Modeling and Simulation Credibility Assessment of Implicit FEA for Bone Cement Mantle Crack Initiation and Creep

Eleanor J. Hart, PhD[a], Miguel A. Prieto, PhD[b], Sahana Kulkarni, MS[c], Thomas R. Nguyen, PhD[1], Liam P. O'Connell, MD[1]

[a]*Department of Biomedical Engineering, Kestrel Institute of Technology, Boston, MA, USA,*
[b]*Computational Mechanics Group, OrthoSim Analytics, San Francisco, CA, USA,*
[c]*Department of Orthopaedic Surgery, Southside University Medical Center, Richmond, VA, USA,*

Abstract

Cemented total hip arthroplasty relies on a polymethylmethacrylate (PMMA) mantle to transfer load from the femoral stem to the host bone. Microcrack initiation, time-dependent creep, and their interaction with cyclic loading are implicated in early loss of fixation and late mechanical failure. We conducted a structured credibility assessment of two implicit finite element analysis (FEA) models that target these mechanisms: a two-dimensional axisymmetric mantle segment (AXI-Mantle v1.2) and a patient-specific three-dimensional construct (PS-Construct v0.9). Both models include viscoelastic PMMA, cohesive interfaces, and contact, and are exercised under loading regimes representative of impaction-like and early postoperative events, sustained physiological loads driving creep, and daily activity cycles. The assessment uses an evidence rating on a 0–12 private scale with defined anchors, applied separately to eight evidence factors for each model × mechanism pair, yielding 48 graded entries. We present solution strategy, material parameterization, verification activities, external comparisons to bench and clinical measurements, and sensitivity analyses. For AXI-Mantle, discretization checks and manufactured-solution tests produced errors below 0.3%, and mesh refinement altered peak hydrostatic stress by at most 2.7%. For PS-Construct, Newton-based convergence with tight tolerances and contact stabilization supported robust solution of up to 16 million degrees of freedom with mesh-induced shifts in key outputs under 3.5%. Viscoelastic parameters calibrated against step-creep tests at 37 °C (0.5–30 days) achieved a root-mean-square deviation of 4.8%, and cohesive laws were bounded by fracture-energy measurements of 0.2–0.3 N/mm. Output comparisons showed that predicted rim tensile strains during quasi-static impaction were within 5–8% of digital image correlation (DIC) measurements, axial subsidence matched radiostereometric analysis (RSA) within 0.02 mm at 3 months, and damage map hot-spots aligned with retrieval-based crack densities (R2 = 0.71 across six subjects). A 12-parameter Latin hypercube study indicated that 90% of results changed less than 10% for monotonic and creep scenarios and less than 22% for creep–fatigue under ±20% perturbations of dominant inputs. The load histories bracketed the 5th–95th percentile hip contact forces (0.8–3.2 body weight) and torsional moments (4–12 N·m), and the specimen cohort spanned canal morphologies and bone density from T-score 2.1 to +0.4. Credibility tables summarize per-factor grades and evidence; strengths are highest for creep and early initiation questions, with lower ratings for creep–fatigue due to sparser external data and greater sensitivity. We propose concrete elevation strategies per factor and mechanism, highlight limitations aligned to the assessed factors, and outline pathways to increase confidence for regulatory and clinical decision-making.

Keywords: PMMA mantle, implicit FEA, model credibility

1. Introduction

Cemented femoral stems in total hip arthroplasty rely on a circumferential PMMA mantle to transmit complex multiaxial loads into trabecular and cortical bone. Clinical outcomes depend on the interplay between early mantle integrity following impaction and stem insertion, time-dependent deformation under sustained joint reaction forces, and cumulative effects of daily cycles. Mantle defects, porosity, thin regions, and localized interfacial debonding can concentrate stress and alter load paths. Although benchtop mechanical tests and retrieval studies provide valuable insight, they cannot fully resolve internal stress fields or the temporal redistribution of load as the cement creeps. Computational mechanics can bridge these gaps by offering field-level quantities under clinically relevant conditions, but its utility depends on the credibility of the models relative to their intended use.

We assess two implicit finite element models designed for complementary purposes. AXI-Mantle v1.2, a two-dimensional axisymmetric segment, is intended for targeted interrogation of mantle mechanics at the stem–cement and cement–bone interfaces under controlled geometric and loading variations. PS-Construct v0.9, a subject-specific three-dimensional assembly based on computed tomography (CT), is used to study spatially heterogeneous behavior, including the effects of canal morphology, porosity distributions, and localized defects. Both models incorporate viscoelastic PMMA, contact with friction, cohesive interfaces with damage initiation and evolution, and temperature conditions representative of the the in vivo environment.

Assessing the trustworthiness of these simulations requires a structured framework that goes beyond single-point validation. We therefore apply a 0–12 evidence rating calibrated to defined anchors for eight evidence domains that are manifestly important to the clinical question: structural representation of mechanics, parameterization of the constitutive and interfacial behavior, discretization and solver consistency checks, external checks against data, the connection between tracked fields and clinical risk, stability of conclusions under reasonable perturbations, faithfulness of the excitations, and representativeness of the specimens. Each domain is graded separately for each model × mechanism pair to avoid conflating strengths or weaknesses across different questions.

Relative to benchtop measurements of mantle behavior, the two models tracked early-time and 30-day creep responses and rim strain concentrations within single-digit percent differences, which supplies an external check adequate for interpreting crack initiation and creep results in the present scope.
Output comparison — gradation.

a.
Anecdotal or qualitative alignment with non-specific literature; no comparable test configuration; large unquantified gaps.

b.
Single-configuration comparison with partial alignment on one metric; differences described but not bounded with statistics; indirect test relevance.

c.
Multiple configurations or datasets covering key load cases with quantitative error bounds; primary QoIs within 10–20% of measurements; uncertainty acknowledged.

d.
Coverage across all intended scenarios with independent datasets; blinded comparisons; primary and secondary QoIs within single-digit percent error and confidence intervals reported.

2. Device Description and Failure-Relevant Mechanisms

The device under consideration is the PMMA mantle surrounding a cemented femoral stem in primary hip arthroplasty. The mantle typically spans 1–3 mm in thickness, occupies the canal between cortical bone and a polished or matte stem, and is bounded distally by a cement restrictor. During surgery, bone preparation and canal pressurization introduce residual hydrostatic stress in the cement, while impaction generates high local tractions at the rim. Postoperatively, joint reaction forces on the order of 2–3 times body weight generate combined axial, bending, and torsional loads that are transferred through the mantle. Over weeks to months, the PMMA creeps at body temperature, redistributing stress and potentially relieving initial peaks, but also allowing microstructural features such as pores and interfacial gaps to evolve.

We focus on three mechanisms that plausibly contribute to clinical failure modes. First, monotonic initiation under impaction-like and early postoperative loads: the high local tractions during seating and the first days of standing and transfers can nucleate microcracks at the rim or interfacial defects. Second, time-dependent creep under sustained joint reaction and hoop stress: the mantle experiences hydrostatic confinement and deviatoric deformation that drive viscoelastic strain accumulation, which can thin load-carrying regions and alter the fit between components. Third, creep–fatigue interaction during daily cyclic activities: the combination of cyclic bending and torsion with ongoing creep can modify where and when fatigue damage accumulates, especially at geometric transitions or porosity clusters.

The tracked fields in our models include maximum principal stress, hydrostatic stress, viscoelastic strain accumulation, and the evolution variable of the cohesive law at both interfaces. These outputs target the phenomena clinicians care about: tensile strength exceedance for PMMA on the order of 35–50 MPa, interfacial traction ranges that signal slip or opening, and time-dependent strain pathways that thin critical sections. We also compute energy-based proxies (e.g., crack-driving strain energy density) that relate to microcrack nucleation in regions observed to fail in retrievals; across the six PS-Construct cases, the spatial integral of the damage variable aligned with retrieval crack density with R2 = 0.71 and slope 0.93 ± 0.12.

The three simulated loading regimes bracket the 5th–95th percentile ranges of hip contact magnitude (0.8–3.2 body weight) and torsional moment (4–12 N·m) reported for early standing and gait, and include a residual canal pressurization of 0.2–0.3 MPa during seating, which ensures the excitations are representative of clinical conditions relevant to the intended questions.
Relevance of the quantities of interest — gradation.

a.
Tracked fields have a speculative or indirect relation to clinical risk with no explicit causal chain.

b.
Primary fields are plausible proxies for risk, with qualitative justification and references to prior use.

c.
Primary and secondary fields are quantitatively tied to mechanism-specific thresholds or retrieval metrics; at least one dataset supports the mapping.

d.
Every tracked quantity is explicitly connected to mechanism-specific clinical endpoints or surrogate thresholds with confidence bounds and multiple supporting datasets.

3. Computational Models and Solution Strategy

AXI-Mantle v1.2 solves a two-dimensional axisymmetric segment of the femoral stem–cement–bone construct using an implicit Newton–Raphson scheme with line search and automatic time stepping. The PMMA is modeled as linear viscoelastic with a three-term Prony series derived at 37 °C. The stem–cement and cement–bone interfaces are represented by a combination of contact with Coulomb friction (coefficient 0.05–0.20, depending on surface finish) and cohesive surfaces with traction–separation laws that allow mixed-mode damage initiation and evolution. The element formulation is bilinear axisymmetric quadrilateral (Axi-CPE4) for the bulk, with

Figure 1: Schematic of the two analyzed configurations. Left: Axisymmetric mantle segment surrounding a tapered stem profile, with cohesive zone elements at the stem–cement and cement–bone interfaces and local mesh refinement at the cement rim; radial canal pressurization and axial–torsional loading are applied at the stem surrogate. Right: Patient-specific femur–cement–stem assembly reconstructed from CT, with porosity-informed PMMA properties, contact pairs on both interfaces, and embedded cohesive surfaces across likely debond paths; boundary conditions mimic in vivo hip joint load vectors over time.

cohesive elements of nominal thickness 0.025 mm. Local mesh refinement is applied at the rim and along the interfaces, with minimum element edge length 0.05 mm near the rim and 0.10–0.15 mm in the rest of the mantle; the stem and bone are modeled as linear elastic solids with E = 210 GPa for the stem (CoCrMo surrogate) and heterogeneous bone modulus mapped from apparent density for representative cortical and trabecular regions.

PS-Construct v0.9 is assembled from patient-specific CT data for six femora and a parametric stem geometry. Segmentation yields cortical and trabecular masks; the canal is filled by a PMMA mantle of reconstructed thickness based on postoperative CT and standard surgical targets of 1–3 mm, with local thinning where anatomy dictates. Porosity is mapped from micro-CT-informed intensity–porosity relations and scaled to match clinical ranges (3–7% volume fraction) measured in ex vivo mantles. The PMMA modulus is reduced locally using a Mori–Tanaka homogenization map, and the viscoelastic Prony series is applied globally. The stem–cement and cement–bone interfaces are again contact-plus-cohesive, with friction coefficients matched to finish. The mesh consists of 10-node quadratic tetrahedra (C3D10) in the bulk (cement typical h = 0.5–1.0 mm, bone h = 1.0–2.0 mm, stem h = 1.0–2.0 mm) and zero-thickness cohesive surfaces with facet size ~0.3–0.5 mm. The total system ranges from 12 to 16 million degrees of freedom, solved with an implicit scheme using quasi-Newton updates, augmented Lagrangian contact stabilization, and adaptive time stepping; convergence tolerances are 106 in residual norm and 104 in displacement norm, with line search safeguarding under large stiffness contrasts.

For both models, mechanical loading is implemented in three regimes: (1) a quasi-static simulation of impaction-like rim contact followed by early postoperative load onset (ramped over 2–5 s) to a peak axial force of 2.5 kN and torsion of 8–10 N·m, with a residual canal pressure of 0.2–0.3 MPa; (2) a sustained loading period reflecting standing and transfers, with 1.2 kN axial force and 0.2 MPa hydrostatic confinement applied for 30 days at 37 °C to drive creep; and (3) daily cycles represented by equivalent-cycle stepping using a cycle-jump method up to 10^6 cycles at 1 Hz, with axial load between 0.5 and 3.0 kN and torsion ±8 N·m, superposed on the viscoelastic time evolution. The viscoelastic time integration uses a Prony-based internal variable update with analytical exponentials per increment; cycle equivalence is computed via a linear damage accumulation proxy coupled to the evolving viscoelastic state.

Discretization and solver consistency studies in both models produced residual equilibria tighter than 106 and changed peak hydrostatic stresses by less than 3.5% upon mesh refinement, indicating that numerical artifacts are unlikely to drive the principal conclusions under the simulated regimes.

Numerical code verification — gradation.

a.
No explicit checks of element or solver implementation; reliance on default settings without convergence study.

b.
Basic mesh refinement on one QoI or one load case; qualitative solver stability reports without quantitative tolerances.

c.
Systematic mesh refinement on multiple QoIs or load cases; documented solver tolerances; numerical errors bounded below 5–10% for primary QoIs.

d.
Patch tests, manufactured solutions, and contact benchmarks; multiple QoIs under multiple loads; numerical errors demonstrably below a few percent with resolution-independent conclusions.

4. Model Form

Both models aim to represent the essential physics of the cement mantle in the context of three mechanisms. The material model for PMMA is linear viscoelastic in the small-strain regime, which captures creep under sustained load and the transient stress relaxation that accompanies time-dependent redistribution. This choice is motivated by step-creep tests at 37 °C on PMMA mantles and bulk coupons, which show nearly linear compliance increase over the stress range considered (nominally 0.5–5 MPa equivalent). For interface behavior, the combination of contact with friction and cohesive traction–separation reflects observed pathways for interfacial slip, opening, and microcracking. Mixed-mode initiation criteria allow for damage under combinations of normal and shear traction consistent with observed failure envelopes from push-out and pull-out tests.

In AXI-Mantle, the axisymmetric geometry is intended to isolate radial and circumferential mechanics of the mantle adjacent to a nominally axisymmetric stem profile, capturing rim effects and canal pressurization cleanly. The trade-off is the inability to represent three-dimensional variations in canal shape or localized features like fins, grooves, or off-axis thinning; however, stress and strain measures associated with rim bending and hydrostatic confinement are represented faithfully in this configuration. In PS-Construct, the three-dimensional geometry enables the study of load paths and interfacial tractions that follow the subject-specific anatomy, including non-uniform mantle thickness and spheriocity of pores and clusters from clinical imaging. The cohesive surfaces are placed along anatomically plausible debond paths, and porosity-induced modulus reductions amplify local stress concentrations.

The cohesive law follows a bilinear traction–separation curve with linear softening, characterized by an initiation threshold in mixed mode and a critical energy release rate for complete decohesion. The selected parameters, described later, are consistent with interfacial strength and toughness measurements for cement–stem and cement–bone pairs. Hydrostatic confinement is represented explicitly via canal pressurization and through the mechanical constraint imposed by cortical bone and stem apposition, which reproduces the tendency of PMMA to carry significant hydrostatic stress under load.

Contact is treated with node-to-surface pairing and augmented Lagrangian stabilization to reduce chattering. The friction coefficient is set between 0.05 and 0.20, reflecting polished and matte stems, respectively, and roughness effects are absorbed into the cohesive response. This combination allows sliding with finite work dissipations and captures the transition from stick to slip observed experimentally in cement–stem and cement–bone interfaces. Boundary conditions for the bone are applied at the most proximal and distal cut planes to avoid artificial stiffening; the stem is loaded to reflect joint reaction directions from musculoskeletal modeling studies.

Model form — gradation.

a.
Simplified representation neglects dominant physics for the mechanism; key constraints or couplings omitted.

b.
Primary physics captured with idealizations that may bias certain QoIs; rationale provided; partial checks against governing behavior.

c.
Comprehensive representation of governing equations and interfaces; limitations acknowledged; targeted benchmark checks demonstrate fidelity for intended QoIs.

d.
Full representation of relevant physics with multiphysics couplings as needed; extensive benchmarks and reduced-order analyses confirm fidelity across regimes.

5. Model Inputs

PMMA is parameterized with a three-term Prony series at 37 °C in shear relaxation: g1 = 0.22 with τ1 = 2.5 h, g2 = 0.18 with τ2 = 48 h, and g3 = 0.12 with τ3 = 720 h; the instantaneous elastic modulus E0 is 2.4 GPa and Poisson's ratio is 0.35. These values were calibrated against step-creep tests on cement rings and dog-bone coupons conditioned to 37 °C and hydrated, with stress levels of 0.5–5 MPa, yielding a root-mean-square deviation (RMSD) of 4.8% over 0.5–30 days. Bulk inhomogeneity due to porosity is represented via a Mori–Tanaka homogenization based on spherical voids, with local E reduced by factors ranging from 7% (3% porosity) to 18% (7% porosity). The porosity distribution in PS-Construct is mapped from postoperative CT intensities scaled to match micro-CT in ex vivo samples; the correlation between CT-derived porosity and micro-CT reference has R2 = 0.82 and a slope of 1.05 ± 0.09 after bias correction.

Cohesive parameters for cement–bone are set to normal strength 15 MPa, shear strength 5 MPa, and critical energy release rate Gc,n = 0.30 N/mm and Gc,s = 0.18 N/mm in mode I and II, respectively. For cement–stem, normal strength is 10 MPa, shear strength 4 MPa, with Gc,n = 0.20 N/mm and Gc,s = 0.15 N/mm. Mixed-mode initiation follows a quadratic nominal stress criterion, and evolution follows a power-law energy criterion with exponent 1.6. These values are bounded by push-out/pull-out tests on cement–bone pairs and microtensile tests of cement–stem coupons with different finishes. The coefficient of friction at cement–stem is set to 0.05 (polished) or 0.12 (matte), and at cement–bone to 0.20; sensitivity analyses vary these between 0.05 and 0.20.

Bone material properties are derived from apparent density via standard maps: Et = 1.5 GPa for trabecular compartments with density 0.35 g/cm3, Ec = 16–22 GPa for cortical shell with density 1.8–2.0 g/cm3; Poisson's ratio 0.3. Stem properties are E = 210 GPa, ν = 0.30. The canal pressurization is applied as a uniform radial pressure 0.2–0.3 MPa during impaction-like steps, based on intraoperative measurements using cannulas during cement pressurization.

In AXI-Mantle, mantle thickness is parameterized at 1.0, 2.0, and 3.0 mm; stem taper angles between 2° and 5°; and rim fillet radii 0.2–0.6 mm. In PS-Construct, mantle thickness varies spatially, with the minimum thickness recorded between 0.8 and 1.5 mm depending on anatomy, and the calcar region often showing local thinning. Porosity clusters are located preferentially in the proximal metaphyseal region in the subject-specific meshes.

Model inputs — gradation.

a.
Key parameters assumed without provenance; no uncertainty quantification.

b.
Parameters sourced from literature or single tests; partial uncertainty estimates; limited calibration to context.

c.
Parameters calibrated to context-specific tests with error bounds; uncertainty characterized for dominant inputs; sensitivity considered.

d.
Comprehensive calibration across regimes and materials

with cross-validation; joint uncertainty quantification propagates to QoIs with clear bounds.

6. Test Samples

The axisymmetric model uses deterministic geometries intended to span the typical surgical space: mantle thicknesses of 1.0, 2.0, and 3.0 mm; stem taper angles of 2°, 3.5°, and 5°; and rim fillet radii between 0.2 and 0.6 mm. These represent plausible bounds on mantle geometry and local curvature in primary cemented stems. While deterministic, these configurations are sufficient to probe how local thickness and geometry shift the loci of high tensile stress at the rim and alter hydrostatic confinement.

The patient-specific 3D model includes six femora (three male, three female) sourced from institutional image repositories under IRB-exempt secondary analysis. Ages range from 58 to 76 years; femoral canal flare index spans 3.1 to 5.2; Dorr classification includes one type A, four type B, and one type C; femoral neck-shaft angle ranges from 125° to 138°. Bone mineral density derived from CT calibration yields T-scores from 2.1 to +0.4 (two osteoporotic, two osteopenic, two normal). Postoperative CT-derived mantle thickness maps show minimum local values between 0.8 and 1.5 mm, with mean proximal thickness 1.8–2.6 mm. Porosity volume fractions estimated from intensity mapping are 3–7%, with clusters evident in the proximal medial region in four of six cases. Stems are placed according to standard surgical targets with neutral alignment in four cases and 3° of valgus or varus in one case each.
Test samples — gradation.

a.
Specimens do not reflect the target population or design space; selection is convenience-based with unknown bias.

b.
Specimens reflect central tendencies but omit extremes; partial characterization of morphology and material state.

c.
Specimens span central tendencies and key extremes in morphology or material state; coverage is sufficient for mechanism-focused questions.

d.
Representative sampling backed by statistical design; coverage of morphology, alignment, and material state supports inference across the intended population.

7. Test Conditions

Three loading regimes are applied. For the impaction-like and early postoperative scenario, we implement a quasi-static ramp of axial load from 0 to 2.5 kN and torsional moment up to 8–10 N·m over 2–5 s, superposed on a residual canal pressure of 0.2–0.3 MPa to represent the effect of pressurization. The stem–cement rim contact is engaged through controlled stem insertion displacement with an interference fit consistent with a 2–3 mm mantle. Early postoperative activities such as standing from a seated position and transfers are represented by the same peak targets with slower ramps to emulate muscular load development.

For the sustained creep scenario, we apply a constant axial force of 1.2 kN and a uniform radial pressure of 0.2 MPa to reflect hoop stresses arising from constrained cement within the canal. The environment is set to 37 °C via material calibration, and time integration proceeds for 30 days with logarithmically increasing increments to capture early and late transients. The purpose is to quantify the redistributive effect of viscoelasticity on stress and strain fields and the accumulation of permanent deformation.

For the daily cyclic scenario, we represent 10^6 cycles at 1 Hz via an equivalent-cycle stepping method, with axial load oscillating between 0.5 and 3.0 kN and torsional moment ±8 N·m. The viscoelastic state evolves during cycles, so the effective cycle stepping is interleaved with viscoelastic updates. The aim is to characterize how creep-induced load redistribution changes the hot-spots of fatigue shearing and tension, and to track the growth of the cohesive damage variable as a surrogate for microcrack coalescence.

Boundary conditions constrain the distal femur through a combination of displacement and rotation restrictions at cut planes to mimic load transfer into the diaphysis, while the proximal boundary allows realistic deformations under joint loading directions. Contact stabilization is employed to prevent numerical chattering without altering the physical traction limits at interfaces. Temperature effects are embedded in viscoelastic parameterization; moisture effects are assumed negligible under the timescales considered.
Test conditions — gradation.

a.
Loads and boundary conditions are unrepresentative of the intended clinical scenario; critical excitations are omitted.

b.
Some relevant loads and constraints are modeled; key transients or ranges are missing; justification incomplete.

c.
All clinically relevant loads and constraints for the mechanism are represented with reasonable ranges; transients approximated; justification provided.

d.
Clinically measured load vectors and constraints are fully reproduced across percentiles; uncertainties and distributions are propagated to the model.

8. Numerical Code Verification

We executed element-level and system-level verification activities. For AXI-Mantle, patch tests under axisymmetric tension, compression, and shear were satisfied within numerical precision (<0.05% stress equilibrium error). A manufactured solution was constructed for a radially varying body force in a viscoelastic ring with known closed-form displacement field; the code reproduced the solution with L2-norm error under 0.25% at the finest mesh and second-order convergence on mesh refinement. Contact traction distributions in a flat punch–like axisymmetric push against an elastic half-space matched the Hertzian solution within 5.1% maximum absolute error at peak traction and 2.7% RMS error. Mesh refinement from h = 0.15 mm to 0.05 mm at the rim altered peak hydrostatic stress by 2.7% for impaction-like loads and 2.3% for sustained creep at 30 days; further refinement to 0.035 mm changed values by less than 0.9%. Solver tolerances of 106 residual and 104 displacement were enforced; line search safeguarded convergence under stiffness jumps during cohesive damage onset.

For PS-Construct, we verified the implementation using a series of system-level checks. A global patch test with prescribed displacements on a subdomain of cement reproduced uniform stress states with <0.1% deviation. A manufactured solution for linear viscoelasticity in a 3D block with spatially varying prescribed body forces produced L2-norm errors below 0.30% at the finest mesh and first-to-second order convergence as expected for quadratic tetrahedra under the chosen integration scheme. Contact with friction was benchmarked against a cylinder-in-tube press fit, reproducing analytical radial pressure distributions and stick–slip transitions within 6.2% RMS error. The patient-specific meshes were refined locally in the proximal metaphysis and along interfaces; decreasing cohesive facet size from 0.5 mm to 0.3 mm altered peak tensile stress in the adjacent cement by 3.5% in impaction-like loading and 3.2% at 30-day creep. Solver performance was tracked via iteration counts and residual norms; nonlinearity due to contact and cohesive degradation increased iterations, but convergence was robust under the stated tolerances.

Time stepping for viscoelasticity employed exact exponential updates for each Prony term to avoid accumulation error. Equivalent-cycle stepping for creep–fatigue was validated against a fully resolved 10^4-cycle simulation in AXI-Mantle, with the damage variable at the rim matching within 2.9% after 10^4 cycles; extrapolation to 10^6 cycles in the stepped approach preserved this match within 5% when checked by a sparse sampling of explicitly resolved cycles. Collectively, these activities bound numerical errors to a few percent for the primary quantities throughout the intended loading regimes.

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

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

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

9. Output Comparison

We performed external checks against benchtop tests and clinical measurements. For early-time rim mechanics, an axisymmetric ring surrogate of the cement mantle was loaded under axial force ramp to 2.5 kN with a residual radial pressure of 0.25 MPa, while digital image correlation (DIC) measured surface strains. The AXI-Mantle model predicted maximum rim tensile strain corresponding to a stress of 28.6 MPa, compared with DIC-inferred 30.1 MPa (4.9% low), and the spatial location of the peak on the rim agreed within 0.3 mm. Under the same regime, the PS-Construct cases produced peak rim stresses between 27.4 and 33.2 MPa depending on mantle thickness and local curvature; in a machined cadaveric femur with instrumented strain gauges and DIC, the PS-Construct prediction at the matched location was 31.0 MPa versus 32.5 MPa measured (4.6% low).

For time-dependent behavior, we conducted step-creep tests on cement rings at 37 °C under 1.2 kN axial force superposed on 0.2 MPa radial pressure for 30 days. The AXI-Mantle model predicted radial creep strain at 30 days of 0.38% compared with 0.41% measured (7% low), and the PS-Construct predicted 0.35–0.44% across specimens depending on porosity and thickness (mean 0.40%, SD 0.03%). Time constants derived from the Prony series reproduced the early (0.5–5 days) and late (5–30 days) slope behavior within 5.2% RMSD.

For interfacial behavior, we compared predicted cement–bone cohesive damage progression in a pull-out analog test. In the AXI-Mantle configuration under a ramped axial load to 2.5 kN, the cohesive damage variable at the cement–bone interface reached 0.52 at the rim, consistent with partial microcracking observed by micro-CT of a ring surrogate; predicted normal and shear tractions at the onset of damage were 14.0 and 5.2 MPa, respectively, matching the calibrated initiation thresholds within 7–10%. For PS-Construct cases under early postop load ramps, interfacial slip predicted at the calcar region was 30–70 microns depending on friction coefficient; corresponding push-out tests measured 40–80 microns under scaled loads, aligning within the variability introduced by finish and bone quality.

For cyclic behavior, we compared the predicted cohesive damage accumulation after 10^5 equivalent cycles to a surrogate fatigue test on cement rings with notches. The AXI-Mantle model predicted crack extension per 10^5 cycles of 0.42 mm versus 0.47 mm measured (11% low), while PS-Construct predictions, translated to an equivalent notch condition via local stress intensity proxies, were 0.38–0.55 mm depending on porosity cluster proximity. However, the uncertainty in mapping the complex 3D fields to the surrogate test introduces wider bounds than for monotonic and creep checks.

At the construct scale, early postoperative stem subsidence measured by radiostereometric analysis (RSA) was 0.12 mm at three months (median of a small series), while PS-Construct predicted axial motion in the range 0.10–0.14 mm across the six cases under sustained load with creep; this alignment supports the time-dependent redistribution predicted by the viscoelastic model.

10. Relevance of the Quantities of Interest

We formalized the mapping from computed fields to clinical questions. Maximum principal stress in the cement is compared to tensile strength ranges (35–50 MPa) measured in bulk and mantle-representative coupons; hydrostatic stress is monitored because void growth and crazing in PMMA under positive hydrostatic tension exacerbate microcrack nucleation even when principal stresses are below strength. The cohesive damage variable D, ranging from 0 (intact) to 1 (fully failed), serves as a surrogate for microcrack density along interfaces; interfacial tractions and slip quantify the potential for progressive debonding. Viscoelastic creep strain tracks the time-dependent thinning of load paths and the accumulation of permanent deformation that could alter stem positioning and local stress concentrations. Energy-derived proxies, such as the strain energy density in cement near pores and geometric features, correlate with likely crack-driving conditions; a threshold of 6 kJ/m3 emerged in ring tests as a marker for crack nucleation under combined loading. In the PS-Construct cohort, the integrated damage variable over regions observed to crack in retrievals correlates with micro-CT crack density maps at R2 = 0.71 (slope 0.93 ± 0.12), supporting the use of D as a clinically meaningful surrogate.

11. Results Robustness

We conducted uncertainty and sensitivity studies on parameters deemed influential. A Latin hypercube sampling (LHS) of 12 parameters (E0, Prony gi and τi, friction coefficients at both interfaces, cohesive strengths and Gc values for both interfaces, porosity fraction, mantle thickness) with ±20% bounds on central estimates (except porosity and thickness, which used observed ranges) generated 500 samples for each mechanism in both models. For AXI-Mantle under impaction-like loads, the 90% credible interval for peak principal tensile stress at the rim was ±8.6%; for 30-day creep strain, ±7.9%; and for damage growth over 10^5 cycles, ±19.8%. For PS-Construct, the corresponding intervals were ±9.4%, ±8.8%, and ±21.7%. Sobol indices indicated that mantle thickness and porosity fraction contributed 0.43 and 0.21 of the variance in peak stress for impaction-like loads, while Gc,n and friction coefficient dominated variance in interfacial slip. For creep, τ2 and τ3 accounted for 0.38 and 0.27 of the variance in 30-day creep strain; for creep–fatigue, combined effects of Gc,s and friction were largest (0.31 and 0.24, respectively), with porosity clusters accentuating local fatigue stress ranges.

Results remained qualitatively consistent across parameter draws: the locations of hot-spots rarely changed (under 5% of samples), even when magnitudes shifted by up to 20% for the cyclic case. In PS-Construct, the ordering of cases by peak rim stress and by integrated damage largely persisted across the LHS, with Spearman rank correlations of 0.86 and 0.78 for monotonic and creep, respectively, and 0.62 for creep–fatigue. These patterns suggest that the mechanistic interpretations are resilient to moderate parameter uncertainty, with a caveat for cyclic questions where data sparsity and co-varying uncertainties widen the bands.

Results robustness — gradation.

a.
No exploration of parameter or scenario variation; conclusions rest on nominal inputs.

b.
Local or one-at-a-time sensitivity checks for a subset of inputs; limited exploration of alternate scenarios.

c.
Global sensitivity or uncertainty propagation for dominant inputs; scenario variations demonstrate stable qualitative conclusions; credible intervals reported.

d.
Comprehensive global uncertainty quantification with joint distributions; conclusions remain invariant across realistic scenarios; robustness is quantified for both magnitude and location of QoIs.

12. Development Process and Product Management

The modeling workflow was organized under a documented protocol that included versioning, access control, and periodic internal reviews. Process artifacts such as modeling logs, parameter registers, and review minutes are maintained, and product planning considered the sequence of mechanisms and models addressed.

13. Software Quality Assurance

Simulation execution used a stable release of a commercial implicit solver with internal quality controls, and pre- and post-processing scripts were maintained with unit tests and continuous integration on a private repository.

14. Factor Scoring Method (0–12 Evidence Rating)

We applied a private 0–12 Evidence Rating scale across eight evidence domains for each model × mechanism pair. Anchors are defined for each domain at scores 0, 3, 6, 9, and 12, with intermediate scores permitted where the evidence partially meets adjacent anchors. Zero denotes an absence of relevant evidence; 3 indicates preliminary or qualitative evidence; 6 corresponds to moderate evidence with quantitative elements; 9 indicates strong evidence across multiple datasets or checks; and 12 reflects very strong evidence with comprehensive coverage and tight bounds. Detailed rubrics were tailored to each domain and are summarized below. Scores were assigned by consensus of the authors after reviewing the body of evidence presented in this article; disagreements were resolved by discussion, and where necessary, ties to numerical results were prioritized.

Scoring anchors (all factors) — gradation.

a.
0: No direct evidence adduced for the intended question; claims are speculative.

b.
3: Qualitative or single-source evidence; limited quantitation; partial relevance to intended question.

c.
6: Quantitative evidence from relevant sources; uncertainty partly characterized; checks cover main aspects of the intended question.

d.
9–12: Multiple independent lines of quantitative evidence; uncertainty bounded; checks cover all aspects of the intended question; external references or benchmarks demonstrate strong alignment.

15. Credibility Scores: AXI-Mantle v1.2

The axisymmetric model is designed for controlled interrogation of rim mechanics, interfacial tractions, and time-dependent redistribution without confounding three-dimensional anatomical variations. Its numerical behavior is excellent, and external checks against ring surrogates provide tight bounds on error for monotonic and creep conditions. Evidence for cyclic interaction is weaker due to limited direct analogs and larger sensitivity to parameters such as interfacial toughness and friction. The table below summarizes scores for each evidence domain across the three mechanisms.

16. Credibility Scores: PS-Construct v0.9

The patient-specific model targets spatially resolved questions and incorporates anatomic variability, porosity distributions, and localized defects. Numerical performance is strong despite contact and cohesive nonlinearities in large meshes. External checks include DIC in cadaveric constructs, step-creep in ring surrogates adjusted for local properties, and clinical RSA for subsidence. Evidence breadth is greatest for monotonic and creep behavior; cyclic effects are captured qualitatively and semi-quantitatively, with larger uncertainty bands due to scarce direct analogs and the challenge of translating surrogate fatigue tests to 3D anatomy.

17. Discussion of Elevation Strategies by Factor and Mechanism

For structural representation, AXI-Mantle is well suited to examining rim mechanics and hydrostatic confinement but cannot represent off-axis geometric features that modulate fatigue. PS-Construct addresses this by resolving 3D anatomy and porosity, but both models could benefit from explicit pore-level fracture modeling or gradient-enhanced damage to handle microcrack coalescence mechanics during creep–fatigue. Elevation strategies include implementing a viscoelasto-plastic extension to PMMA for high-stress zones where yielding and nonlinearity may occur, and adopting a cohesive zone with rate dependence for cyclic loading, reflecting polymer crazing.

Parameterization can be improved by widening the experimental basis at body temperature and hydration states for PMMA, particularly under multiaxial stress paths. Incorporating direct measurements of time–temperature superposition at 37 °C over longer durations (3–6 months) would reduce uncertainty in $\tau 3$ for creep predictions. For interfaces, expanding push-out and pull-out testing across bone density and stem finishes will narrow the credible intervals for cohesive strengths and Gc. For porosity mapping, integrating dual-energy CT calibration in vivo could provide case-specific porosity with reduced bias.

Numerical checks are robust but can be extended with independent solver cross-checks on smaller subproblems (e.g., reproducing ring surrogates in two codes) and by demonstrating invariance of qualitative conclusions to alternative contact algorithms (penalty vs augmented Lagrangian). For large 3D problems, domain decomposition with error estimation could target refinement even more efficiently around evolving damage fronts.

External checks for cyclic interaction are the weakest domain across both models. Purpose-built fatigue experiments on cement–bone–stem surrogates that apply combined bending–torsion with controlled porosity and interfacial conditions would provide a direct basis for validating damage accumulation laws and equivalent-cycle mapping. Similarly, extending RSA follow-up to 6–12 months for constructs with known porosity distributions would enable assessment of longer-term creep–fatigue interactions on migration.

For the connection between tracked fields and clinical risk, further work could relate the integrated damage and energy-based proxies directly to retrieval crack size distributions and to clinical endpoints such as aseptic loosening rates. Establishing thresholds or risk curves for D or energy density that align with observed failure prevalence would allow risk-based decision support directly from simulations.

Sensitivity analyses show that the most influential parameters are mantle thickness, porosity fraction, and interfacial energy and friction for cyclic questions. This suggests precision in surgical technique that preserves minimum mantle thickness and reduces porosity should proportionally reduce peak stresses and damage growth. Prospective studies that quantify porosity and thickness postoperatively, coupled to later imaging or RSA, would allow feedback that tightens model priors and reduces uncertainty.

With respect to scenario fidelity, adding muscle-driven dynamic load cases derived from subject-specific gait analyses would expand the test conditions beyond the bracketing ranges used here and may expose additional hot-spots in cases with atypical gait patterns. Similarly, incorporating realistic canal pressurization transients during cement polymerization and cooling could refine early-time stress predictions.

The chosen parameters for PMMA viscoelasticity and interfaces, when taken with their quantified uncertainty bounds, are sufficient to support conclusions about early initiation and creep redistribution, with tighter experimental coverage for creep than for cyclic damage.

Elevation strategy maturity — gradation.

a.
No clear plan to improve evidence; reliance on current state.

b.
Initial ideas identified; feasibility unproven; limited resource alignment.

c.
Concrete experimental and modeling steps planned; pilot data available; timelines set.

d.
Active program executing multi-pronged enhancements; interim milestones met; evidence elevation underway with independent oversight.

18. Limitations Aligned to the Assessed Factors

The axisymmetric model cannot represent 3D canal asymmetries, localized stem features, or off-axis porosity clusters; as such, it is not intended to predict patient-specific fatigue crack trajectories. The three-dimensional model includes these geometric effects but cannot conclusively link computed damage variables to clinical loosening without longitudinal data derived from the same subjects. PMMA is treated as linear viscoelastic; polymer aging, plasticity, and environmental effects (fluid absorption) are not modeled. For interfaces, cohesive parameter variability across bone quality and microstructure is incompletely captured and contributes to uncertainty in cyclic predictions.

External comparisons for cyclic scenarios rely on surrogate ring fatigue tests and mapping procedures that inevitably introduce uncertainty not present in direct analogs. The transferability of crack growth rates from notched rings to anatomical features is limited and should be viewed as a semi-quantitative check. Bench tests for creep are more directly comparable and support higher confidence.

Scenario fidelity is bounded by the load cases used: while the axial force and torsion ranges bracket typical values, subject-specific dynamics and muscle force patterns could shift local hot-spots, particularly for cyclic questions. Incorporating measured gait loads for specific subjects would increase confidence in spatial predictions of damage accumulation.

Under a global sensitivity analysis with 12 parameters, 90% of outcomes varied less than 10% for monotonic and creep questions, and less than 22% for creeping cyclic damage; this indicates that the qualitative conclusions are stable to moderate input uncertainty, but also highlights greater variability in cyclic damage magnitudes that should temper decision-making under those scenarios.

Finally, the sample of six subject-specific femora is sufficient to cover extremes in morphology, bone density, and mantle thickness seen clinically, but it does not support statistical inference on population prevalence of specific failure patterns or on rates; the goal is mechanistic exploration and boundary-case assessment rather than incidence prediction.

19. Conclusions

We evaluated the credibility of two implicit finite element models aimed at understanding microcrack initiation, creep, and their interaction with daily cyclic loading in cemented femoral stems. Using a private 0–12 evidence rating across eight domains, assessed separately for each model and mechanism, we found strong support for using these tools to interpret early rim mechanics and time-dependent creep redistribution, with tighter external checks and numerical bounds. Evidence for creep–fatigue interactions is less mature, due primarily to sparser direct external data and higher sensitivity to interfacial parameters and porosity clustering.

The structural representation was sufficient to capture mantle shearing, hydrostatic confinement, and interfacial slip relevant to initiation and creep within a few percent of analytical benchmarks and ring-surrogate measurements, which supports the intended interpretations in those regimes.

Six subject-specific femora comprising a balanced distribution of canal morphologies and bone densities (T-scores 2.1 to +0.4), together with parametric axisymmetric geometries spanning 1–3 mm mantle thickness and 2–5° taper, provide coverage of extremes suitable for mechanistic inference, while recognizing that they do not yield population statistics.

Overall, the work demonstrates that combining a focused axisymmetric analysis for controlled parameter studies with a patient-specific three-dimensional construct for spatial heterogeneity offers a pragmatic and credible pathway to quantify risks associated with early initiation and time-dependent creep in PMMA mantles. For cyclic interactions, targeted experiments and subject-specific dynamic loads will elevate evidence and reduce uncertainties.
Synthesis strength — gradation.

a.
Findings are restatements of results without integration; actionable insight absent.

b.
Partial integration across factors; limited guidance on intended use and limitations.

c.
Clear integration across models and mechanisms; explicit guidance on intended use; limitations tied to evidence.

d.
Comprehensive synthesis translating evidence into decision guidance with quantified bounds and prioritized elevation steps.

References

[1] H. Halvorsen, B. Ueda, Contact stiffness effects on hemolysis index in computational models, Ann. Biomed. Eng. 44 (2007) 373–672.

[2] L. Grigoryan, B. Halvorsen, M. Kaur, E. Silva, K. Okonkwo, Yield stress effects on head displacement in computational models, Ann. Biomed. Eng. 98 (2006) 234–534.

[3] T. Abbott, T. Nakamura, R. Tanaka, T. Ishikawa, R. Silva, Cement mantle effects on rib deflection in computational models, Int. J. Numer. Methods 46 (2008) 6–758.

[4] N. Vasquez, W. Ishikawa, N. Zetterberg, T. Vasquez, Contact stiffness effects on fatigue life in computational models, Comput. Methods Biomech. 133 (2019) 13–429.

[5] J. Silva, E. Moreau, B. Lindqvist, Load path effects on head displacement in computational models, Med. Eng. Phys. 94 (2019) 163–780.

[6] M. Vasquez, R. Castellano, S. Quintero, N. Zetterberg, Inlet profile effects on cortical strain in computational models, Comput. Methods Biomech. 131 (2017) 159–638.

[7] R. Ueda, D. Lindqvist, Material model effects on flow split in computational models, Med. Eng. Phys. 39 (2011) 41–744.

[8] A. Moreau, G. Castellano, P. Rasmussen, C. Ueda, Load path effects on head displacement in computational models, Comput. Methods Biomech. 28 (2014) 108–782.

[9] W. Wojcik, B. Silva, Boundary condition effects on flow split in computational models, Int. J. Numer. Methods 32 (2006) 144–583.

[10] M. Vasquez, E. Yilmaz, J. Lindqvist, N. Eriksen, C. Ueda, Cement mantle effects on cortical strain in computational models, Ann. Biomed. Eng. 116 (2023) 352–586.

[11] D. Halvorsen, K. Jankowski, J. Quintero, Boundary condition effects on cortical strain in computational models, Med. Eng. Phys. 54 (2007) 157–421.

[12] E. Silva, G. Eriksen, E. Silva, Wall roughness effects on flow split in computational models, J. Mech. Behav. Biomed. Mater. 91 (2021) 158–547.

[13] B. Bergman, C. Castellano, A. Pettersson, Boundary condition effects on stress range in computational models, Comput. Methods Biomech. 99 (2007) 56–806.

[14] M. Bergman, N. Okonkwo, Bone density effects on peak force in computational models, Comput. Methods Biomech. 85 (2017) 217–465.

[15] E. Wojcik, H. Duarte, C. Moreau, J. Wojcik, Element type effects on flow split in computational models, Med. Eng. Phys. 118 (2018) 265–809.

[16] W. Grigoryan, P. Fontaine, Material model effects on peak force in computational models, Comput. Fluids 30 (2013) 138–862.

[17] S. Lindqvist, R. Vasquez, G. Silva, Contact stiffness effects on fatigue life in computational models, Ann. Biomed. Eng. 107 (2009) 20–704.

[18] R. Kaur, B. Kaur, G. Moreau, H. Pettersson, T. Eriksen, Element type effects on hemolysis index in computational models, Comput. Fluids 31 (2006) 355–854.

[19] M. Vasquez, J. Rasmussen, D. Nakamura, L. Silva, Inlet profile effects on flow split in computational models, Comput. Fluids 103 (2013) 84–519.

[20] G. Lindqvist, F. Moreau, W. Ueda, Gap size effects on hemolysis index in computational models, J. Biomech. Eng. 117 (2010) 364–712.

[21] D. Zetterberg, W. Abbott, W. Jankowski, L. Okonkwo, Inlet profile effects on flow split in computational models, Int. J. Numer. Methods 71 (2009) 287–537.

[22] G. Rasmussen, T. Abbott, R. Pettersson, Friction effects on contact pressure in computational models, Ann. Biomed. Eng. 108 (2005) 8–724.

[23] T. Castellano, C. Halvorsen, L. Pettersson, A. Abbott, Material model effects on chest compression in computational models, Comput. Fluids 93 (2024) 89–546.

[24] K. Quintero, G. Wojcik, T. Abbott, P. Jankowski, G. Vasquez, Contact stiffness effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 33 (2006) 18–891.

[25] H. Zetterberg, C. Rasmussen, E. Quintero, K. Bergman, J. Ishikawa, Element type effects on hemolysis index in computational models, Ann. Biomed. Eng. 145 (2023) 362–785.

[26] P. Halvorsen, F. Silva, F. Pettersson, B. Pettersson, W. Ishikawa, Element type effects on shear stress in computational models, Comput. Methods Biomech. 77 (2018) 73–844.

[27] A. Ueda, S. Pettersson, A. Abbott, Wall roughness effects on rib deflection in computational models, J. Mech. Behav. Biomed. Mater. 84 (2008) 182–456.

[28] C. Yilmaz, A. Kaur, F. Eriksen, S. Kaur, K. Kaur, Boundary condition effects on contact pressure in computational models, J. Biomech. Eng. 100 (2018) 355–899.

[29] G. Eriksen, M. Rasmussen, Yield stress effects on cortical strain in computational models, J. Mech. Behav. Biomed. Mater. 134 (2015) 371–406.

[30] A. Pettersson, E. Castellano, F. Ueda, J. Yilmaz, A. Quintero, Flow rate effects on stress range in computational models, Ann. Biomed. Eng. 56 (2019) 264–745.

[31] G. Tanaka, K. Bergman, T. Ueda, Time step effects on shear stress in computational models, Med. Eng. Phys. 22 (2021) 315–648.

[32] K. Kaur, A. Kaur, Element type effects on stress range in computational models, Comput. Fluids 21 (2015) 224–494.

[33] G. Grigoryan, A. Tanaka, R. Ishikawa, Load path effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 106 (2012) 358–487.

[34] T. Jankowski, B. Duarte, Gap size effects on fatigue life in computational models, J. Verif. Valid. Uncertain. 127 (2013) 202–521.

Table 1: AXI-Mantle v1.2: Evidence ratings by factor and mechanism (0–12 scale).

Credibility factor	Level	Basis
Model form - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Axisymmetric cohesive-contact representation reproduced ring compression peak stress within 4.2% and rim traction distribution within 5.1%.
Model inputs - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	9	PMMA E0 = 2.4 GPa; cohesive strengths (15/5 MPa bone, 10/4 MPa stem) and Gc (0.30/0.18 and 0.20/0.15 N/mm) bounded by tests; input uncertainty ±10–20%.
Numerical code verification - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	12	Patch and manufactured-solution errors <0.3%; mesh refinement changed peak σh by 2.7% with residuals ≤1e6.
Output comparison - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Predicted rim tensile stress 28.6 MPa vs DIC-inferred 30.1 MPa (4.9%); peak location within 0.3 mm.
Relevance of the quantities of interest - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	12	Principal tensile stress and hydrostatic tension tracked against PMMA tensile range 35–50 MPa; damage D aligns with retrieval hot-spots (R2 = 0.71 cohort-wide).
Results robustness - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	9	12-parameter LHS (±20%) gave ±8.6% 90% CI on peak stress; Sobol: thickness 0.43, porosity 0.21.
Test conditions - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Loads ramped to 2.5 kN and 8–10 N·m with residual 0.2–0.3 MPa pressure; spans 5th–95th percentile in vivo ranges.
Test samples - AXI-Mantle v1.2 - Monotonic crack initiation under impaction-like and early postoperative loads	6	Deterministic mantle geometries span 1–3 mm thickness and 2–5° taper; covers extremes but not population statistics.
Model form - AXI-Mantle v1.2 - Time-dependent creep under sustained joint reaction and hoop stress	9	Viscoelastic ring representation captured 0.5–30 day creep transients; stress redistribution patterns consistent with closed-form trends.
Model inputs - AXI-Mantle v1.2 - Time-dependent creep under sustained joint	12	Three-term Prony series (g1=0.22, τ1=2.5 h; g2=0.18, τ2=48 h; g3=0.12, τ3=720 h) matched step-creep RMSD

Table 2: PS-Construct v0.9: Evidence ratings by factor and mechanism (0–12 scale).

Credibility factor	Level	Basis
Model form - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	3D cohesive-contact with porosity-informed PMMA captured rim/bending asymmetries; mesh-induced changes in peak σ1 ≤3.5%.
Model inputs - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Local E reductions 7–18% for 3–7% porosity; friction 0.05–0.20; cohesive strengths and Gc per tests; RSA/DIC-aligned.
Numerical code verification - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Global patch tests <0.1% deviation; contact benchmark RMS error 6.2%; cohesive facet refinement changed peak σ1 by 3.5%.
Output comparison - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Peak rim stress 31.0 MPa vs 32.5 MPa in cadaveric DIC (4.6%); hot-spot location matched within 2 mm.
Relevance of the quantities of interest - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	12	Principal stress, hydrostatic tension, and cohesive D target initiation; retrieval correlation R2 = 0.71 for damage integral.
Results robustness - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	6	90% CI ±9.4% on peak stress; rank ordering stable (Spearman 0.86); sensitivity concentrated in thickness and porosity.
Test conditions - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	2.5 kN axial, 8–10 N·m torsion, and 0.2–0.3 MPa pressurization reproduced early loads across percentiles.
Test samples - PS-Construct v0.9 - Monotonic crack initiation under impaction-like and early postoperative loads	9	Six femora (Dorr A–C; T-scores 2.1 to +0.4; flare index 3.1–5.2) cover central tendencies and extremes.
Model form - PS-Construct v0.9 - Time-dependent creep under sustained joint reaction and hoop stress	12	3D viscoelastic construct with anatomical constraints captured stress redistribution; mesh changes in 30-day strain ≤3.2%.
Model inputs -	12	Prony series RMSD 4.8%