# Credibility Assessment of Implicit Finite Element Models of a Shoulder Glenoid Component for Rocking-Horse Loosening


Elena M. Duarte, PhD[a], Rahul Venkataraman, MS[b], Katherine L. Song, PhD[c], Michael J. Anders, MD[1]

[a]*Computational Orthopaedics Laboratory, OrthoBio Mechanics Institute, Austin, TX, USA,*
[b]*Department of Mechanical Engineering, University of Texas at Austin, Austin, TX, USA,*
[c]*Department of Orthopaedic Surgery, Central Texas Medical Center, Austin, TX, USA,*



## Abstract

Objective: To assess the credibility of two implicit finite element analysis (FEA) models designed to evaluate rocking-horse loosening mechanisms of a cemented anatomic shoulder glenoid component. The Continuum Glenoid–Scapula Model (CGSM) resolves the full implant–cement–bone construct under eccentric toggling to quantify rim opening/closure, cement hydrostatic tension, contact patch migration, and cement–bone micromotion. The Pegged Interface Cohesive Model (PICM) interrogates local peg–cement–bone subassemblies using cohesive zone elements to estimate interface tractions, slip/opening, a damage variable, and cycles-to-initiation surrogates for shear fatigue. Methods: The models were developed for quasi-static envelopes representative of clinical rocking-horse loading and were verified for mesh/time increment convergence. Validation used cadaveric rocking tests (n = 7 specimens) with digital image correlation (DIC) rim lift-off tracking, subassembly shear tests, and peg–cement shear fatigue coupons (n = 12). Model outputs were compared against experimental metrics under matched boundary conditions (37 ± 0.5 C saline, 1 Hz, 46 mm humeral head, 5 mm superior/inferior eccentricity). Credibility was graded on a four-point CredMark-12 scale (0, 4, 8, 12) and summarized per factor for each (model × mechanism) pair. Results: For rim toggle–induced tensile debonding, CGSM predictions of rim lift-off (41 µm median) matched DIC measurements (45 ± 12 µm) within experimental variability, and contact patch excursion (7.6 mm) tracked pressure film estimations (8.2 ± 1.4 mm). Peak cement hydrostatic tension (4.1 MPa) was consistent with the onset of microcracking on micro-CT near 80k cycles in 3/7 specimens. Mesh/time refinement studies bounded discretization errors for rim lift-off to 2.7%. Monte Carlo propagation (n = 200) produced 95% predictive intervals of 31–55 µm for rim lift-off and 3.0–5.4 MPa for hydrostatic tension. For peg–cement shear fatigue, PICM matched subassembly slip amplitudes (model 6.5 µm vs measured 7.1 µm, 8.5% error) and produced cohesive damage accumulation consistent with coupon cycles-to-initiation (model 1.3 × 10^6 vs median 1.2 × 10^6, within 12%). Credibility scores for CGSM were highest for the rim toggle mechanism, particularly for alignment of outputs with the mechanism and relevance of validation conditions; they were lower for peg–cement fatigue because the continuum model lacks an explicit fatigue law. PICM achieved high scores for the peg–cement mechanism but only moderate scores for rim debonding due to scope. Conclusions: The CGSM is fit to estimate global toggling metrics and cement stress states driving rim debonding, while the PICM is fit to estimate peg–cement interface shear fatigue initiation under rocking. Together, they provide complementary evidence for preclinical evaluation of rocking-horse loosening. The assessment identifies priority upgrades: direct measurement of peg-layer boundary conditions during whole-construct toggling, expansion of cohesive calibration to mixed-mode at physiological frequencies, and explicit uncertainty quantification for long-term cyclic degradation.



*Keywords:* shoulder arthroplasty, finite element analysis, rocking-horse loosening


## 1. Introduction

A significant portion of aseptic glenoid loosening in anatomic total shoulder arthroplasty is associated with the so-called rocking-horse mechanism, in which alternating eccentric loading of the polyethylene articular surface toggles the component on its cement mantle. The superior–inferior eccentricity of joint force during activities of daily living promotes alternating compression and lift-off at the glenoid rim. This toggling can induce tensile debonding at the cement–bone interface near the lifted rim and accumulate shear damage around cemented pegs or posts. Early-stage loosening often presents radiographically as progressive radiolucent lines and clinically as increasing pain and reduced range of motion; retrievals frequently show fissures in the cement adjacent to the rim and polished wear at peg–cement contacts.

Finite element models are widely used to explore the biomechanics of these processes under controlled scenarios that complement experimental tests. Yet, the credibility of such models for decision-making about failure risk hinges on transparent evidence that the chosen model form, input data, numerical solution quality, and alignment with physical tests together provide sufficient assurance for the stated context of use. In this work, we report a structured assessment for two implicit finite element models tailored to the dominant mechanisms believed to underlie rocking-horse loosening: rim toggle–induced tensile debonding and peg–cement shear fatigue during rocking. The first model, a continuum representation of the full implant–cement–bone construct (CGSM), aims to quantify global toggling kinematics and cement stress states. The second, a pegged interface cohesive model (PICM), focuses on local mixed-mode tractions and damage at the peg–cement and cement–bone interfaces under cyclic shear.

The assessment follows a factor-by-factor approach, scoring the sufficiency of evidence along several dimensions that practitioners routinely weigh: provenance and traceability of experimental and input data, fidelity and appropriateness of the model inputs, the control and reporting of numerical solution errors, how directly the chosen outputs relate to the failure mechanisms, the degree to which validation exercises mirror the intended use, characterization of uncertainties, and the rigor of development and change management. We adopt a four-point, narrative-anchored 0–12 scale where 0 denotes absent or unusable evidence, 4 denotes partial but limited support, 8 denotes sound and traceable support, and 12 denotes comprehensive and persuasive support aligned with the intended decisions.

We deliberately treat the two mechanisms as distinct contexts of evaluation for each model. For example, a continuum model that calculates cement hydrostatic tension may map well to rim debonding but not fully capture the cyclic shear damage processes local to peg–cement interfaces; conversely, a cohesive submodel may faithfully estimate interface slip histories without portraying the global kinematics that position the contact patch. The factor scores, narrative justifications, and tables that follow are organized to convey this nuance clearly. Our goal is to guide both internal model use and external review by making the basis for each judgment explicit in the body of the paper and concise in summary tables that can be read independently.

*Scoring framework overview — gradation.*

a.
0: No stated scale; ad hoc language and undocumented judgments.

b.
4: Scale is stated but without anchors or examples; limited mapping to decisions.

c.
8: Scale is stated with narrative anchors and representative examples per factor.

d.
12: As at 8, plus quantitative thresholds tied to decisions and cross-checked with stakeholders.

## 2. Device and Failure Mechanisms

The device of interest is a cemented anatomic glenoid component featuring an ultra-high molecular weight polyethylene (UHMWPE) articular surface, an elliptical footprint 46 × 38 mm, a nominal thickness of 6 mm centrally tapering to 3 mm at the rim, and three cemented pegs (central peg 6 mm diameter × 18 mm length; two peripheral pegs 5 mm diameter × 14 mm length) with circumferential grooves to enhance cement keying. The component is intended for use with a 46 mm cobalt–chromium humeral head. The surgical technique prepares a shallow biconcave bed in the glenoid fossa, irrigates and dries the bone, and pressurizes polymethyl methacrylate (PMMA) bone cement into the cancellous bed prior to seating the component with manual compression.

Two mechanisms are the focus of this study. First, superior–inferior eccentric joint force toggles the component about the cement mantle, lifting the inferior rim during superior loading and the superior rim during inferior loading. This alters the contact patch on the articular surface, introduces tensile stresses within the cement adjacent to the lifted rim, and can open the cement–bone interface locally. Repeated cycles may initiate cracks at the rim, which can coalesce and propagate, compromising fixation. Second, even as the cement mantle overall remains bonded, alternating shear develops around the pegs as the implant rocks, especially near the peg flanks. This shear may accumulate damage at the peg–cement interface, causing micro-slip and progressive degradation in shear capacity, a process that can ultimately reduce construct stability and aid loosening.

*Mechanism-to-output mapping — gradation.*

a.
0: Model outputs are unrelated to hypothesized mechanisms.

b.
4: Model outputs are qualitative indicators of mechanisms but lack quantitative linkage.

c.
8: Model outputs include direct surrogates that have been linked to mechanism onset or progression.

d.
12: As at 8, with multiple, triangulated outputs corroborated by mechanistic tests and retrieval analyses.

## 3. Computational Models and Assumptions

The Continuum Glenoid–Scapula Model (CGSM) is a three-dimensional implicit finite element model of the implant, cement mantle, and scapular bone. It resolves the UHMWPE articular surface with hexahedral and tetrahedral continuum elements and represents the cortical and cancellous regions of the scapula with region-specific elastic properties informed by computed tomography (CT). The cement mantle is modeled as a uniform layer whose thickness follows the implanted morphology; in the nominal case this thickness varies from 1.8 to 2.6 mm under the central footprint and thins toward the rim. Nonlinear contact between the UHMWPE articular surface and a rigid 46 mm humeral head surrogate is computed under a low friction coefficient, and the implant–cement–bone stack allows separation and sliding at the cement–bone interface to admit micromotion calculations. The model computes quasi-static responses at the load extremes (superior and inferior eccentricities) and interpolates cyclic behavior from these two stable configurations. Assumptions include linearly elastic materials for UHMWPE, PMMA, and bone, small strains for bone and cement regions, and large-deformation contact at the articular surface.

Figure 1: Schematic depiction of the rocking-horse mechanism with superior and inferior eccentric loading, showing rim lift-off, contact patch migration, and shear around a peripheral peg.

The Pegged Interface Cohesive Model (PICM) is a local submodel that extracts a subassembly around a single peg and its surrounding cement and cancellous bone. It uses cohesive zone elements to represent the peg–cement and cement–bone interfaces explicitly. The traction–separation laws are mixed-mode, with separate strengths and fracture energies for normal and shear directions, and are calibrated on coupon-level fatigue tests. The surrounding cement and bone are elastic, and frictional contact may also be included between debonded faces to represent post-initiation sliding. The PICM is driven by displacement or traction boundary conditions mapped from the CGSM at the submodel boundary to ensure that the local fields reflect global toggling-induced shear and opening/closing. Cyclic loading is represented either by steady-state amplitude sweeps or by a cycle-jump strategy where damage evolves per cycle according to the cohesive law, allowing estimation of a cycles-to-initiation surrogate when the scalar damage variable first reaches unity at a critical point.

Across both models we include constraints to reflect surgical seating and physiological support: the scapula is constrained at the medial border and inferior angle in a potting block boundary that replicates the test fixture, while the joint reaction is applied by a rigid humeral head that translates to impose a 5 mm superior or inferior offset relative to the glenoid center. Both models share the same geometric source data and boundary conditions but differ in resolution and constitutive treatment of interfaces. The CGSM is intended to provide kinematics and stress states across the construct; the PICM is intended to quantify the local mixed-mode conditions and fatigue accumulation relevant to peg–cement shear.

*Model form suitability — gradation.*

a.
0: Model form cannot represent the mechanism of interest.

b.
4: Model form captures the mechanism qualitatively with idealizations that limit quantitative use.

c.
8: Model form captures dominant physics with justifiable simplifications; secondary effects acknowledged.

d.
12: Model form captures dominant and secondary physics validated against mechanism-specific observables.

## 4. Geometry, Materials, Contacts, and Boundary Conditions

Geometry was derived from a high-resolution CT scan (0.3 mm isotropic voxels) of a cadaveric scapula implanted with the target glenoid component using the manufacturer's guide. The implant CAD was registered to the scan using an iterative closest-point algorithm, and the cement mantle was segmented by thresholding followed by manual cleanup to conform the mantle to the component pegs and rim. The cortical shell was segmented via a 1.0 mm thresholded layer around the glenoid cavity and lateral border; cancellous bone occupied the inner volume. The resultant mantle thickness was 2.2 ± 0.3 mm under the central footprint (mean ± SD across 100 sampled points) and 1.1 ± 0.2 mm near the rim. The three pegs were orthogonal to the bed; the central peg was 6 mm in diameter and 18 mm long; peripheral pegs were 5 mm in diameter and 14 mm long, with four circumferential grooves each spaced 3 mm apart and 0.4 mm deep.

Material models were linear elastic and isotropic in the CGSM: UHMWPE with elastic modulus E = 1.0 GPa and Poisson's ratio $\nu$ = 0.40; PMMA cement with E = 2.4 GPa and $\nu$ = 0.30; cortical bone with E = 14 GPa and $\nu$ = 0.30; cancellous bone with E = 0.5 GPa and $\nu$ = 0.20. For the PICM, the same continuum material properties were assigned, and the cohesive interfaces used bilinear traction–separation relations with shear strength $\tau 0$ = 12 MPa and shear fracture energy Gc,s = 0.35 N/mm, normal strength $\sigma 0$ = 6 MPa and normal fracture energy Gc,n = 0.15 N/mm. These parameters were calibrated against coupon shear fatigue data as described later. After complete decohesion, a Coulomb friction coefficient of 0.30 governed tangential sliding with hard contact in the normal direction.

Contacts and boundary conditions were defined to reproduce the test configuration. A rigid analytical surface represented the 46 mm humeral head; the UHMWPE articular surface contacted the head with a friction coefficient of 0.04. In the CGSM, the cement–bone interface allowed separation in the normal direction with penalty stiffness chosen to prevent overclosure (established via sensitivity) and tangential sliding with friction coefficient 0.25; the implant–cement interface was tied. The scapula was embedded proximally in a potting block representation with encastre boundary at the medial border and inferior angle, and with its articular plane oriented according to surgical technique. The humeral head was displaced to achieve a 5 mm superior offset relative to the glenoid centroid while applying a compressive load ramp up to 700 N and then subjected to a mirrored inferior offset. The PICM inherited displacement or traction fields along its outer boundary from the CGSM corresponding to the same superior and inferior offset extremes. Frequency-dependent effects were neglected in both models because the intended quasi-static extremes were assumed to dominate damage initiation for the ranges tested.

Test and model environments were matched: components were hydrated in isotonic saline at 37 ± 0.5 C during the experimental campaigns and the model boundary conditions mimicked the same 1 Hz sinusoidal toggling by solving the superior and inferior extreme poses; load application rate in the rig was 20 N/s and the same end states were reproduced in the solver by accompanying displacement control that approached the target load without oscillation.

*Test conditions alignment — gradation.*

a.
0: Loading, geometry, or environment differ substantially from the intended use.

b.
4: Some aspects match; key aspects (e.g., geometry or environment) differ or are uncontrolled.

c.
8: All key aspects match within stated tolerances; minor differences documented.

d.
12: Key aspects match and are verified by instrumented measurements of boundary conditions.

Table 1: Device, material, and boundary condition summary used in the computational models and in matching experiments.

| Credibility factor | Level | Basis |
|---|---|---|
| Implant geometry | — | Elliptical footprint 46 × 38 mm, central thickness 6 mm tapering to 3 mm at rim; 3 pegs (central 6 × 18 mm; peripheral 5 × 14 mm) |
| Cement mantle | — | PMMA mantle thickness 2.2 ± 0.3 mm under footprint; 1.1 ± 0.2 mm at rim |
| Materials (continuum) | — | UHMWPE E = 1.0 GPa, $\nu$ = 0.40; PMMA E = 2.4 GPa, $\nu$ = 0.30; Cortical E = 14 GPa, $\nu$ = 0.30; Cancellous E = 0.5 GPa, $\nu$ = 0.20 |
| Cohesive (PICM) | — | Shear $\tau$0 = 12 MPa, Gc,s = 0.35 N/mm; Normal $\sigma$0 = 6 MPa, Gc,n = 0.15 N/mm; post-failure friction μ = 0.30 |
| Contacts | — | UHMWPE–head friction μ = 0.04; cement–bone μ = 0.25 with normal separation allowed (CGSM); cohesive interfaces at peg–cement and cement–bone (PICM) |
| Boundary conditions | — | Rigid 46 mm humeral head; 700 N compressive load with 5 mm superior/inferior offset; scapula potted at medial border and inferior angle |
| Environment | — | 37 ± 0.5 C saline, 1 Hz sinusoidal toggling in tests; quasi-static extremes solved in models |

## 5. Model Inputs and Pre-processing

Model generation began with segmentation of a CT volume to delineate scapular cortical and cancellous domains and the implanted glenoid component with its cement mantle. Semi-automatic thresholding and morphological operations were used to isolate the cement, aided by the high contrast between PMMA and bone. The implant CAD was registered and boolean operations were performed to create a watertight cement volume. Meshes were generated using an advancing-front algorithm, with element sizing specified by proximity to features of interest. For the CGSM, quadratic tetrahedra (C3D10) were used for bone and cement, and a mix of C3D10 and reduced-integration hexahedra (C3D8R) were used in the UHMWPE to improve contact smoothness with the rigid head. Target edge lengths were 0.5 mm near the rim and peg–cement transitions, 0.8 mm in the central mantle, and 1.5–2.0 mm in the cortical shell and peripheral cancellous bone. The final CGSM contained approximately 1.8 million degrees of freedom.

For the PICM, a submodel domain was extracted encompassing a peripheral peg, the surrounding cement out to 8 mm radius, and an 18 mm depth of cancellous bone. Meshes used C3D10 elements with 0.25 mm sizing near the cohesive interfaces, transitioning to 0.5–0.75 mm in the outer cement and cancellous bone. Zero-thickness cohesive elements with 0.25 mm facet edge length were inserted at the peg–cement and cement–bone interfaces. Boundary conditions for the submodel were derived by interpolating displacements from the CGSM at nodes on the cut boundary for both superior and inferior toggling extremes; for cyclic analyses, a sinusoidal variation at 1 Hz of the mapped shear component was imposed with amplitude matching the extracted shear at the submodel boundary.

Material properties were assigned based on a combination of manufacturer specifications and literature values. The UHMWPE elastic modulus of 1.0 GPa was corroborated by tensile coupons cut from production-stock sheets tested at 37 C (n = 5; 0.98 ± 0.06 GPa). PMMA cement modulus of 2.4 GPa was selected from compressive testing of cylindrical casts prepared from the same manufacturer's lot used in the cadaveric implants (n = 8; 2.42 ± 0.15 GPa). Cortical and cancellous bone moduli followed published ranges for the glenoid; localization of cortical thickness and cancellous density was achieved from the CT by segmentation but elastic moduli were not spatially graded in this version. Friction coefficients for UHMWPE–head (0.04) and cement–bone (0.25) were taken from tribological studies under saline lubrication with machined PMMA and bovine cancellous analogs, respectively.

Interface parameters in the PICM were calibrated using peg–cement shear coupons machined from off-the-shelf pegs embedded centrally in PMMA cylinders with grooved surfaces matching the component geometry. Stainless steel adapters clamped the peg core. Under fully reversed shear (R = 1) at 2 Hz in 37 C saline, the cyclic shear amplitude was stepped across specimens; we recorded cycles to a 5% drop in secant stiffness for n = 12 coupons. The median cycles-to-initiation was 1.2 × 10^6 at 0.6 MPa shear amplitude (coupon nominal contact area basis). Shear traction–separation parameters $\tau$0 and Gc,s were adjusted to replicate the stiffness drop and dissipated energy per cycle at that amplitude; normal parameters $\sigma$0 and Gc,n were set to replicate a single-load peel test with 0.1 MPa tension at which initial separation was observed at the cement–bone interface.

Pre-processing quality checks included visual inspection of mesh quality (Jacobian > 0.5 for tetrahedra), node-to-surface projection for cohesive insertion, and verification of displacement mapping from the CGSM to the PICM outer boundary through a synthetic rigid-body motion test. Contacts were verified by probing normal gaps at the UHMWPE–head interface to ensure non-penetration post-solve, and interfacial penetration at cement–bone was restricted to < 2 μm under peak compressive load by tuning penalty stiffness with sensitivity sweeps.
*Model inputs — gradation.*

a.
0: Key parameters are guessed or undocumented.

b.
4: Key parameters are documented; some are estimated without justification or outside known ranges.

c.
8: All key parameters are measured or sourced from peer-reviewed studies; estimates are justified and within known ranges.

d.
12: As at 8, plus independent corroboration of parameters and traceable linkage to the specific lot/specimen used.

## 6. Numerical Methods and Solver Characterization

All analyses used an implicit finite element solver with a Newton–Raphson algorithm for nonlinear equilibrium and an augmented Lagrangian formulation for contact. For the CGSM, static general steps ramped the load to 700 N with simultaneous displacement control to achieve the 5 mm eccentric offset. Contact stabilization used automatic penalty update with a target overclosure tolerance below 1 μm at the UHMWPE–head interface. Tangential contact used a penalty formulation with the stated friction coefficients. Convergence criteria were set to a residual force norm of 10^8 times the reference load and a displacement norm of 10^8 times a characteristic length (1 mm). A line search with maximum damping factor of 0.8 was enabled to improve convergence near contact patch migration. The time stepping strategy used 50 quasi-time increments for each extreme, with automatic sub-stepping when the cutback criterion was triggered by contact chattering or stiffness changes.

For the PICM, the cohesive softening demanded robust control of path-following. The analyses used either static general steps for monotonic loading or a Riks arc-length method for post-peak response segments when cohesive damage localized. Damage evolution per cycle was computed using a cycle-jump approach: a representative cycle was solved at amplitude, and then damage was incremented per cycle using the cohesive law, with recalculation of stiffness and slip fields at cycle milestones (e.g., 10^2, 10^3, 10^4 cycles). Convergence tolerances matched those of the CGSM. A small viscosity regularization ($\eta$ = 0.001 s) in the cohesive constitutive law helped suppress snap-back instabilities without altering the steady-state amplitudes beyond 1% relative error, verified via parameter sweep.

Verification activities addressed mesh and increment sensitivity. In the CGSM, rim lift-off at the inferior edge under 700 N superior offset was computed on a baseline mesh (0.5 mm local size) and a refined mesh (0.35 mm local size). The displacement decreased from 42.1 μm to 41.0 μm, a 2.7% change. Halving the maximum increment size reduced lift-off by a further 0.9% on the refined mesh. Peak cement hydrostatic tension at the inferior rim decreased from 4.23 MPa to 4.09 MPa with mesh refinement (3.3%) and 0.7% with step refinement. In the PICM, cohesive slip amplitude at the peg–cement interface under mapped shear boundary conditions decreased from 6.8 μm to 6.5 μm (4.3%) when the cohesive facet edge length was reduced from 0.25 mm to 0.18 mm; halving the step size in the Riks arc-length method changed slip amplitude by 1.6%. These results bounded numerical solution errors for the chosen meshes and increments to below 5% on the primary outputs.
*Numerical solver error — gradation.*

a.
0: No verification; mesh and time-step dependence unknown.

b.
4: Single sensitivity check; limited metrics; impact on outputs not quantified.

c.
8: Systematic mesh and step-size studies reported; impact on primary outputs bounded.

d.
12: As at 8, with Richardson extrapolation or asymptotic error estimates and target-based adaptivity.

## 7. Quantities of Interest and Post-processing

The CGSM reports the following quantities of interest (QoIs) under superior and inferior toggling extremes. Rim lift-off or closure displacement is computed as the normal gap or overlap at nodes along the glenoid rim in the lifted or compressed quadrant, respectively, with the inferior edge under superior load and the superior edge under inferior load. The maximum rim lift-off across circumferential nodes is extracted as a scalar QoI. Cement hydrostatic tension is computed as the positive mean stress within elements immediately subjacent to the lifted rim (1 mm rim-proximal band); the peak value in that band is reported. Contact patch migration is quantified by mapping the contact pressure footprint on the UHMWPE surface onto the component plane and measuring the superior–inferior excursion of the centroid between the two extremes. Cement–bone micromotion is computed as the relative tangential displacement between coincident cement and bone surface nodes integrated along the interface; the maximum amplitude under the superior–inferior toggling is reported. The PICM reports cohesive tractions (normal and shear), the scalar damage variable at cohesive integration points, and relative slip or opening displacements at the peg–cement and cement–bone interfaces. A cycles-to-initiation surrogate is estimated by counting cycles until the damage variable first reaches unity at any cohesive integration point under the imposed cyclic shear amplitude.

Post-processing included scripts to collate node sets along the rim, normal vector evaluation to ensure that the rim lift-off was measured in the local outward normal direction, and spatial averaging in the 1 mm rim-proximal band to mitigate spurious peaks from element-level stress oscillations. The hydrostatic tension was verified to be insensitive to the exact band thickness over the range 0.5–1.5 mm, changing less than 4% in the CGSM rim region. For the PICM, cohesive damage fields were examined to identify the initiation point; in most cases, the first D = 1 location occurred near the lateral flank of the peripheral peg at roughly mid-depth. For mixed-mode conditions, an equivalent energy release metric was computed from the cohesive law to ensure that the calibration from shear coupons extended smoothly to the combined loading states in the subassembly.

Outputs were exported for comparison to experimental measures: rim lift-off was aligned to DIC-measured out-of-plane displacements at corresponding rim regions; contact patch centroid excursion was compared to pressure-sensitive film data collected under matching loads in the cadaveric tests; interfacial micromotion was compared qualitatively to radiographic and micro-CT evidence of early debond and, where available, to DVRT (differential variable reluctance transducer) measurements of cement–bone relative motion on three specimens instrumented with miniature transducers.

*Relevance of the quantities of interest — gradation.*

a.
0: Reported outputs have no clear link to decisions.

b.
4: Reported outputs are qualitatively related to decisions; quantitative mapping not established.

c.
8: Reported outputs have a documented quantitative relationship to decisions or mechanistic thresholds.

d.
12: Multiple outputs jointly triangulate decisions with documented thresholds and external corroboration.

## 8. Data Sources and Pedigree

Input parameters and validation data were obtained from a combination of in-house tests and published sources. The principal validation set comprised cadaveric scapulae (n = 7) implanted with the study glenoid using a single experienced surgeon's technique. Each specimen was potted in a standardized block and tested in a servo-hydraulic frame with a spherical 46 mm cobalt–chromium head applying compressive load up to 700 N at 1 Hz with 5 mm superior/inferior eccentricity. The test environment was thermostatted saline at 37 ± 0.5 C. DIC speckle patterns on the glenoid rim enabled measurement of out-of-plane displacements with a spatial resolution of ~0.1 mm and a displacement resolution of ~5 µm. Three specimens were instrumented with miniature DVRTs spanning the cement–bone interface at the inferior rim. Pressure-sensitive film (0.2–0.6 MPa range) inserted between the head and the UHMWPE during a separate set of monotonic tests on two additional implants provided a check of contact patch movement. Additionally, micro-CT scans (49 µm voxel) were acquired before testing and at 20k, 50k, and 100k cycles to detect new radiolucencies and microcrack-like features in the cement near the rim.

Cohesive parameter calibration data were drawn from peg–cement shear coupons (n = 12) prepared using the same PMMA lot used in the cadaveric implants to limit material variability. Coupon geometry replicated the peg grooves; the PMMA cylinders were 20 mm in diameter and 25 mm in length, centrally containing the peg. Fully reversed shear was applied at 2 Hz in saline at 37 C through a torsional fixture that converted rotational oscillation into circumferential shear along the peg–cement interface. A run-out criterion of 2 × 10^6 cycles was used; cycles to 5% stiffness drop were recorded for each coupon.

Published sources were used to corroborate friction coefficients and material properties. UHMWPE elastic modulus was cross-checked against a pooled literature mean of 0.9–1.2 GPa at 37 C; PMMA elastic modulus against 2.0–2.5 GPa in compression; cortical and cancellous bone moduli against 12–17 GPa and 0.2–1.0 GPa ranges respectively. Tribology studies under saline with cobalt–chromium heads reported UHMWPE–metal friction coefficients of 0.02–0.06, and laminar sliding tests between PMMA and cancellous analogs reported friction coefficients of 0.2–0.4 depending on surface roughness and preparation.

Data management included version-controlled storage of all test procedures, sensor calibrations, and raw files; each specimen dataset included a signed test report, DIC image sets, load–displacement histories, and environmental logs. Peer-reviewed literature citations were cataloged with full bibliographic metadata and archived PDFs. While the body of this paper focuses on the analysis and interpretation needed for credibility assessment, the underlying datasets are preserved and traceable to their sources by specimen and test identifiers.

*Data pedigree — gradation.*

a.
0: Sources are anecdotal or missing; no traceability to raw data.

b.
4: Sources are partially documented; traceability to some raw data exists.

c.
8: Sources are fully documented; raw data, calibrations,

and metadata are traceable for all key datasets.

d.
12: As at 8, plus independent audit or third-party certification of data integrity.

## 9. Validation Evidence and Its Relevance

CGSM validation compared predicted rim lift-off, contact patch movement, and cement–bone micromotion to cadaveric measurements under matched loading and environmental conditions. The DIC analysis yielded inferior rim lift-off under 700 N superior loading with a specimen-averaged mean of 45 μm and standard deviation 12 μm. The CGSM predicted 41 μm rim lift-off for the representative specimen geometry and boundary conditions, which falls within the experimental variability. Contact patch centroid excursion between superior and inferior extremes was estimated as 8.2 ± 1.4 mm from pressure-sensitive film tests; the CGSM computed 7.6 mm for the same load. DVRT measurements at the inferior rim recorded cement–bone relative slips of 17 ± 6 μm at peak load; the CGSM predicted 15 μm. Qualitative comparison of micro-CT before and after 80k cycles revealed three specimens with new low-attenuation lines near the inferior rim consistent with cement cracks; the CGSM's peak hydrostatic tension in that region was 4.1 MPa, consistent with reported thresholds for cement tensile cracking under cyclic conditions when coupled with stress concentrations.

PICM validation proceeded at two levels. First, cohesive parameters were calibrated to match coupon cycles-to-initiation under cyclic shear of 0.6 MPa amplitude, achieving a predicted initiation at 1.3 × 10^6 cycles compared to the median coupon value of 1.2 × 10^6 cycles (8% difference). Second, a subassembly specimen comprising a peripheral peg with surrounding cement and a block of cancellous bone was excised from a separate cadaveric implant and loaded in a custom shear fixture to impose mapped boundary conditions extracted from CGSM superior toggling. The subassembly demonstrated an interface slip amplitude of 7.1 μm by LVDT measurement at the peg flank under the mapped shear amplitude; the PICM under identical boundary conditions predicted 6.5 μm (8.5% error). Cohesive damage localized initially at the mid-depth of the peg flank, which is consistent with post-mortem micro-CT indicating subtle deformations of the cement grooves in that region.

We also conducted sensitivity checks on validation metrics: the DIC displacement resolution of ~5 μm sets a lower bound on the rim lift-off measurement error; the pressure-sensitive film's low-pressure threshold (0.2 MPa) limited the accuracy of contact patch edges, but centroid movement remained robust across repeated prints. The DVRTs, while precise, sampled only a small arc length of the interface. Coupon tests, run at 2 Hz to achieve practical test durations, may not replicate frequency effects experienced in vivo; nonetheless, the traction–separation calibration produced plausible shear strengths and energies that yielded consistent predictions in the subassembly context.

Collectively, the validation touches the global toggling kinematics and stress states central to rim debonding and the local interface tractions and damage accumulation pertinent to peg–cement shear fatigue, using test conditions that adhere to the modeled boundary states and environments.
*Relevance of the validation activities to the COU — gradation.*

a.
0: Validation conditions bear little resemblance to intended use.

b.
4: Some validation elements match use; key loads or constraints differ.

c.
8: Most validation elements match within tolerances; remaining gaps are justified and bounded.

d.
12: Validation replicates intended use with matched geometry, loads, constraints, and environment; metrics map directly to QoIs.

## 10. Results Uncertainty Characterization

Uncertainty in model outputs arises from variability in material properties, geometry (cement mantle thickness and peg seating), friction coefficients, and measurement noise in validation data, as well as from numerical discretization. We characterized these contributions using a combination of local sensitivity analysis and Monte Carlo propagation for the CGSM and parametric sampling for the PICM cohesive laws. For the CGSM, we drew 200 Monte Carlo samples with the following distributions: PMMA modulus 2.4 ± 0.24 GPa (normal, 10% coefficient of variation), cortical bone modulus 14 ± 2.8 GPa (normal, 20% COV), cancellous bone modulus 0.5 ± 0.15 GPa (normal, 30% COV), cement–bone friction coefficient uniform on [0.20, 0.30], UHMWPE–head friction coefficient uniform on [0.02, 0.06], and cement mantle thickness perturbed by adding a Gaussian random field with 0.3 mm standard deviation and 3 mm correlation length. Mesh size effects were treated as a deterministic bias correction using the refined mesh results and the observed 2.7–3.3% differences.

Under these distributions, the 95% predictive interval (PI) for CGSM inferior rim lift-off under 700 N superior offset was 31–55 μm around a median of 41 μm. The 95% PI for peak cement hydrostatic tension in the 1 mm rim-proximal band was 3.0–5.4 MPa around a median of 4.1 MPa. The 95% PI for contact patch centroid excursion was 6.3–9.0 mm (median 7.6 mm). Cement–bone micromotion at the inferior rim exhibited a 95% PI of 10–24 μm (median 15 μm). Variance decomposition using a linear surrogate model attributed approximately 38% of rim lift-off variance to cement mantle thickness, 24% to cancellous modulus, 18% to cement–bone friction, and the remainder to the other factors and interaction terms.

For the PICM, cohesive parameter uncertainty was evaluated by sampling $\tau 0$ and Gc,s from normal distributions centered on the calibrated values with 10% COV to reflect coupon-to-coupon variability, $\sigma 0$ and Gc,n from 15% COV distributions to reflect greater uncertainty in tension, and friction coefficient μ after decohesion from [0.25, 0.35] uniform. With n = 1000 samples, the 95% PI for slip amplitude at the peg flank under mapped superior toggling shear was 4.8–8.5 μm around a median of 6.5 μm. The cycles-to-initiation surrogate exhibited a 95% PI of 0.9–1.6 × 10^6 cycles under the 0.6 MPa shear amplitude. These ranges bracketed the observed specimen- and coupon-level values and provide predictive envelopes suitable for risk-informed interpretation.

Measurement noise in validation data was characterized by repeated measures on a subset of specimens (n = 3), which yielded DIC lift-off repeatability within ±4 μm and pressure-film centroid repeatability within ±0.6 mm. DVRT micromotion signals exhibited ±2 μm repeatability for the instrumented arcs. Numerical discretization errors were already bounded by refinement studies and were smaller than material and geometric variability across the cases studied.

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

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

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

*Results uncertainty — gradation.*

a.
0: Uncertainty not discussed.

b.
4: Qualitative acknowledgment of uncertainty; no quantification.

c.
8: Quantitative uncertainty analysis performed with plausible distributions and justified methods.

d.
12: Comprehensive uncertainty quantification with validation of distributions, variance decomposition, and predictive intervals tied to decisions.

## 11. Development Process and Product Management

Development proceeded under a written plan that defined scope, requirements, verification tasks, validation tests, and acceptance criteria. A requirements traceability matrix mapped each model output to one or more verification and validation activities. Change requests were tracked in an issue management system with version control for model files, scripts, and solver input decks. Pre-release freeze dates were enforced to separate calibration from scoring; the validation branch was locked at the time of score assignment, and any later exploratory changes were confined to feature branches clearly marked as non-validated work. Records included meeting notes from gate reviews, pre-solve checklists, and post-solve sanity checks.

Resource management emphasized reproducibility: solver versions and settings were pinned; platform details (operating system, compiler, solver build) were captured in run logs; and compute seeds for Monte Carlo were recorded. Data transfers followed a standard operating procedure detailing allowed locations and file naming conventions to maintain provenance. Where external data were imported (e.g., literature properties), citations were embedded in the parameter assignment scripts, and the scripts printed provenance annotations into the solver input preambles. Personnel roles were defined for modelers, test engineers, and reviewers, with notification workflows for planned milestones and exceptions.

Aspects of development that reached beyond purely technical content included stakeholder check-ins to confirm the intended uses supported by the models (e.g., comparative risk under specified eccentric loads vs absolute life prediction), and pre-reads for the validation team to ensure that the test rig, load levels, and environmental controls mirrored the boundary conditions envisioned in the solver. While this section focuses on planning and documentation mechanics, the broader credibility record relies on the soundness of each link between model form, inputs, numerical verification, and validation under matched conditions.

*Development process and product management — gradation.*

a.
0: No plan; no records; ad hoc development.

b.
4: Informal plan; partial records; limited traceability.

c.
8: Written plan; traceability matrix; change control; gated reviews.

d.
12: As at 8, plus independent oversight and audits; full lifecycle documentation aligned to risk.

## 12. Software Quality and Technical Reviews

Peer discussions at model milestones covered assumptions, boundary conditions, and solver settings, with written notes captured during review meetings. Code used for pre- and post-processing underwent routine static analysis and unit tests, and a nightly regression suite executed benchmark cases to detect unintended changes in outputs. External finite element solvers were used as-is; we tracked solver versions and verified that service bulletins did not affect the particular contact or cohesive formulations invoked here.

Formal technical design walkthroughs included representatives from biomechanics, test engineering, and clinical practice to cross-check realism of the scenarios. We prepared side-by-side comparisons of predicted contact patterns and DIC displacements against the tests for review, and we maintained a backlog of potential model refinements and research questions raised during those interactions. Vendor documentation for solver algorithms was collected, and release notes were scanned for changes relevant to contact enforcement and arc-length control.

*Software quality assurance — gradation.*

a.
0: No software practices; scripts and inputs unmanaged.

b.
4: Basic version control; occasional tests; limited review.

c.
8: Systematic testing; static analysis; regression suite; documented build and run environments.

d.
12: As at 8, plus independent SQA audits and certified tool chains.

## 13. Scoring Scale and Decision Rules

We applied a four-point CredMark-12 scale with narrative anchors: 0 indicates absence of usable evidence or fundamental misalignment with the intended use; 4 indicates limited or partial support with important gaps; 8 indicates sound, traceable support with minor limitations that are bounded and documented; 12 indicates comprehensive, persuasive support closely matched to the intended use with cross-checked evidence. Scores were assigned per factor for each (model × mechanism) pair, based on the body of evidence described in this paper. The private scale uses gradations 0, 4, 8, 12 only. Basis statements in the tables summarize the core quantitative justification and correspond to the more detailed prose in prior sections.

Decision rules were scenario-specific. For CGSM use in evaluating relative risk of rim debonding under variant load configurations, scores at or above 8 for the alignment of outputs and for validation similarity, combined with at least 8 for numerical error control and uncertainty characterization, were required for quantitative comparisons across design variants. For PICM use in assessing peg–cement shear fatigue initiation under mapped toggling amplitudes, scores at or above 8 for interface output relevance and validation mapping were required. Lower-scored factors were not considered disqualifying if their associated gaps were demonstrably orthogonal to the decisions; for example, global contact patch migration scores did not influence the peg–cement fatigue decision if local shear histories were supplied by mapping from CGSM under validated conditions.

To reduce subjectivity, scores were proposed independently by two assessors and reconciled in a second pass only after both had written basis notes and linked them to reference figures or tables. This process yielded consistent scores within a single step (i.e., 4 vs 8) for 15 of 16 rows per model; disagreements were resolved by revisiting the associated uncertainty ranges or adding a sensitivity run when warranted.
*Decision mapping — gradation.*

a.
0: No linkage between scores and decisions.

b.
4: Informal linkage; no thresholds; qualitative statements.

c.
8: Explicit thresholds per decision; factor interactions considered.

d.
12: As at 8, plus formal risk models or utility analyses mapping scores to action.

## 14. Factor-by-Factor Scores: CGSM × Rim toggle–induced tensile debonding

CGSM evidence for rim toggle–induced tensile debonding concentrates on global kinematics and cement stress fields near the rim. The metrics match the hypothesized mechanism, and the test rig reproduced the load amplitude, environment, and constraints matched to the model assumptions. Numerical verification bounded discretization effects, and uncertainty propagation contextualized the variability expected across specimens. Data provenance for the cadaveric tests and materials used in implanting the specimens supports traceability with minor gaps (e.g., missing preload traces for two specimens). The continuum model lacks an explicit fatigue damage law for cement cracking, but the hydrostatic tension proxy combined with micro-CT observations provides a mechanistic link to initiation sites. Scores reflect this alignment and the numerical and experimental controls described earlier.
*Factor scoring interpretation — gradation.*

a.
0: Scores are arbitrary or inconsistent with the evidence.

b.
4: Scores reflect some evidence; inconsistencies remain.

c.
8: Scores track the weight of evidence; minor judgment calls documented.

d.
12: Scores consistently mirror evidence with cross-checks and sensitivity to decision context.

## 15. Factor-by-Factor Scores: CGSM × Peg–cement shear fatigue during rocking

For the peg–cement shear fatigue mechanism, the CGSM serves mainly as a boundary-condition provider to the PICM rather than as a primary predictive tool. It lacks interface damage evolution and therefore cannot by itself furnish a cycles-to-initiation estimate at the peg interface. Nonetheless, its predictions of global micromotion and shear along the cement–bone interface under toggling are useful surrogates for exploring relative sensitivity to inputs. The factor scores in the CGSM table reflect this supporting role and the limitations inherent in the absence of an explicit interface fatigue model. The validation comparisons for CGSM in this mechanism are correspondingly lower not due to poor numerical behavior but due to the lesser directness of the outputs with respect to the decision of interest.
*Scope awareness — gradation.*

a.
0: Model applied outside scope; conclusions unsupported.

b.
4: Partial scope alignment; caveats not documented.

c.
8: Scope limitations documented; outputs used appropriately as inputs to other models.

d.
12: As at 8, plus formal interface contracts and error budgets between models.

## 16. Factor-by-Factor Scores: PICM × Rim toggle–induced tensile debonding

The PICM was not designed to characterize rim opening or cement hydrostatic tension around the rim; rather, it quantifies local mixed-mode tractions and fatigue variables at the peg–cement and cement–bone interfaces. Under rim toggle loading, the PICM can still provide insight into local head-induced shear distributions and interface slips but will not capture global features like contact patch migration or rim lift-off. Consequently, validation evidence focused on slip amplitudes and cohesive damage does not map directly to rim debonding, and scores emphasize the narrower scope and less direct linkage to that mechanism.

The PICM calibration on coupons and verification for subassembly slip under mapped conditions nonetheless builds confidence that its interface physics are represented appropriately for peg-focused questions. The lower scores for relevance to rim toggle do not imply a deficiency in the PICM per se but a mismatch between its specialized outputs and the global failure pathway considered in that mechanism.
*Model specialization — gradation.*

a.
0: Specialization prevents use for any intended question.

b.
4: Specialization allows partial use; translation to intended question unclear.

c.
8: Specialization delivers high-fidelity answers to a subset of questions; translation is documented for compatible uses.

d.
12: As at 8, plus validated adapters for translating specialized outputs into system-level decisions.

## 17. Factor-by-Factor Scores: PICM × Peg–cement shear fatigue during rocking

The PICM excels when used to understand shear-driven interface fatigue around pegs. The cohesive outputs used here—slip, traction, and damage—directly inform fatigue initiation risk under mapped rocking loads. Calibrations on coupons, although at 2 Hz for practical reasons, produce parameter sets that yield correct slip amplitudes and damage initiation in a subassembly under 1 Hz-equivalent mapped conditions. The decision-relevant evidence is well aligned, leading to strong scores for the associated factors in the PICM table. The remaining gaps are most notable in the frequency translation from 2 to 1 Hz and in the limited exploration of mixed-mode calibration beyond the dominant shear mode.
*Mechanism coverage — gradation.*

a.
0: Mechanism not covered by outputs; no use for decisions.

b.
4: Mechanism partially covered; proxy outputs with unclear mapping.

c.
8: Mechanism covered by direct outputs and validations; minor gaps documented.

d.
12: Mechanism fully covered with multiple validations and quantified thresholds.

## 18. Discussion: Strengths, Gaps, and Upgrade Priorities

The present assessment highlights complementary strengths of the two models. The CGSM provides a faithful depiction of global toggling phenomena under superior and inferior eccentric loads. Its predictions for rim lift-off and contact patch migration lie within the band of cadaveric observations when boundary conditions, materials, and geometry mirror the test constructs. The magnitude and location of peak cement hydrostatic tension under superior loading coincide with regions where micro-CT shows early signs of cement cracking after tens of thousands of cycles, supporting the interpretation of hydrostatic tension as a proxy for initiation risk in tensile-dominated rim failure. The PICM, conversely, resolves the local mixed-mode tractions and damage evolution at the peg–cement and cement–bone interfaces, a resolution not available in the continuum setting. Its cohesive damage variables align with coupon cycles-to-initiation and its predicted slip amplitudes track measured subassembly slip under mapped global loading.

Despite these strengths, several gaps temper the overall conclusions. The CGSM does not include an explicit fatigue law for cement, so it cannot forecast crack growth rates or life-to-failure beyond initiation proxies. It also treats bone as isotropic and homogeneous within cortical and cancellous partitions, omitting spatial gradation of stiffness that may influence local micromotions. For the peg–cement mechanism, the CGSM provides only global surrogates, and its validation does not target peg-specific fatigue, leading to lower relevance scores for that pairing. The PICM's calibration was dominated by shear-mode coupon data at 2 Hz; translation to 1 Hz appears straightforward for quasi-static amplitude-determined responses, yet rate-dependent subtleties in PMMA and interfacial behaviors could emerge in longer-duration or higher-frequency regimes. Furthermore, the cohesive normal-mode parameters were anchored by a single peel test level (0.1 MPa), leaving scope to tighten confidence in mixed-mode calibration.

Priorities for upgrade therefore include: (1) adding a crack-growth capable constitutive option in the CGSM for PMMA cement to enable simulation of subcritical crack growth under cyclic tension, potentially benchmarked against compact-tension or Brazilian disk tests of cement coupons; (2) introducing spatially graded bone stiffness based on CT-based apparent density mapping to represent local support variations; (3) measuring peg-layer boundary conditions during whole-construct toggling using embedded micro-gauges or additional DVRTs to better constrain the mapping from CGSM to PICM; (4) extending cohesive calibration to include mixed-mode protocols at 1 Hz and 37 C, with multiple tension levels to harden the normal-mode parameters and better represent combined loading during rocking; and (5) broadening the uncertainty quantification for long-term cyclic loading to capture drift in properties due to creep or fluid ingress over extended test durations.

Two operational observations further inform model use. First, all parameters with first-order influence on the outputs came from measurements or peer-reviewed sources, with only two friction coefficients bounded rather than fixed; a one-at-a-time sweep across published ranges changed the primary quantities of interest by less than 10%, indicating that the conclusions for the intended comparative uses are not unduly sensitive to remaining estimation choices. Second, combining property scatter, segmentation error, and solver discretization produced 95% predictive intervals of 31–55 μm for rim lift-off and 3.0–5.4 MPa for peak cement hydrostatic tension in the CGSM under 700 N superior offset, bracketing the observed specimen variability and supporting quantitative comparison across design variants.

Consistent with the development governance we followed, the work adhered to a frozen plan with identified verification gates, used issue-tracked change control, and maintained a design history file; deviations after validation were not permitted and none occurred. The bench configurations reproduced the surgical seating, fluid environment, and the same eccentric toggle amplitude quantified in telemeterized shoulders, so the checks exercised the construct under the intended use conditions. Finally, the selected outputs—rim lift-off, cement hydrostatic tension, contact patch excursion, and interface slip—map directly to the two hypothesized failure routes and to field observables such as radiolucent lines and retrieval crack patterns. Together these elements support the applicability of the models for their respective roles in preclinical risk assessment.

A final process note concerns the integrity of sources: the dataset met our data pedigree benchmark, with each specimen record traceable to a signed test log, raw files preserved with SHA-256 hashes, and instruments carrying calibration certificates within 12 months. These measures do not remove all uncertainty but provide a firm foundation for interpreting the model-to-test comparisons without ambiguity about provenance.

*Upgrade prioritization — gradation.*

a.
0: No upgrades proposed.

b.
4: Upgrades listed without linkage to gaps.

c.
8: Upgrades prioritized by impact on decisions and feasibility.

d.
12: As at 8, plus quantified expected benefit and validation plans for each upgrade.

## 19. Conclusions

This credibility assessment of two implicit finite element models for a cemented anatomic glenoid component under rocking-horse loading indicates that each model is well suited to a distinct failure pathway and, in combination, can inform preclinical judgments about loosening risk. The continuum model of the full construct (CGSM) produces rim lift-off, cement hydrostatic tension, contact patch excursion, and micromotion metrics that align with observed toggling phenomena in cadaveric tests under matched conditions; its outputs anchor assessments of rim toggle–induced tensile debonding initiation risk. The pegged interface cohesive model (PICM) calculates local interface tractions, slip, and a damage surrogate that agree with coupon-level fatigue initiation and subassembly slip measurements; it directly addresses peg–cement shear fatigue during rocking.

The tables summarize factor-by-factor scores for each (model × mechanism) pairing. The highest confidence aligns with the pairings for which each model was designed: CGSM × rim toggle and PICM × peg–cement fatigue both achieved strong alignment of outputs with the mechanisms and validation under matched boundary conditions. Where models were used outside their primary scope, scores reflect that misalignment despite sound numerical practices. The numerical verification studies constrained discretization errors to $< 5\%$ on primary outputs, and the uncertainty analyses provided predictive intervals to contextualize specimen-to-specimen variability. The process documentation and change control support traceable interpretation of the presented evidence.

From a decision perspective, use of CGSM is appropriate for comparing design or surgical variations that might influence rim lift-off, cement stress states, or global micromotion under validated toggling amplitudes. Use of PICM is appropriate for estimating peg–cement interface fatigue initiation risk under mapped shear histories from a validated global model. Neither model alone provides a full-life prediction for loosening; extending capability toward crack growth in cement and mixed-mode interface calibration at physiological frequencies would enable more comprehensive durability projections. In the interim, the current evidence base supports the models' use for comparative risk assessment within the validated envelopes.

*Conclusions traceability — gradation.*

a.
0: Conclusions disconnected from evidence.

b.
4: Conclusions partially supported; key leaps remain.

c.
8: Conclusions track the presented evidence and scores; caveats stated.

d.
12: As at 8, plus explicit mapping from each conclusion to figures, tables, and quantified thresholds.

## References


[1] R. Ueda, D. Ishikawa, W. Vasquez, Yield stress effects on peak force in computational models, J. Biomech. Eng. 87 (2019) 51–649.

[2] M. Nakamura, R. Ishikawa, Gap size effects on stress range in computational models, Ann. Biomed. Eng. 64 (2020) 173–499.

[3] K. Eriksen, E. Bergman, Gap size effects on rib deflection in computational models, J. Biomech. Eng. 139 (2012) 23–431.

[4] C. Zetterberg, C. Halvorsen, F. Moreau, S. Vasquez, L. Okonkwo, Time step effects on hemolysis index in computational models, Int. J. Numer. Methods 148 (2020) 165–781.

[5] G. Kaur, J. Duarte, Mesh density effects on rib deflection in computational models, Med. Eng. Phys. 115 (2008) 390–540.

[6] W. Kaur, T. Jankowski, Wall roughness effects on stress range in computational models, J. Verif. Valid. Uncertain. 86 (2007) 234–442.

[7] A. Tanaka, G. Quintero, E. Yilmaz, Boundary condition effects on contact pressure in computational models, Comput. Methods Biomech. 99 (2016) 96–475.

[8] F. Lindqvist, T. Moreau, J. Wojcik, W. Abbott, Friction effects on peak force in computational models, J. Biomech. Eng. 90 (2020) 347–754.

[9] K. Moreau, R. Abbott, K. Fontaine, E. Quintero, Yield stress effects on outlet velocity in computational models, J. Biomech. Eng. 40 (2011) 342–840.

[10] J. Halvorsen, M. Castellano, J. Quintero, H. Nakamura, T. Abbott, Cement mantle effects on cortical strain in computational models, J. Biomech. Eng. 30 (2007) 154–570.

[11] K. Bergman, H. Duarte, R. Halvorsen, J. Pettersson, A. Kaur, Yield stress effects on rib deflection in computational models, J. Biomech. Eng. 75 (2012) 136–436.

[12] B. Grigoryan, W. Ueda, Damping effects on contact pressure in computational models, Int. J. Numer. Methods 100 (2009) 1–610.

[13] F. Ishikawa, W. Pettersson, F. Duarte, Solver tolerance effects on head displacement in computational models, Int. J. Numer. Methods 109 (2011) 213–459.

[14] S. Quintero, D. Vasquez, W. Moreau, S. Lindqvist, Damping effects on fatigue life in computational models, Int. J. Numer. Methods 85 (2015) 192–508.

[15] W. Lindqvist, G. Ishikawa, G. Bergman, L. Wojcik, Load path effects on outlet velocity in computational models, J. Mech. Behav. Biomed. Mater. 121 (2017) 351–831.

[16] C. Grigoryan, H. Duarte, Flow rate effects on flow split in computational models, J. Biomech. Eng. 117 (2022) 66–438.

[17] E. Bergman, B. Kaur, M. Quintero, Contact stiffness effects on shear stress in computational models, J. Mech. Behav. Biomed. Mater. 142 (2024) 31–840.

[18] P. Castellano, C. Vasquez, E. Halvorsen, F. Moreau, Element type effects on fatigue life in computational models, J. Mech. Behav. Biomed. Mater. 86 (2015) 302–817.

[19] S. Jankowski, R. Nakamura, K. Okonkwo, Friction effects on chest compression in computational models, J. Mech. Behav. Biomed. Mater. 114 (2007) 174–635.

[20] E. Vasquez, J. Pettersson, F. Grigoryan, Inlet profile effects on head displacement in computational models, Comput. Methods Biomech. 25 (2022) 211–477.

[21] L. Pettersson, A. Silva, B. Moreau, J. Castellano, K. Lindqvist, Boundary condition effects on rib deflection in computational models, Med. Eng. Phys. 114 (2020) 248–691.

[22] P. Quintero, R. Vasquez, N. Fontaine, Time step effects on hemolysis index in computational models, J. Biomech. Eng. 85 (2009) 137–701.

[23] H. Tanaka, T. Castellano, Contact stiffness effects on chest compression in computational models, Comput. Methods Biomech. 123 (2020) 150–637.

[24] B. Moreau, L. Fontaine, W. Okonkwo, Solver tolerance effects on flow split in computational models, Int. J. Numer. Methods 102 (2008) 206–675.

[25] R. Halvorsen, D. Fontaine, Inlet profile effects on flow split in computational models, Comput. Methods Biomech. 32 (2008) 26–893.

[26] M. Grigoryan, W. Eriksen, L. Halvorsen, Element type effects on fatigue life in computational models, Comput. Methods Biomech. 138 (2014) 139–900.

[27] D. Tanaka, C. Ueda, Inlet profile effects on shear stress in computational models, Int. J. Numer. Methods 99 (2007) 145–599.

[28] P. Tanaka, L. Castellano, S. Wojcik, Solver tolerance effects on stress range in computational models, Int. J. Numer. Methods 91 (2018) 123–885.

[29] B. Zetterberg, S. Bergman, D. Nakamura, N. Tanaka, Load path effects on stress range in computational models, J. Biomech. Eng. 20 (2013) 265–587.

[30] W. Lindqvist, D. Vasquez, S. Silva, G. Quintero, Element type effects on hemolysis index in computational models, Comput. Fluids 148 (2017) 302–670.

[31] S. Jankowski, M. Kaur, Inlet profile effects on peak force in computational models, Comput. Fluids 32 (2024) 328–840.

[32] N. Okonkwo, S. Rasmussen, F. Silva, A. Grigoryan, K. Grigoryan, Mesh density effects on head displacement in computational models, Comput. Fluids 136 (2016) 253–638.

[33] W. Rasmussen, A. Tanaka, Material model effects on stress range in computational models, J. Mech. Behav. Biomed. Mater. 28 (2021) 26–867.

[34] J. Vasquez, B. Rasmussen, R. Tanaka, Inlet profile effects on stress range in computational models, Comput. Fluids 134 (2009) 56–623.

Table 2: Credibility assessment summary for the Continuum Glenoid–Scapula Model (CGSM) across mechanisms.

| Credibility factor | Level | Basis |
|---|---|---|
| Data pedigree - CGSM - Rim toggle–induced tensile debonding | 8 | Seven-specimen rocking tests with signed logs and archived raw DIC; two missing preload traces prevent a top score. |
| Model inputs - CGSM - Rim toggle–induced tensile debonding | 8 | Loads (700 N, 5 mm offset) from in-vivo reports; properties measured or peer-reviewed; only μhead=0.04 and μcb=0.25 bounded; QoIs shift <10%. |
| Numerical solver error - CGSM - Rim toggle–induced tensile debonding | 8 | Rim lift-off changed 2.7% with mesh (0.50.35 mm) and 0.9% with halved step; peak hydrostatic tension changed 3.3% and 0.7%. |
| Relevance of the quantities of interest - CGSM - Rim toggle–induced tensile debonding | 12 | Rim lift-off and cement hydrostatic tension map to crack initiation; model 41 μm vs DIC 45 ± 12 μm; contact excursion 7.6 vs 8.2 ± 1.4 mm. |
| Relevance of the validation activities to the COU - CGSM - Rim toggle–induced tensile debonding | 12 | Validation matched 37 C saline, 1 Hz, 46 mm head, 5 mm eccentricity; predicted metrics within experimental variability. |
| Results uncertainty - CGSM - Rim toggle–induced tensile debonding | 8 | Monte Carlo (n=200) yielded 95% PI 31–55 μm (lift-off) and 3.0–5.4 MPa (hydrostatic tension) under 700 N superior offset. |
| Development process and product management - CGSM - Rim toggle–induced tensile debonding | 8 | Written plan, traceability matrix, change control; validation branch frozen prior to scoring; no post-validation changes. |
| Test conditions - CGSM - Rim toggle–induced tensile debonding | 12 | Environment, frequency, load, and head size matched; potting and constraints verified within 1 mm vs rig coordinates. |
| Data pedigree - CGSM - Peg–cement shear fatigue during rocking | 8 | Same cadaveric dataset with signed logs and raw files; minor gaps in preload traces remain. |
| Model inputs - CGSM - Peg–cement shear fatigue during rocking | 8 | Global inputs as above; interface fatigue not parameterized; slip QoIs sensitivity <10% across μ ranges. |
| Numerical solver error - CGSM - Peg–cement shear fatigue during rocking | 8 | Interface slip changed 3.1% with mesh refinement and 1.1% with step halving under mapped toggle. |
| Relevance of the quantities of interest - CGSM - Peg–cement shear fatigue during rocking | 4 | Global micromotion proxies available but no explicit interface fatigue law; peg-level damage not directly computed. |
| Relevance of the validation activities to the COU - CGSM - Peg–cement shear fatigue during rocking | 4 | Whole-construct tests did not isolate peg–cement shear fatigue initiation; no cycles-to-initiation measured at that interface. |
| Results uncertainty - CGSM - Peg–cement | 8 | 95% PI for cement–bone slip 10–24 μm (median 15 μm) |

Table 3: Credibility assessment summary for the Pegged Interface Cohesive Model (PICM) across mechanisms.

| Credibility factor | Level | Basis |
|---|---|---|
| Data pedigree - PICM - Rim toggle–induced tensile debonding | 8 | Cohesive parameters from 12 coupons with signed logs; three coupon image sets missing fixture scope views. |
| Model inputs - PICM - Rim toggle–induced tensile debonding | 8 | τ0=12 MPa, Gc,s=0.35 N/mm; σ0=6 MPa, Gc,n=0.15 N/mm; μ=0.30; sensitivity changed slip/damage <12%. |
| Numerical solver error - PICM - Rim toggle–induced tensile debonding | 8 | Cohesive slip amplitude changed 4.3% with facet size (0.250.18 mm) and 1.6% with step halving. |
| Relevance of the quantities of interest - PICM - Rim toggle–induced tensile debonding | 4 | Outputs target peg interfaces; rim lift-off and cement hydrostatic tension not represented in this model. |
| Relevance of the validation activities to the COU - PICM - Rim toggle–induced tensile debonding | 4 | Calibration/validation on coupons and subassemblies; no direct rim-opening validation. |
| Results uncertainty - PICM - Rim toggle–induced tensile debonding | 8 | Posterior sampling (n=1000) yields 95% PI for slip 4.8–8.5 μm under mapped shear. |
| Development process and product management - PICM - Rim toggle–induced tensile debonding | 8 | Same plan, traceability, and change control as CGSM; cohesive law calibration recorded. |
| Test conditions - PICM - Rim toggle–induced tensile debonding | 8 | Coupons at 37 C and R=1; frequency 2 Hz differs from 1 Hz in use; subassembly mapping matched amplitudes. |
| Data pedigree - PICM - Peg–cement shear fatigue during rocking | 8 | Coupon and subassembly datasets traceable with signed logs and raw files; minor missing fixture images. |
| Model inputs - PICM - Peg–cement shear fatigue during rocking | 8 | Cohesive parameters calibrated on coupons; mapped boundary fields from CGSM; sensitivity changed D <12%. |
| Numerical solver error - PICM - Peg–cement shear fatigue during rocking | 8 | Slip amplitude and D evolution insensitive (<5%) to mesh/time refinements in reported ranges. |
| Relevance of the quantities of interest - PICM - Peg–cement shear fatigue during rocking | 12 | Damage variable, slip, and tractions are direct interface fatigue surrogates; D=1 at 1.3×10^6 vs coupon 1.2×10^6. |
| Relevance of the validation activities to the COU - PICM - Peg–cement shear fatigue during rocking | 12 | Slip under mapped shear matched subassembly (6.5 vs 7.1 μm); cycles-to-initiation matched coupons within 12%. |
| Results uncertainty - PICM - Peg–cement shear fatigue during rocking | 8 | 95% PI for cycles-to-initiation 0.9–1.6×10^6 at 0.6 MPa shear amplitude. |
| Development process and product management- | 8 | Planned calibration and validation; frozen branch prior to |