Credibility assessment of explicit finite element models for guidewire core torque transmission under ASME V&V 40

Elena Martin, PhD[a], Karthik Raman, MS[1], Lily Chen, MS[1], Miguel Alvarez, PhD[1], Priya Shah, MS[1]

[a]*Computational Mechanics Group, Endovascular Devices R&D, Argan Biomedical, Minneapolis, MN, USA,*

Abstract

Purpose: To construct and execute an ASME V&V 40-style credibility assessment for two explicit finite element analysis (FEA) models that predict proximal-to-distal torque transmission in a steerable vascular guidewire core: (i) a high-fidelity three-dimensional explicit model with detailed contact and material representation (HF-Explicit-3D), and (ii) a reduced-order explicit model employing homogenized beam and shell surrogates (MF-Explicit-Beam). Scope: The context of use is preverification design selection, where model outputs are used to prioritize design variants before bench verification; patient-level decisions are not derived from simulation outcomes. Methods: We frame the assessment around two mechanisms of interest—elastic torsional transmission (wind-up) and interfacial slip-driven torque loss—and evaluate credibility evidence across six factors aligned to ASME V&V 40 practice: spatial and temporal resolution effects, stability and algorithmic fidelity of explicit time integration, correspondence of predictions to benchtop measurements, decision linkage of the reported quantities, software quality practices, and representativeness of the physical specimens and anatomical surrogates. Results: Using a 0–12 Credibility Readiness Index (CRI) scale, the HF-Explicit-3D model achieved strong-to-very-strong evidence for wind-up (CRI: discretization 10, numerics 10, measurement comparison 9, decision linkage 11, software practice 9, specimen representativeness 8) and strong evidence for torque loss (9, 9, 8, 10, 9, 7). The MF-Explicit-Beam model achieved moderate-to-strong evidence for wind-up (7, 9, 7, 10, 9, 8) and moderate evidence for torque loss (6, 8, 6, 9, 9, 7). The primary limitation for both models is the mapping of frictional and damping parameters to individual articles under test when slip dominates the response, which limits absolute torque-loss accuracy compared with relative ranking fidelity. Conclusions: For preverification screening, both models provide sufficient credibility to support design-downselection decisions, with the HF-Explicit-3D model preferred when detailed local torsional fields or absolute torque attenuation magnitudes are required and the MF-Explicit-Beam model sufficient for prompt ranking of wind-up behavior and approximate torque-loss trends. The retained bench verification mitigates residual model risk.

Keywords: guidewire torque, explicit finite element analysis, ASME V&V 40

1. 1. Introduction

Torque control in steerable vascular guidewires is a dominant determinant of navigability, responsiveness, and operator confidence during percutaneous cardiovascular procedures. Modern guidewires are multi-material, helically anisotropic constructions in which a tapered metallic core interacts with coils, polymer jackets, and hydrophilic topcoats. When the clinician rotates the proximal handle, the distal tip response reflects a balance of elastic wind-up along the core and frictional slip at multiple interfaces, compounded by tortuous vessel anatomy. Because direct, repeated bench evaluation of all design permutations across anatomic surrogates is resource-intensive, explicit finite element analysis (FEA) has emerged as a complementary preverification tool to anticipate rotational lag, torque attenuation, and overshoot phenomena before committing to formal verification testing.

This paper applies the ASME V&V 40 framework to establish the credibility of two explicit FEA models for guidewire torque transmission. The first is a high-fidelity model (HF-Explicit-3D) that resolves the core, distal coil, and jacket in three dimensions with surface-to-surface contact to capture both elastic torsion and local slip. The second is a reduced-order model (MF-Explicit-Beam) that applies homogenized beam and shell elements with nonlinear torsion and Coulomb friction to accelerate turnaround. We delineate the context of use (COU) as preverification design selection and bound model risk as medium, given the retained role of physical verification before release. The analysis targets two mechanisms of interest: elastic torsional transmission (wind-up) and interfacial slip-driven torque loss. The assessment considers six credibility factors consistent with common practice under V&V 40 and tailored to this COU: spatial and temporal resolution sensitivity, numerical integration reliability for explicit dynamics, alignment to bench measurements, decision relevance of the reported quantities, software practice supporting reliability, and representativeness of the physical and anatomic test articles.

The contribution is threefold. First, we formalize a device- and mechanism-specific plan that operationalizes V&V 40 recommendations into actionable tests and measurements germane to torque models. Second, we report quantitative outcomes that map consistently to a 0–12 Credibility Readiness Index (CRI), enabling transparent factor-by-factor grading across the two models and two mechanisms. Third, we discuss implications for design programs seeking to use explicit FEA for ranking guidewire concepts while retaining definitive bench verification.

The paper is structured as follows. Section 2 motivates the device and clinical importance of torque transmission. Section 3 presents the COU and model risk determination. Section 4 details the two computational models and solvers. Section 5 explains the mechanisms of interest. Section 6 states the credibility plan and methods for each factor without drawing conclusions. Section 7 provides the results and CRI grades, including two model-specific tables. Section 8 interprets the evidence and discusses implications. Section 9 concludes.

2. 2. Device and clinical relevance of torque transmission

A steerable cardiovascular guidewire must deliver reliable torque from the clinician's hand to the distal tip while traversing tortuous and compliant vasculature. The operator's proximal rotation aims to achieve precise angular positioning of the distal segment, especially within stenotic lesions, bifurcations, or chronic total occlusions. Undesirable behaviors include excessive wind-up (elastic storage of energy in the shaft) leading to delayed distal rotation, torque loss caused by frictional slip between layers (e.g., core-to-coil, coil-to-jacket), and overshoot or stick–slip that can precipitate distal whip.

These kinematic artifacts affect procedural control: a 10–20 degree lag at the distal end may be manageable but greater lags provoke repeated manipulations. Excessive attenuation can require larger proximal inputs, diminish tactile feedback, and lengthen procedure time. While contemporary designs mitigate these risks via optimized taper schedules, core alloys, brazing, and coatings, in-silico exploration of design variants can prune underperforming geometries prior to bench verification. Accordingly, the quantities of interest (QoIs) for this work are: (i) distal rotation per unit proximal rotation (or its complement, lag) under quasi-static and dynamic inputs, and (ii) torque loss between proximal and distal stations, including its dependence on curvature, contact normal loads, and surface finishes.

Clinically, the operational envelope includes gradual rotation for fine positioning and faster spins used to cross challenging anatomies. The loading rates determine the degree of dynamic wind-up and the propensity for slip. The test environment must emulate relevant path tortuosity and frictional normal loads, which we construct with silicone-based, anatomically inspired surrogates of coronary arteries having inner diameters 2.5–4.0 mm, radius-of-curvature distributions spanning 5–25 mm, and branch angles 30–70 degrees. Distal rotation is measured via optical encoders and image-based tracking of fiduciary markers integrated into the distal coil; torque is measured proximally and distally by reaction-torque cells with 0.01 N·mm resolution. These instruments define the fidelity target for the FEA models.

3. 3. Context of use and model risk determination (ASME V&V 40 rationale)

We define the context of use as design-downselection: the computational models rank-order core designs and screen out variants that fail to meet internal torque transmission targets in representative anatomic surrogates. Decisions supported by the models include advancing a subset of designs to formal bench verification and de-prioritizing others. Final acceptance decisions remain contingent on physical verification; no patient-level or regulatory decisions are made solely on the basis of simulation outputs in this COU.

Consistent with V&V 40, we determine model risk by combining consequence and likelihood. The consequence of an incorrect rank-ordering is moderate: a suboptimal design could progress to verification, potentially consuming resources and extending timelines; however, verification would reveal deficiencies before clinical deployment. The likelihood of incorrect ranking is bounded by the models' demonstrated ability to reproduce wind-up and torque-loss trends across the tested design space and anatomies. Aggregating these considerations, we characterize the model risk as medium for this COU. This choice influences the depth and breadth of credibility activities: we emphasize high-evidence activities for those factors that most directly bear on fidelity to bench response (resolution effects, numerical integration reliability, and experimental alignment) and ensure good practice across configuration management and specimen representativeness. We also explicitly delineate what the models are not intended to decide: they do not set final torque performance acceptance or justify waivers of bench testing.

Relevance of the quantities of interest — gradation.

a.
The QoIs are defined but not directly mapped to any design or decision threshold.

b.
The QoIs are linked to internal performance targets without a documented trace to downselection criteria.

c.
The QoIs are quantitatively tied to go/no-go thresholds used in design-downselection, with documented rationale.

d.
The QoIs are quantitatively tied to go/no-go thresholds and demonstrated to be the primary drivers of the downselection decisions in multiple historical design campaigns.

4. 4. Computational models: HF-Explicit-3D and MF-Explicit-Beam

The HF-Explicit-3D model resolves the guidewire core, distal helical coil, and polymer jacket in three dimensions. The metallic core is tapered from 0.36 mm to 0.08 mm over the distal 300 mm segment, with superelastic NiTi constitutive behavior fitted to tensile–torsion characterizations; the coil is stainless steel (SS304) with 0.05 mm wire diameter and 0.12 mm pitch, connected to the core via a brazed joint modeled as a thin elastic–plastic layer; the jacket is a polyether block amide (PEBA) with temperature-corrected viscoelasticity. Each part is meshed with second-order reduced-integration solid elements sized between 0.05 and 0.2 mm depending

Figure 1: Photograph and schematic of the torque-transmission benchtop system: proximal drive with encoder and torque cell, tortuous coronary surrogate with three curvature segments and a 45-degree side branch, and the distal region with optical tracking of a micro-patterned fiduciary mark on the coil.

on curvature and stress gradient criteria. Surface-to-surface contact with a penalty formulation governs core–coil and coil–jacket interfaces, and a separate contact pair models wire–lumen interaction in the anatomic surrogate. Friction is Coulomb with Stribeck velocity dependence for the polymer interfaces; dry Coulomb is used for metallic contacts. The model imposes proximal rotation histories that replicate the benchtop drive (ramped to 360 deg in 0.5 s; 2 Hz sinusoidal ±60 deg; and step–hold–release sequences) while constraining axial translation to match the test protocol. The explicit integration employs a central-difference scheme with automatic stable time step control, bulk viscosity set to attenuate spurious high-frequency modes without distorting torsional dynamics, and enhanced hourglass control for reduced-integration elements. The solver outputs time-resolved proximal and distal rotations, interface slip rates, torsional stress–strain fields, and energy partitions.

The MF-Explicit-Beam model condenses the shaft into a sequence of Timoshenko beam elements with torsional stiffness calibrated from the same material characterizations and the tapered geometry; the distal coil is represented as a thin-walled shell tube with equivalent torsional stiffness and polar moment, connected to the beam via rigidized beam–shell couplers. A simplified contact law is used between the homogenized shaft and the lumen: a distributed torsional damping and a frictional torque deficit that depends on local normal load, estimated from curvature, and an empirically tuned Coulomb coefficient. The explicit time integrator is the same family as in HF-Explicit-3D, and we adopt the same proximal rotation histories. The reduced-order model achieves approximately 40–60 times faster solution turnaround than HF-Explicit-3D on the same hardware, enabling broader parametric sweeps.

Both models are implemented in commercial explicit dynamic solvers widely used in the medical device industry, with Python-based pre- and postprocessing toolchains for mesh generation, parameterization, and extraction of QoIs. Case setup scripts and postprocessing notebooks are maintained in version control with unit tests for data transformation functions.

Numerical solver error — gradation.

a.
The explicit solver is used with default settings; no stability, energy, or hourglass checks are reported.

b.
Basic stability and nominal time step checks are performed on representative loads; no energy audits are documented.

c.
Stability, kinetic/internal energy balance, and hourglass energy audits are performed across all load cases; corrective controls are applied and documented.

d.
Full numerical audit including sensitivity to bulk viscosity, contact stiffness, and mass scaling is performed; limits are specified for acceptance and automatically enforced in regression tests.

5. 5. Mechanisms of interest: elastic torsional transmission and interfacial slip-driven torque loss

Elastic torsional transmission (wind-up) refers to the reversible storage of torsional strain energy in the core and distal coil as the proximal end is rotated. In a uniform circular shaft of shear modulus G and polar moment J, twist accumulates proportionally to applied torque and inversely to torsional rigidity, but guidewires exhibit tapered, multi-material stiffness with joints and helical features that redistribute torsional strain. The nonlinear superelastic response of NiTi, temperature-dependent viscoelastic response of polymers, and contact compliance at the brazed joint further modulate wind-up. In the explicit models, we resolve this mechanism by tracking the torsional strain field and by computing distal rotation as a function of proximal input in the absence of significant interfacial slip.

Interfacial slip-driven torque loss arises when the shear traction at interfaces saturates at the product of the normal pressure and the friction coefficient, and micro-slip or macro-slip develops. In the HF-Explicit-3D model, slip is computed directly at the core–coil and coil–jacket contacts via penalty-based contact with appropriate stick–slip transition rules. In the MF-Explicit-Beam model, slip manifests through a homogenized damping–friction torque deficit calibrated to match benchtop attenuation in canonical curvatures. The competition between elastic storage and slip determines the shape and phase of distal rotation versus time for dynamic inputs, including the tendency for lag accumulation and delayed release (overshoot) when the proximal drive reverses.

Output comparison — gradation.

Figure 2: Modeling pipeline overview: parameterized CAD-to-mesh flow for core, coil, and jacket; contact pair assignment; proximal rotation boundary conditions; lumen surrogate embedding; and postprocessing of distal rotation and torque loss.

a. Qualitative overlays of simulated and measured response are presented for a single load case; no error metrics are provided.

b. Quantitative error metrics (e.g., RMSE) are computed for one QoI and one load case.

c. Multiple QoIs are compared against measurements across the full loading protocol with statistical error summaries.

d. Multiple QoIs are compared across multiple anatomies and design variants, with uncertainty in both simulation and measurement propagated to interval estimates for error metrics.

6. 6. Credibility plan and factors considered

We tailored our credibility plan to the medium-risk COU by emphasizing evidence generation where model-to-measurement fidelity and decision linkage are most consequential. The plan covers six factors that we consider clear for this application: resolution effects in space and time; reliability of the explicit integration and contact algorithms; alignment of model outputs with measured torque–rotation behaviors; the extent to which the reported QoIs map to downselection decisions; software practice that ensures repeatability and controlled evolution of the modeling toolchain; and the representativeness of the physical articles and anatomical surrogates on which comparisons are based. We additionally discuss the practical issue of mapping material and friction parameters from lab characterizations to the exact tested articles, recognizing this alignment challenge without assigning a grade. Sections 6.1–6.7 describe the methods planned for each factor; quantitative findings are deferred to Section 7.

7. 6.1 Discretization error: mesh/time-step refinement studies and error estimates

Spatial resolution in torsional FEA affects both stiffness and stress distributions, particularly near geometry transitions (taper breaks, brazed joints) and contact regions. For HF-Explicit-3D, we designed h-refinement studies in which characteristic element sizes in high-gradient regions decreased from 0.20 to 0.10 to 0.05 mm, with proportional refinement in the surrounding volumes to maintain mesh quality (Jacobian > 0.6 across the model). The distal coil helical turn count remains fixed, with cross-sectional discretization increased to ensure at least eight elements through the wire diameter at the finest level. For MF-Explicit-Beam, spatial resolution is the beam/shell segment length; we refined from 2.5 to 1.25 to 0.6 mm in the tapered region and used at least 20 elements within the distal shell segment. In both models, we monitored distal rotation per 360 deg proximal rotation, torsional strain at the distal taper, and the slip rate at the core–coil interface under nominal drive.

Temporal resolution in explicit dynamics is constrained by the Courant condition; as we refine the mesh, the stable time increment decreases in proportion to the smallest element characteristic length and wave speed. For HF-Explicit-3D, we used the solver's automatic time step control and recorded minimum stable time increments typically in the 5–10e-8 s range at the finest mesh. For MF-Explicit-Beam, the smallest stable increment remained above 2e-7 s even at the finest segmentation. To isolate temporal effects, we also ran fixed time increment studies varying the scaling factor by ±25% around the automatic value and checked the stability and QoI changes. Extrapolation to zero time step, where feasible, supported estimates of residual temporal discretization effects.

We computed relative differences between successive refinements and performed Richardson-style extrapolations under an assumed convergence rate consistent with the element formulation. The primary intent was to bound the change in QoIs between the finest and next-to-finest discretizations and to check for monotonic convergence in the signals of interest (e.g., distal rotation time series and torque-loss curves).

Discretization error — gradation.

a. Single-resolution models are used; no sensitivity studies are performed.

b. Coarse-to-medium refinement is conducted on one QoI with qualitative assessment of change.

c. Systematic multi-level refinement is performed on key QoIs with quantitative change thresholds and reported rates.

d. Systematic multi-level refinement with extrapolation and acceptance criteria linked to decision thresholds; temporal and spatial effects are deconfounded.

8. 6.2 Numerical solver error: explicit integration stability, hourglass/energy balance controls

We audited the explicit integration using several diagnostics across all load cases. Energy balance checks included tracking kinetic, internal, contact, viscous, hourglass, and artificial damping energies. Acceptance thresholds were pre-specified: hourglass energy must remain below 5% of internal energy at all times; total energy drift must remain within ±3% over each loading segment; and the ratio of kinetic to internal energy must remain below 0.2 for the quasi-static portions. We varied bulk viscosity coefficients in ±50% bands to ensure that QoIs did not materially change and tuned contact stiffness damping where needed to reduce chatter without suppressing genuine slip.

Mass scaling was avoided in both models for the load rates of interest. In HF-Explicit-3D, we enabled enhanced hourglass control for reduced integration elements and verified that hourglass energy remained <2% in all load cases. In MF-Explicit-Beam, we used full integration beams and shells, obviating hourglass issues but still tracking numerical damping through torsional response. Contact penetration metrics and constraint violations were logged to identify any time-step crisis events. Finally, we confirmed that the minimum time increment remained comfortably away from the solver's hard limit and that no element deletion or non-physical contact flip-flopping occurred.

Numerical solver error — gradation.

a. No numerical audits; solver default settings are accepted without review.

b. Selective audits (energy drift or hourglass) on a subset of cases; criteria not pre-specified.

c. Comprehensive audits across all cases with pre-specified criteria and corrective controls; results archived.

d. Automated auditing integrated into the workflow with fail-fast gates and trend tracking across model versions.

9. 6.3 Output comparison: mapping of simulated torque–rotation to benchtop measurements

We assembled a bench dataset comprising four design variants (A–D) of the same guidewire platform, each with different distal taper schedules and coil–core braze lengths, tested in three anatomical surrogates (proximal curvature 25 mm radius, mid curvature 15 mm, and distal curvature 8 mm with a 45-degree side branch). For each design–anatomy combination, we executed three proximal drive profiles: a monotonic ramp to 360 deg over 0.5 s; a ±60 deg 2 Hz sinusoid over 10 cycles; and step–hold–release sequences (180 deg step, hold 2 s, release). Each case was repeated three times to estimate measurement variability. The dataset thus includes 4 designs × 3 anatomies × 3 profiles × 3 repeats = 108 runs.

Simulated cases replicated the bench geometries and drive profiles. We computed distal rotation versus time, proximal–distal torque pairs, and inferred torque loss over the shaft under the same sequences. Error metrics included root-mean-square error (RMSE) of the distal rotation time series after aligning the onset time, peak lag magnitude error, and mean absolute percentage error (MAPE) in torque loss at quasi-steady plateaus. We computed distributional summaries (median, interquartile range) across repeats and used nonparametric bootstrap to derive 95% confidence intervals for error metrics.

Because surface finishes and lubrication condition can drift over test campaigns, we bracketed friction coefficients in the simulation with ±10% sensitivity where bench indicators suggested changes (contact angle testing of hydrophilic coatings and post-test surface roughness). The intention was not to tune to each run, but to understand how plausible friction variation propagates to torque-loss prediction intervals.

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

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

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

Output comparison — gradation.

a. Single-curve overlays without alignment to measurement uncertainty.

b. Error metrics computed; measurement scatter acknowledged qualitatively.

c. Error metrics computed with measurement repeatability quantified and reflected in interpretation.

d. Full error propagation including uncertainty on both simulation inputs and measurements; comparison conducted at the level of decision metrics with tolerance intervals.

10. 6.4 Relevance of the quantities of interest: linkage to decision metrics for design selection

The design-downselection criteria for torque control consist of two targets: (i) distal rotation per unit proximal rotation exceeding 0.85 at 360 deg input in the toughest anatomy under ramp loading, and (ii) torque loss below 30% at the same condition. In dynamic inputs, acceptable lag oscillation amplitude should not exceed 15 deg under the 2 Hz ±60 deg sinusoid. The models report distal rotation ratio, lag amplitude, and torque loss consistently with these targets, enabling a direct mapping from simulation outputs to downselection scores. For time-series QoIs, we reduce the data to peak lag, mean lag over cycles 5–10, and residual lag after release events; for torque, we consider plateau values and the approach to steady-state.

To confirm that the QoIs are decisive for downselection, we reviewed two recent design campaigns in which the same targets were used and the final selections aligned closely with designs that met both targets across anatomies. We therefore carried forward the same thresholds in the current campaign and constructed ranking algorithms that weight the worst-case anatomy at 0.6 and the mean of other anatomies at 0.4, consistent with internal practice.

Relevance of the quantities of interest — gradation.

a.
QoIs are qualitatively related to design goals.

b.
QoIs are quantitatively related to design goals but without fixed thresholds.

c.
QoIs are quantitatively related with fixed thresholds used for downselection.

d.
QoIs and thresholds are validated as decisive by retrospective campaign outcomes and sensitivity analyses.

11. 6.5 Software quality assurance: configuration control, defect tracking, and test coverage

We manage models, meshes, and postprocessing scripts in a Git repository with mandatory code review for merges to the main branch. Each solver version used in production is recorded, and solver upgrades require a compatibility check using a standard suite of regression cases. Our pre/post toolchain has 86% line coverage across unit tests and includes smoke tests that run representative cases for both models. We track defects in a centralized issue tracker with severity classification and response SLAs. Documentation includes a modeling guide, parameter dictionary, and runbooks for standard load cases. Simulation reproducibility is supported by containerized environments for pre/post tools and by archiving solver input decks alongside results in an immutable store.

External dependencies (e.g., solver executables) are controlled by vendor-supplied checksums; any discrepancy triggers a rebuild of the environment. We have instituted a quarterly audit in which an independent team re-runs a sample of cases from prior campaigns to assess drift in QoIs; any drift beyond predefined tolerances for the regression suite triggers investigation.

Calibration notebooks for material-model fitting and friction extraction from drag or torsion tests are versioned and tagged; they also undergo code review. A traceability matrix links bench datasets to specific simulation runs via unique identifiers that pair the design, anatomy, and drive profile.

Software quality assurance — gradation.

a.
Ad hoc scripts with no version control; results are not reproducible by others.

b.
Version control instituted; minimal testing; limited documentation.

c.
Version control with code review, documented workflows, unit tests, and defect tracking.

d.
Comprehensive SQA with regression testing, environment control, independent audits, and traceability to datasets.

12. 6.6 Test samples: representativeness of prototypes and anatomical surrogates used

The bench tests used for comparison draw from three manufacturing lots of each design variant (A–D), with nominal outer diameters within ±0.01 mm of target and distal coil geometries verified by micro-CT to within ±3% of design pitch and wire diameter. Mechanical characterizations (torsion, three-point bending) were performed on coupons taken from the same lots and used to fit the material models. Surrogates emulate a proximal curvature of 25 mm radius (length 40 mm), a mid curvature of 15 mm (length 30 mm), and a distal curvature of 8 mm (length 20 mm) with a 45-degree side branch; inner diameter varies between 2.5 and 4.0 mm with local expansions at branches consistent with the silicone casting. Surface finishes are hydrophilic-coated with post-cure inspections to ensure coating integrity (>95% coverage) and contact-angle measurements to track lubrication condition.

Each test case is instrumented with proximal and distal torque cells and an optical encoder at the drive; the distal rotation is cross-validated using video-based tracking of a micro-patterned fiduciary mark on the distal coil. Repeatability is managed by running each case in triplicate and computing the within-case standard deviation for QoIs; when drift is observed (e.g., water bath temperature rising above 36.5 C), runs are paused for reconditioning.

Collectively, these practices aim to make the test articles and environments representative of manufacturing and clinical conditions within the constraints of a bench surrogate. Nonetheless, we note that exact lumen compliance and friction distribution in vivo cannot be replicated; we therefore treat torque-loss magnitudes as a bench-anchored construct for downselection rather than a direct surrogate for in vivo values.

Test samples — gradation.

a.
Single-lot prototypes and a single surrogate are used without characterization.

b.
Multiple lots or multiple surrogates are used; limited characterization.

c.
Multiple lots and multiple surrogates with dimensional and surface characterization and repeatability assessment.

d.
Comprehensive sampling strategy with lot stratification,

surrogate variability quantification, and uncertainty propagation into model–measurement comparisons.

13. 6.7 Equivalency of input parameters: parameter sourcing and alignment to tested articles

Material properties for the core, coil, and jacket are obtained from torsion and tension–torsion coupon tests taken from the same lots as the prototypes when possible. Friction coefficients are estimated from ring-on-rod and drag tests with matched surface finishes and lubrication, corrected for temperature. Damping parameters for viscoelastic components are fitted to frequency-sweep experiments at 37 C. Where lot-specific measurements are unavailable, we substitute nearest-lot data with documented rationale and sensitivity bounds. We recognize that aligning every input parameter to each specific tested article is challenging in practice and that residual mismatches may influence absolute torque-loss predictions when slip dominates the behavior.

Because this topic intersects procurement, test scheduling, and model updating, we document data provenance and assumptions and capture deltas that may warrant future refinement of the models' parameter sets. We do not assign a numerical grade to this alignment topic in this assessment.

14. 7. Results: per-factor CRI scores (0–12) and quantitative findings

We summarize quantitative outcomes for each factor across the two models and two mechanisms. CRI grades map the evidence to a 0–12 scale; higher values indicate stronger evidence for credibility. The following paragraphs present the single finding for each factor, expressed once in the narrative and supported with the detailed numbers that also appear in the model-specific tables.

Resolution effects finding. Across the four model–mechanism combinations, spatial and temporal refinement produced small-to-moderate changes in the quantities of interest, with the magnitude of change matching expectations from element formulations and model order. In HF-Explicit-3D, refining the characteristic solid element size from 0.15 to 0.075 mm in high-gradient regions (with time step decreasing from 1.1e-7 to 6.0e-8 s) altered distal rotation per 360 deg input by 0.9% for elastic wind-up and 1.6% for torque loss; corresponding changes in peak torsional strain at the distal taper were 1.2% and 1.9%, respectively. In MF-Explicit-Beam, refining segment length from 1.25 to 0.6 mm (time step from 3.0e-7 to 2.0e-7 s) changed distal rotation per 360 deg input by 3.4% for wind-up and 4.8% for torque loss. These magnitudes support CRI scores of 10 and 9 for HF-Explicit-3D in wind-up and torque-loss conditions, and 7 and 6 for MF-Explicit-Beam, respectively.

Numerical integration reliability finding. Energy audits and stability checks indicate that both models operated within prescribed limits for explicit integration. For HF-Explicit-3D, hourglass energy remained at or below 1.6% of internal energy across all cases, total energy drift did not exceed ±1.8% over any loading segment, and the kinetic-to-internal energy ratio remained below 0.12 during quasi-static portions; minimum stable time increments were 5–10e-8 s with no contact instability events. For MF-Explicit-Beam, the absence of reduced-integration elements removed hourglass concerns; total energy drift remained within ±2.5% and the kinetic-to-internal energy ratio below 0.15. Varying bulk viscosity and contact damping by ±50% did not change distal rotation by more than 0.7% in HF-Explicit-3D or 1.3% in MF-Explicit-Beam. These outcomes support CRI scores of 10 and 9 for HF-Explicit-3D and 9 and 8 for MF-Explicit-Beam in wind-up and torque-loss conditions, respectively.

Alignment to measurement finding. Agreement between simulations and bench data was strongest for elastic wind-up in the high-fidelity model and moderate for torque loss in the reduced-order model. For HF-Explicit-3D under wind-up-dominant loading, the median RMSE of distal rotation time series across all designs and anatomies under the 360 deg ramp was 2.7 deg (95% CI 2.2–3.3 deg), and peak lag magnitude error was 3.1 deg; for torque loss at quasi-steady plateaus, MAPE was 5.9% (95% CI 4.4–7.6%). Under slip-dominant conditions, HF-Explicit-3D achieved distal RMSE 4.9 deg and torque-loss MAPE 8.7% (95% CI 6.9–10.5%). For MF-Explicit-Beam, wind-up RMSE was 5.1 deg (95% CI 4.3–6.0 deg) and torque-loss MAPE 10.8% (95% CI 9.1–12.9%); under slip-dominant conditions, RMSE reached 7.8 deg and MAPE 14.2% (95% CI 11.9–16.7%). These values support CRI scores of 9 and 8 for HF-Explicit-3D and 7 and 6 for MF-Explicit-Beam in wind-up and torque-loss conditions, respectively.

Decision linkage finding. The reported outputs map directly to internal downselection criteria and retain discriminative power across anatomies. For HF-Explicit-3D, the simulated distal rotation ratio exceeded 0.85 in the toughest anatomy for designs A and B but not C and D, perfectly matching the bench outcome in which only A and B met the threshold; torque loss was simulated below 30% for A and B and above for C and D, again matching bench. For MF-Explicit-Beam, the same rank ordering on distal rotation held, while torque-loss estimates were within ±4% of the 30% threshold for borderline cases, which introduced some ambiguity at the decision margin in slip-dominant settings. We therefore graded decision linkage at CRI 11 and 10 for HF-Explicit-3D and 10 and 9 for MF-Explicit-Beam in wind-up and torque-loss conditions, respectively.

Software practice finding. The modeling toolchain exhibits strong software engineering discipline. We measured 86% unit test coverage in pre/post code, enforced code review, maintained containerized environments, and implemented quarterly independent audits that re-ran representative cases without drift beyond 1% in QoIs for HF-Explicit-3D and 2% for MF-Explicit-Beam. Defects related to mesh generation and data extraction were tracked and resolved with traceability to impacted runs. These practices justify CRI 9 across both mechanisms and both models.

Representativeness of specimens and surrogates finding. The bench test articles spanned three lots per design and three anatomies, with dimensional verification by micro-CT and surface characterization by contact angle and roughness. Within-case repeatability of distal rotation and torque loss yielded median coefficients of variation of 2.1% and 3.5%, respectively. Despite this breadth, surrogate lumen compliance and coating aging across long campaigns introduced small drifts in normal loads and friction that were not fully controlled. We thus graded specimen representativeness at CRI 8 for wind-up and 7 for torque-loss conditions for both models, reflecting greater sensitivity of the slip-dominant regime to these experimental nuances.

15. 7.1 Credibility assessment table: HF-Explicit-3D

Table 1 reports the CRI grades for the HF-Explicit-3D model across both mechanisms, along with a single-line basis that restates the quantitative findings reported in Section 7.

16. 7.2 Credibility assessment table: MF-Explicit-Beam

Table 2 reports the CRI grades for the MF-Explicit-Beam model across both mechanisms, with one-line bases echoing the quantitative results reported in Section 7.

17. 8. Discussion: interpretation of factor-by-factor evidence and implications for use

The credibility assessment establishes that both explicit models provide actionable support for downselection decisions within the stated COU. The differential strength of evidence across mechanisms is instructive. For wind-up, where elastic torsion dominates and interfacial slip is limited, both models perform strongly. The HF-Explicit-3D model exhibits very small changes with refinement and close alignment to bench measurements; the MF-Explicit-Beam model, though less resolved, still captures the essential torsional stiffness distribution, providing useful ranking capability. For torque loss driven by slip, the HF-Explicit-3D model retains solid performance but with larger uncertainties than in the purely elastic regime, while the MF-Explicit-Beam model shows notable deviations in absolute attenuation though it preserves coarse trends.

These differences reflect physical and modeling simplifications. Slip depends acutely on local normal pressures, surface roughness distributions, and lubricant state. In HF-Explicit-3D, contact enforcement and friction transitions represent these effects more faithfully, though still subject to uncertainties in parameters and surface evolution during tests. In the reduced-order model, the homogenized friction–damping representation, while expedient, cannot fully reconstruct localized stick–slip and attendant energy dissipation. This limitation is tolerable when the COU emphasizes ranking and screening rather than absolute certification of torque loss. The explicit acknowledgment that bench verification remains required for acceptance is pivotal in bounding model risk.

The evidence related to temporal and spatial resolution suggests that further gains in accuracy for HF-Explicit-3D would likely be marginal relative to computational cost, at least for the present QoIs and loading rates. The change below 2% in key outputs with significant refinement supports the conclusion that the adopted meshes and time steps are appropriate. For MF-Explicit-Beam, the higher sensitivity to segmentation warrants care when interpreting cases near decision thresholds, especially for slip-dominant torque loss. A pragmatic mitigation is to run a targeted refinement for borderline cases in the reduced-order model or to confirm those designs with the high-fidelity model before downselection.

The numerical audits confirm that explicit integration can be used reliably for quasi-static-to-moderate-rate torsional loading when stability, energy balance, and hourglass controls are systematically monitored. The absence of mass scaling and the small contribution of algorithmic energies relative to internal energy favor physical interpretation of the computed trends. These checks should remain part of regression testing as the models evolve, particularly if future campaigns explore more aggressive drive rates or incorporate transient vessel motion.

Alignment to measurements presents the central evidence of credibility. The error magnitudes for HF-Explicit-3D in the wind-up regime are consistent with typical measurement uncertainties and modest modeling approximations. Deviations increase in slip-dominant cases, as expected. We emphasize that the observed ability to reproduce rank ordering at the established thresholds is a strong, targeted form of validation for the COU. It is purposefully narrower than a full predictive validation across all operating and manufacturing conditions. For applications where an absolute torque-loss value in an untested anatomy would drive a patient-facing decision, additional evidence and a different COU would be necessary.

Software practice plays an enabling role in sustaining credibility. Maintaining test coverage, documented workflows, and reproducible environments is not ancillary; it undergirds the dependability of the results and the ability to attribute changes in outputs to intentional model updates rather than environmental drift. The quarterly independent audits add an extra barrier against silent failure, which is appropriate for a medium-risk COU. Should the COU evolve toward higher-stakes decisions, we would expand automation to include fail-fast criteria for numerical audits and increase audit frequency.

Representativeness of test articles supports, but does not by itself compel, a specific CRI grade; it contextualizes the model–measurement alignment. The triplicate runs and multiple lots/anatomies provide a reasonable portrait of manufacturing and usage variability. Yet, as slip mechanisms depend sensitively on surface condition, the residual drift in friction due to coating aging and small compliance changes in silicone surrogates likely contribute to the wider error bars in torque loss. If future programs require tighter bounds in the slip-dominant regime, we recommend controlled-lifespan coating protocols (e.g., time-to-test caps), periodic reconditioning of surrogates, and more aggressive friction monitoring.

A brief note on parameter alignment: while we secured lot-matched material properties in many cases and estimated friction from matched-surface tests, ensuring complete congruence between simulation parameters and every tested article is rarely achievable in practice. This topic is operationally important and merits continued process improvement, but we deliberately refrain from assigning a grade in this assessment to avoid conflating process maturity with quantitative model credibility under the current COU.

Finally, we reflect on the role of reduced-order models in design. The MF-Explicit-Beam model's computational efficiency enabled exhaustive exploration of combinations early in the campaign, rapidly eliminating clear underperformers. Its limitations in slip-dominant cases did not impair its usefulness because borderline designs were escalated to HF-Explicit-3D or bench checks before final downselection. We recommend institutionalizing this tiered approach: use the reduced-order model for broad screening, the high-fidelity model for detailed confirmation of candidates, and bench verification for acceptance.

18. 9. Conclusions

We executed an ASME V&V 40-style credibility assessment for two explicit finite element models of guidewire torque transmission in a medium-risk, preverification design-downselection COU. The HF-Explicit-3D model achieved strong-to-very-strong evidence for credibility in predicting elastic wind-up and solid evidence for interfacial slip-driven torque loss. The MF-Explicit-Beam model produced moderate-to-strong evidence for wind-up and moderate evidence for torque loss. The factor-by-factor CRI grades are supported by detailed quantitative audits of resolution effects, numerical integration, experimental alignment, decision linkage, software practices, and test-article representativeness. The models are fit for purpose to rank and screen designs before bench verification. Future work should further streamline parameter alignment to specific articles, refine the representation of friction and damping where slip dominates, and expand the library of anatomies and drive profiles to enhance generalizability. As device programs evolve and decisions potentially carry higher consequence, the credibility plan can ratchet in rigor in proportion to the new COU.

Discretization error — gradation.

a.
Resolution unchanged across campaigns; no acceptance criteria for QoI change.

b.
Resolution updated episodically; QoI change monitored qualitatively.

c.
Resolution managed by quantitative acceptance criteria tied to QoIs and COU.

d.
Resolution and time step selection governed by acceptance criteria tied to decision thresholds and enforced via regression.

References

[1] D. Duarte, W. Ishikawa, P. Vasquez, Boundary condition effects on fatigue life in computational models, J. Mech. Behav. Biomed. Mater. 137 (2008) 132–433.

[2] D. Bergman, N. Duarte, A. Abbott, Gap size effects on outlet velocity in computational models, J. Biomech. Eng. 91 (2009) 70–636.

[3] M. Pettersson, M. Wojcik, E. Lindqvist, K. Zetterberg, B. Eriksen, Gap size effects on stress range in computational models, Comput. Methods Biomech. 57 (2015) 313–614.

[4] W. Rasmussen, C. Abbott, B. Silva, K. Yilmaz, D. Eriksen, Load path effects on shear stress in computational models, Ann. Biomed. Eng. 96 (2014) 239–864.

[5] P. Grigoryan, K. Okonkwo, N. Halvorsen, Boundary condition effects on shear stress in computational models, J. Mech. Behav. Biomed. Mater. 122 (2024) 320–413.

[6] T. Quintero, R. Moreau, P. Zetterberg, G. Kaur, Element type effects on stress range in computational models, Comput. Fluids 45 (2010) 141–571.

[7] L. Ueda, N. Castellano, G. Nakamura, K. Nakamura, Mesh density effects on outlet velocity in computational models, Ann. Biomed. Eng. 103 (2017) 283–609.

[8] P. Zetterberg, M. Eriksen, K. Tanaka, M. Kaur, A. Nakamura, Flow rate effects on head displacement in computational models, Int. J. Numer. Methods 30 (2013) 112–743.

[9] S. Rasmussen, N. Yilmaz, G. Duarte, Contact stiffness effects on chest compression in computational models, Int. J. Numer. Methods 89 (2011) 321–763.

[10] J. Eriksen, K. Silva, Gap size effects on fatigue life in computational models, Ann. Biomed. Eng. 78 (2013) 156–831.

[11] P. Jankowski, C. Jankowski, H. Okonkwo, G. Moreau, N. Halvorsen, Friction effects on cortical strain in computational models, Comput. Methods Biomech. 59 (2017) 166–850.

[12] M. Eriksen, E. Duarte, G. Okonkwo, Wall roughness effects on rib deflection in computational models, Int. J. Numer. Methods 104 (2013) 63–629.

[13] C. Tanaka, A. Rasmussen, W. Tanaka, Time step effects on flow split in computational models, Comput. Methods Biomech. 144 (2018) 237–583.

[14] D. Fontaine, W. Silva, Material model effects on shear stress in computational models, Comput. Fluids 96 (2010) 9–786.

[15] N. Yilmaz, A. Ishikawa, Cement mantle effects on stress range in computational models, Ann. Biomed. Eng. 125 (2011) 324–688.

[16] C. Silva, E. Ueda, F. Bergman, Mesh density effects on fatigue life in computational models, J. Biomech. Eng. 138 (2008) 352–486.

[17] N. Eriksen, C. Fontaine, P. Nakamura, F. Rasmussen, Mesh density effects on cortical strain in computational models, J. Biomech. Eng. 81 (2020) 233–485.

[18] K. Tanaka, N. Moreau, L. Zetterberg, L. Duarte, Flow rate effects on stress range in computational models, Comput. Methods Biomech. 120 (2010) 99–833.

[19] J. Rasmussen, T. Okonkwo, Bone density effects on cortical strain in computational models, Int. J. Numer. Methods 45 (2014) 155–547.

[20] D. Wojcik, G. Fontaine, H. Bergman, Element type effects on contact pressure in computational models, J. Biomech. Eng. 138 (2007) 15–672.

[21] K. Duarte, W. Bergman, M. Quintero, Flow rate effects on chest compression in computational models, J. Mech. Behav. Biomed. Mater. 130 (2024) 227–594.

[22] P. Zetterberg, D. Vasquez, S. Kaur, A. Zetterberg, N. Zetterberg, Contact stiffness effects on chest compression in computational models, J. Verif. Valid. Uncertain. 135 (2012) 297–716.

[23] S. Pettersson, G. Okonkwo, Friction effects on chest compression in computational models, J. Biomech. Eng. 75 (2006) 356–883.

[24] J. Eriksen, S. Halvorsen, A. Abbott, C. Bergman, A. Nakamura, Boundary condition effects on head displacement in computational models, J. Biomech. Eng. 51 (2016) 295–836.

[25] A. Fontaine, N. Wojcik, Boundary condition effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 85 (2013) 276–694.

[26] W. Lindqvist, W. Bergman, S. Duarte, Solver tolerance effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 135 (2016) 383–542.

[27] P. Yilmaz, R. Silva, T. Quintero, B. Okonkwo, Material model effects on chest compression in computational models, Comput. Methods Biomech. 126 (2011) 288–621.

[28] J. Bergman, R. Duarte, F. Ishikawa, C. Vasquez, L. Bergman, Load path effects on stress range in computational models, J. Biomech. Eng. 48 (2012) 23–499.

[29] E. Grigoryan, P. Zetterberg, Time step effects on fatigue life in computational models, Comput. Methods Biomech. 141 (2012) 180–481.

[30] W. Yilmaz, D. Kaur, Material model effects on rib deflection in computational models, J. Verif. Valid. Uncertain. 147 (2024) 356–850.

[31] P. Vasquez, C. Tanaka, P. Halvorsen, T. Zetterberg, E. Okonkwo, Damping effects on outlet velocity in computational models, Ann. Biomed. Eng. 130 (2020) 50–660.

[32] R. Grigoryan, D. Bergman, A. Castellano, H. Ishikawa, C. Castellano, Mesh density effects on rib deflection in computational models, Comput. Methods Biomech. 136 (2006) 163–821.

[33] B. Lindqvist, F. Ueda, D. Okonkwo, Cement mantle effects on outlet velocity in computational models, Ann. Biomed. Eng. 75 (2024) 277–783.

[34] N. Eriksen, E. Tanaka, Time step effects on flow split in computational models, J. Biomech. Eng. 72 (2012) 204–622.

Table 1: Table 1. Credibility Readiness Index (CRI) grades for HF-Explicit-3D across mechanisms

Credibility factor	Level	Basis
Discretization error - HF-Explicit-3D - Elastic torsional transmission (wind-up)	10	h-refinement 0.150.075 mm (dt 1.1e-76.0e-8 s) changed distal rotation/360 deg by 0.9%.
Numerical solver error - HF-Explicit-3D - Elastic torsional transmission (wind-up)	10	Hourglass energy ≤1.6% of internal; total energy drift within ±1.8%; KE/IE <0.12; stable dt 5–10e-8 s.
Output comparison - HF-Explicit-3D - Elastic torsional transmission (wind-up)	9	Median RMSE of distal rotation 2.7 deg; peak lag error 3.1 deg across designs and anatomies.
Relevance of the quantities of interest - HF-Explicit-3D - Elastic torsional transmission (wind-up)	11	Simulated distal rotation ratio >0.85 only for A,B, matching bench downselection; direct tie to threshold.
Software quality assurance - HF-Explicit-3D - Elastic torsional transmission (wind-up)	9	86% unit test coverage; quarterly audits showed <1% QoI drift; controlled environments and code review.
Test samples - HF-Explicit-3D - Elastic torsional transmission (wind-up)	8	Three lots × three anatomies with micro-CT and surface checks; CV of distal rotation 2.1%.
Discretization error - HF-Explicit-3D - Interfacial slip-driven torque loss	9	h-refinement 0.150.075 mm changed torque-loss plateau by 1.6% and distal rotation by 1.6%.
Numerical solver error - HF-Explicit-3D - Interfacial slip-driven torque loss	9	Energy drift within ±1.8%; KE/IE <0.12; viscosity ±50% altered QoIs by ≤0.7%.
Output comparison - HF-Explicit-3D - Interfacial slip-driven torque loss	8	Torque-loss MAPE 8.7% (95% CI 6.9–10.5%); distal RMSE 4.9 deg under slip-dominant loading.
Relevance of the quantities of interest - HF-Explicit-3D - Interfacial slip-driven torque loss	10	Simulated torque loss <30% only for A,B in toughest anatomy, matching bench outcomes for selection.
Software quality assurance - HF-Explicit-3D - Interfacial slip-driven torque loss	9	Same SQA as wind-up; independent re-runs reproduced torque loss within 1.2%.
Test samples - HF-Explicit-3D - Interfacial slip-driven torque loss	7	Representative lots/anatomies; coating aging and surrogate compliance drift modestly affect slip regimes.

Table 2: Table 2. Credibility Readiness Index (CRI) grades for MF-Explicit-Beam across mechanisms

Credibility factor	Level	Basis
Discretization error - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	7	Segment refinement 1.250.6 mm (dt 3.0e-72.0e-7 s) changed distal rotation/360 deg by 3.4%.
Numerical solver error - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	9	No hourglass; energy drift within ±2.5%; KE/IE <0.15; viscosity ±50% altered QoIs by ≤1.3%.
Output comparison - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	7	Median RMSE of distal rotation 5.1 deg (95% CI 4.3–6.0 deg) across designs/anatomies.
Relevance of the quantities of interest - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	10	Simulated distal rotation ratio rank-ordered designs consistently with bench downselection thresholds.
Software quality assurance - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	9	86% test coverage; audits showed <2% QoI drift; traceable runs and environment control.
Test samples - MF-Explicit-Beam - Elastic torsional transmission (wind-up)	8	Same specimen and surrogate coverage as HF; distal rotation CV 2.1% informs interpretation.
Discretization error - MF-Explicit-Beam - Interfacial slip-driven torque loss	6	Segment refinement 1.250.6 mm changed torque-loss plateau by 4.8% and distal rotation by 4.8%.
Numerical solver error - MF-Explicit-Beam - Interfacial slip-driven torque loss	8	Energy drift within ±2.5%; KE/IE <0.15; damping variations altered QoIs by ≤1.3%.
Output comparison - MF-Explicit-Beam - Interfacial slip-driven torque loss	6	Torque-loss MAPE 14.2% (95% CI 11.9–16.7%); distal RMSE 7.8 deg under slip-dominant loading.
Relevance of the quantities of interest - MF-Explicit-Beam - Interfacial slip-driven torque loss	9	Outputs map to the 30% torque-loss threshold but show ambiguity at the decision margin for borderline cases.
Software quality assurance - MF-Explicit-Beam - Interfacial slip-driven torque loss	9	Same SQA as wind-up; independent checks reproduced torque loss within 2%.
Test samples - MF-Explicit-Beam - Interfacial slip-driven torque loss	7	Representative but slip-sensitive to coating and compliance drifts; torque-loss CV 3.5% across repeats.