical code verification - Model A - M1 deployment into a compliant sac	4	Manufactured beam/contact tests and pendular benchmark yielded 0.7% energy error at 0.9 µs timestep; penetration under 2% wire diameter.
Output comparison - Model A - M1 deployment into a compliant sac	4	Absolute packing density error 2.8 points (limits 4.1 to +4.7); regional void fraction error 3.9 points.
Relevance of the quantities of interest - Model A - M1 deployment into a compliant sac	5	Computed fill, regional voids, and contact area matched bench observables and map directly to procedural decisions.
Relevance of the validation activities to the COU - Model A - M1 deployment into a compliant sac	4	Bench reproduced deployment workflow, sac compliance 0.6–1.4 kPa1, and pulsation 90/140 mmHg at 1.2 Hz used clinically.
Test conditions - Model A - M1 deployment into a compliant sac	4	Environment controlled at 37 ± 0.2 °C, viscosity 3.2 mPa·s; catheter feed 0.6–1.2 mm/s with traceable calibrations.
Model form - Model A - M2 early post-deployment settling	4	Pressure-driven shell dynamics and frictional contact supported compaction and rearrangement observed after 5 min pulsation.
Numerical code verification - Model A - M2 early post-deployment settling	4	Long-step stability preserved; round-trip benchmark error 0.7% indicates small integration/contact errors relative to data scatter.
Output comparison - Model A - M2 early post-deployment settling	3	Packing density error 7.8 points; regional void fraction error 7.2 points; 82% regional contact change agreement with staining.
Relevance of the quantities of interest - Model A - M2 early post-deployment settling	5	Post-settle fill and regional voids directly support the decision to add coils during the same session.
Relevance of the validation activities to the COU - Model A - M2 early post-deployment settling	3	Five-minute pulsation captures early compaction but not prolonged biologic remodeling; still aligned for immediate decisions.
Test conditions - Model A - M2 early post-deployment settling	4	Pulsation 90/140 mmHg at 1.2 Hz verified; sac wall motion logged to confirm volume oscillation amplitude.
Model form - Model A - M3 friction heterogeneity due to nascent thrombus	3	Node-level spatially varying friction supported; transitions captured but metastability around sharp gradients remained.
Numerical code verification - Model	4	Contact benchmarking maintained <2% penetration across

Table 2: Credibility assessment summary for Model B (Explicit Mass–Spring Surrogate) across three mechanisms using the 0–5 Evidence Strength Scale.

Credibility factor	Level	Basis
Model form - Model B - M1 deployment into a compliant sac	3	Mass–spring–damper chain reproduces bulk compaction; local loop fidelity limited but M1 trends captured.
Numerical code verification - Model B - M1 deployment into a compliant sac	3	Round-trip benchmark energy error 1.9% at 3 µs timestep; residual penetration 3% wire diameter.
Output comparison - Model B - M1 deployment into a compliant sac	3	Absolute packing density error 5.6 points (limits 8.2 to +8.9); regional void fraction error 6.7 points.
Relevance of the quantities of interest - Model B - M1 deployment into a compliant sac	5	Fill, regional voids, and contact area correspond to bench observables and procedural triggers.
Relevance of the validation activities to the COU - Model B - M1 deployment into a compliant sac	4	Deployment workflow, sac compliance 0.6–1.4 kPa1, and physiologic pulsation reproduced in bench tests.
Test conditions - Model B - M1 deployment into a compliant sac	4	Environment 37 ± 0.2 °C, viscosity 3.2 mPa·s, and feed 0.6–1.2 mm/s with traceable calibrations.
Model form - Model B - M2 early post-deployment settling	3	Quasi-static wall compliance and penalty contact express compaction; details of rearrangement smeared.
Numerical code verification - Model B - M2 early post-deployment settling	3	Integrator phase error bounded; 1.9% energy error indicates numerics small vs. experimental scatter.
Output comparison - Model B - M2 early post-deployment settling	2	Packing density error 10.1 points; regional void fraction error 9.5 points; 69% regional contact change agreement.
Relevance of the quantities of interest - Model B - M2 early post-deployment settling	5	Post-settle fill and voids remain the operative quantities for deciding on further coiling.
Relevance of the validation activities to the COU - Model B - M2 early post-deployment settling	3	Five-minute pulsation appropriate for early settling decision; longer-term biology excluded by design.
Test conditions - Model B - M2 early post-deployment settling	4	Pulsation 90/140 mmHg at 1.2 Hz confirmed; sac wall motion tracked to verify oscillation amplitude.
Model form - Model B - M3 friction heterogeneity due to nascent thrombus	2	Segment-level friction state maps capture trends but smear sharp µ transitions; metastability underrepresented.
Numerical code verification - Model B - M3 friction heterogeneity due to nascent thrombus	3	Contact regularization stable across mixed-µ patches with 3% penetration; no defects in tested configs.
Output comparison - Model B - M3 friction heterogeneity	1	Packing density error 14.8 points; regional void fraction error 14.1 points; 58% align-