Credibility assessment under ASME V&V 40 of two implicit FEA models for screw load sharing in an orthopaedic locking plate

Evan J. Marshall[a], Priya K. Desai[1], Luca R. Benedetti[1], Hannah M. Ortiz[1]

[a] *Computational Biomechanics Group, Center for Orthopaedic Engineering, Meridian Biomedical Research Institute, Boston, MA, USA,*

Abstract

Background: Locking plate constructs distribute load among screws in patterns that depend on plate geometry, screw placement, and external loading. Designers increasingly use finite element analysis (FEA) to predict per-screw axial load fractions and the plate's force distribution during bench tests. For decisions that screen new designs prior to costly physical testing, ASME V&V 40 recommends matching model credibility to decision risk and model influence. Objectives: We evaluated two implicit FEA models intended for this screening task under a structured credibility plan that only considered the following factors: Equivalency of input parameters, Model inputs, Numerical code verification, Numerical solver error, Test conditions, and Test samples. The plan deliberately excluded all other factors in the standard. Methods: The device of interest was a 3.5 mm locking compression plate spanning a 10 mm mid-shaft gap fixed with locking screws in near and far cortical purchase. Model S (3D solid-contact) represented plate, screws, and surrogate bone with solid elements and nonlinear contact with friction; screw pretension was applied as bolt load. Model R (beam–spring surrogate) used beam elements for plate and screws and nonlinear axial springs for screw–bone engagement. Three mechanisms were examined: (i) four-point bending across an 80 mm support span with 40 mm load spacing, (ii) axial compression with 5 mm medial eccentricity, and (iii) pure torsion about the long-bone axis under rotational control at 0.5 deg/s. Outputs of interest were per-screw axial load fractions and plate force distribution. Evidence for the six credibility factors was assembled and scored on a 0–7 Credibility Attainment Scale (CAS), where 0 denotes no relevant evidence and 7 denotes comprehensive evidence sufficient for high-risk decisions. Predefined acceptance criterion for this high-risk context required CAS ≥ 6 for all factors across all mechanisms for at least one model. Results: In four-point bending, Model S achieved CAS 6 for the mapping of the apparatus to the model, the selection and provenance of physical properties, and implementation checks, but solver-related variability remained at the 5 level due to the need for mild contact stabilization; sample availability reached level 5. In eccentric compression, Model S exhibited contact-induced variability of approximately 8% and incomplete representation of fixture flexibility, yielding solver and apparatus-mapping levels of 4 and 5, respectively. In torsion, Model S encountered increased solver variability (approximately 12% change with refinement at higher twist) and incomplete modeling of clamping flexibility, yielding lower levels (3–5) for some factors. Model R demonstrated strong algorithmic robustness and reproducibility (levels of 6 for solver behavior and implementation checks) but lacked a faithful mapping of the apparatus for eccentric compression and torsion and did not recalibrate spring parameters for combined loading, yielding levels 2–4 on those dimensions. Conclusion: The acceptance criterion was not met. Neither model attained CAS ≥ 6 simultaneously for all six factors across all three mechanisms. The principal deficits were insufficient mapping of the bench setups to the surrogate model and missing mechanism-specific parameter identification in Model R for eccentric compression and torsion, and excessive solution variability under torsion in Model S. For the stated high-risk context of use, the models are Not accepted.

Keywords: ASME V&V 40, orthopaedic plate FEA, screw load sharing

1. 1. Introduction

Locking plate constructs permit rigid angular fixation of diaphyseal fractures while preserving periosteal blood supply. The distribution of internal force among screws and the plate during standard bench tests is closely linked to construct stiffness, the potential for localized fastener overloading, and overall risk of mechanical failure during the early phases of design iteration. Prediction of per-screw axial load fractions and the plate's force distribution using computational models is increasingly common in design screening before physical prototypes are committed to bench testing.

ASME V&V 40 provides a risk-informed framework for determining whether a computational model is credible for a specified context of use. The present assessment implements that framework to appraise two implicit finite element models intended to predict screw load sharing in a 3.5 mm locking compression plate spanning a mid-shaft gap. We restrict our analysis to the set of credibility dimensions specified a priori and deliberately refrain from offering judgments on any aspects outside this defined set.

Two models that aim to predict per-screw axial and shear force fractions under bench-test mechanisms were examined. The first, referred to here as Model S, is a three-dimensional solid-contact finite element model that explicitly resolves plate, screws, and surrogate bone with nonlinear contact and friction, including bolt preload to represent screw pretension. The second, Model R, is a reduced-order surrogate in which the plate and screws are represented by beam elements and screw–bone engagement is captured through nonlinear axial springs calibrated to reflect axial engagement stiffness.

Three loading mechanisms reflect common bench configurations used for preliminary mechanical characterization of long-bone plating constructs: (i) four-point bending across a 10 mm gap between surrogate bone ends, (ii) axial compression with a medial eccentricity, and (iii) pure torsion about the long-bone axis. In each case, the outputs of interest are the per-screw axial load fractions and the plate's longitudinal force distribution under the specified external load.

The context of use for both models is to support down-selection of new plate configurations by predicting per-screw axial load fractions and plate force distribution for the three bench tests prior to physical testing. Given that the outcome of this analysis could advance or halt a design toward verification testing and that the model outputs directly determine which screw layouts proceed, the decision consequence and model influence are both high. Consequently, this assessment is structured to meet the demands of a high-risk decision under the standard.

The credibility plan applies a 0–7 Credibility Attainment Scale (CAS), where 0 denotes no relevant evidence and 7 denotes comprehensive evidence consistent with high-risk decisions. We summarize evidence across six specified dimensions that together govern whether the models can be relied upon for the stated task. A predefined acceptance criterion requires CAS at or above 6 for all of these dimensions and across all three bench mechanisms for at least one of the two models. Failing to reach that bar yields a not accepted outcome for the stated use.

The remainder of this article first describes the device and the computational models, then the three mechanisms, followed by the context of use and risk rationale, the credibility plan and scoring scale, and the supporting evidence from physical tests and numerical verification activities. We then present the results as model-by-mechanism summaries of CAS levels, discuss why certain dimensions fell short, and close with an explicit decision statement aligned to the predefined criterion.

Numerical code verification — gradation.

a.
No explicit checks that implemented elements or contact reproduce known solutions.

b.
Basic spot checks of element behavior on simple benchmarks with qualitative agreement.

c.
Systematic benchmark suite demonstrating quantitative agreement within specified tolerances across relevant loading states.

d.
Traceable, peer-reviewed benchmark coverage of all formulations used with error metrics below pre-set thresholds in the domains of interest.

2. 2. Device and computational models

The device is a 3.5 mm locking compression plate used for diaphyseal fixation, bridging a 10 mm mid-shaft gap between surrogate bone segments. The plate is a stainless-steel construct with a symmetric hole pattern compatible with 3.5 mm locking screws. The study configurations employed near and far cortical purchase according to the manufacturer's drilling recommendations for mid-diaphyseal application. All plates and screws used in the physical tests were taken from a single manufacturing lot to minimize variability associated with production tolerances.

Surrogate bones were machined from homogeneous polymer cylinders to provide repeatable substrate stiffness and geometry. Each bone segment was turned to match the plate's intended curvature plane with a nominal outer diameter selected to match standard sawbone dimensions for mid-shaft applications. After machining, a 10 mm axial gap was introduced between segments to represent a comminuted fracture scenario where the plate dominates load transfer across the span. The bone surrogates were drilled with consistent hole spacing matching the plate template to position screws with repeatable orthogonal trajectories into both cortices.

Model S creates a high-fidelity representation of this construct. The plate is meshed with quadratic tetrahedral or hexahedral solid elements, depending on local geometric features, with mesh density refined in regions near the screw holes. Screws are modeled as separate solid parts with threaded regions idealized to smooth shanks where necessary to reduce element distortion, while preserving thread pitch in regions where thread–bone interaction is expected to affect contact areas materially. The surrogate bone is represented as a solid with elastic isotropy. All interfaces between plate, screw heads, screw shafts, and bone are defined by surface-to-surface contact with a penalty formulation and friction coefficients measured on representative material pairs. Screw pretension is introduced as a bolt load in the solver to represent installed torque converted to axial preload, applied uniformly at assembly and maintained during loading.

Model R replaces geometric and contact complexity with a reduced-order representation. The plate is modeled as a Timoshenko beam line with sectional properties derived from the plate's CAD geometry. Each screw is a beam connected to the plate node at the corresponding hole, with beam orientation matching the intended drilling axis through the plate hole. Engagement between screw and bone is represented by a nonlinear axial spring that captures clamp-up stiffness along the screw axis; laterally, the screw head-to-plate angular constraint is assumed rigid for locking screws, while no torsional coupling through the thread flank is included. The springs exhibit piecewise linear stiffness profiles identified from prior calibrations that fit force–displacement data from isolated screw pull-up and unloading experiments in the same surrogate material.

Both models were built and solved in widely used commercial finite element software with static implicit solvers. For Model S, contact stabilization parameters and automatic time stepping were used judiciously, with manual control introduced

Figure 1: Overview of the orthopaedic plate geometry, load cases and the mechanisms assessed, shown across the full page width.

when needed to enforce consistent convergence patterns during the higher-curvature portions of the load–displacement curve, particularly in torsion. For Model R, load steps were applied in a purely linear setting for bending and torsion, while in eccentric compression steps were augmented to incorporate the effect of the 5 mm offset on bending-induced axial screw forces. In all simulations, outputs of interest were extracted as reaction forces at screw–plate connections (Model R) or as integral axial contact forces at screw–bone interfaces (Model S), normalized to report each screw's share of the total axial screw-group force.

A critical theme in device-model alignment is the consistent representation of screw pretension. For Model S, torque applied during assembly in the physical tests was converted to bolt preload through a calibration function obtained from instrumented fastener coupons tightened into the same plate and surrogate bone system. The function was then applied to assign initial bolt loads to each modeled screw. For Model R, screw pretension increases the clamp force and thereby affects the axial engagement stiffness of the screw–bone system, which in this surrogate was captured indirectly by setting the initial spring stiffness segment to reflect the preloaded engagement stiffness in the polymer.

The computational cost differed markedly between the two models. A typical Model S run with full contact for six screws and a 10 mm bone gap involved approximately 1.2 million degrees of freedom and required several hours to converge a complete load history for the highest loads in torsion. In contrast, Model R had on the order of a few thousand degrees of freedom and ran in seconds, making it attractive for rapid parametric explorations provided that the mapping of bench conditions to the model is faithful and parameter identification is appropriate for the mechanisms of interest.

Because the decision at stake is to down-select candidate plate configurations ahead of bench testing, it is not enough for a model to be computationally efficient; it must also produce credible predictions of how axial forces are shared among screws across all three mechanisms of interest. The details given above provide the basis for the subsequent evaluation under a set of credibility dimensions tailored to the stated context of use.

Model inputs — gradation.

a.
Assumed material and contact parameters without provenance or stated tolerance.

b.
Parameters drawn from literature sources for similar systems with limited justification.

c.
Parameters measured on representative coupons or provided by supplier with tolerances, and uncertainties acknowledged.

d.
Parameters measured and cross-validated under mechanism-relevant conditions with quantified uncertainties propagated to outputs of interest.

3. 3. Mechanisms of interest

Three laboratory mechanisms were chosen to probe distinct load paths in the plated construct: four-point bending, axial compression with a small medial offset, and pure torsion. Together, these mechanisms interrogate plate-dominant bending, combined compression–bending that elevates demand on screws near the gap, and twisting that can induce differential shear transfer and warping through the screw group.

In four-point bending, the construct is supported on two lower anvils with an 80 mm span and loaded from two upper anvils with 40 mm spacing. This configuration creates a nearly constant bending moment over the inner span that includes the 10 mm gap. For the present work, load was applied quasi-statically in displacement control to prescribed levels that remained in the range where the surrogate materials respond elastically and no irreversible damage occurs at contact interfaces. The key output is how the screw group shares axial force under the resulting plate-dominant bending across the gap.

Axial compression with a 5 mm medial offset was achieved by introducing a spacer that shifted the line of action of the compressive load relative to the surrogate bone axis. The compression load was applied at the proximal end while the distal end was supported to maintain alignment of the bone surrogate segments across the gap. This mechanism introduces a bending component superimposed on axial compression, with the effect of amplifying axial forces in screws closest to the gap on the compressive side. The loading regimen used monotonic displacement control, and instrumentation ensured that the eccentricity remained at 5 mm throughout the load history.

Pure torsion about the long-bone axis was imposed using rotational control at one end while the other end was held to prevent axial translation. The apparatus provided a rotational rate of 0.5 deg/s to avoid dynamic effects, and the angle was increased to prescribed levels that remained below any thresholds that would cause thread slip or bearing damage in the surrogate system. Torsion engages non-intuitive load transfer pathways within the screw group, including local warping and frictional constraints at the screw head–plate interface. Because Model R represents screw engagement via axial springs, it cannot directly represent off-axis coupling and warping modes that may contribute to force sharing under twist.

For all mechanisms, outputs of interest were (i) per-screw axial load fractions, defined as each screw's axial force divided by the sum of the axial forces over all screws in the group, and (ii) the plate's force distribution along its length, which in Model S was obtained from section forces and in Model R from beam element forces. These outputs were extracted at matched external load levels to enable meaningful comparison among models and physical tests, though direct model-to-test comparisons are not part of the present article's scope by design.

To ensure that the simulations track the benches, the boundary conditions were crafted to mirror the contact surfaces, spans, offsets, clamp constraints, and rotational controls of the laboratory fixtures. In Model S, this included representing the anvil–plate interaction surfaces during four-point bending with appropriate contact, while in Model R, equivalent nodal constraints reproduced the positions and directions of supports and loads. Later sections will explain where these translations were complete and where they were approximate.

Equivalency of input parameters — gradation.

a.
No explicit translation of laboratory geometry or load path into model boundary conditions.

b.
Partial translation of key spans, offsets, or constraints with simplifying assumptions left undocumented.

c.
Complete translation of spans, offsets, supports, rates, and pretension mapping with documented approximations to fixture flexibility.

d.
Traceable one-to-one mapping of all laboratory features into the model, including compliance of fixtures, with sensitivity studies showing negligible impact of remaining approximations.

4. 4. Context of use and model risk rationale

The intended use of the computational models is to predict per-screw axial load fractions and the plate's force distribution in three standard bench-test mechanisms for a new configuration of a 3.5 mm locking compression plate. These predictions are used to down-select among alternative designs before committing to physical prototype construction and testing. In this role, the models directly inform whether a candidate configuration advances or is halted prior to verification testing, placing the predictions squarely in the decision path.

In terms of risk classification under the standard, this context exhibits high consequence. If the model's prediction of load sharing is materially inaccurate, a configuration that should be eliminated may be advanced, incurring cost and potentially masking an unfavorable load concentration on a subset of screws; conversely, a design that would have performed adequately could be discarded prematurely, delaying innovation. The models' influence is also high, because the predictions in question directly determine whether and how a design variant proceeds, rather than serving as an auxiliary information source among many.

Given these two dimensions—consequence and influence—the risk category is high. This categorization sets a correspondingly high bar for evidence across the chosen credibility dimensions, motivates stringent acceptance criteria, and guides the level of detail expected of numerical verification and mapping of laboratory conditions into the computational representations.

Accordingly, we adopted a pre-specified acceptance criterion: for at least one of the two models, all six credibility dimensions considered here must reach a CAS of 6 or higher across all three mechanisms for the model to be accepted for the stated use. Failure to attain that level on any dimension in any mechanism leads to a decision of not accepted. This criterion is intentionally conservative for a high-risk context and closes the gap between aspirational model capabilities and the robustness required to guide design screening.

The remainder of this paper details the evidence assembled for each dimension and presents the model-by-mechanism CAS assignments underlying the final decision.

Test conditions — gradation.

a.
Laboratory protocols not documented or not linked to the simulations.

b.
General description of the rigs and loads without quantitative specification of spans, offsets, or control modes.

c.
Quantitative protocols for spans, offsets, rates, and alignment documented and used to guide simulation setups.

d.
Full protocol traceability with measurement of fixture stiffness and alignment used to parameterize model boundary conditions.

5. 5. Credibility plan and scoring scale

The credibility plan was tailored to the needs of high-stakes design screening and focused on six dimensions: the mapping of laboratory features and applied loads into model boundary conditions; the provenance and appropriateness of physical property and contact parameters; checks that the implemented elements and contact formulations reproduce known solutions; quantification of the residual error arising from discretization, nonlinear solution strategies, and step size; quantitative documentation of the laboratory apparatus, protocols, and control; and the size and representativeness of the specimens used to generate the physical evidence base. These choices were made at the outset and constrained the scope of evidence collection and reporting.

We adopted a 0–7 Credibility Attainment Scale. Level 0 indicates that no relevant evidence was available for the dimension in question. Level 3 may correspond to scattered or qualitative support. Level 5 indicates reasonably complete evidence appropriate for moderate-risk decisions, whereas level 6 signifies strong, quantitative, and mechanism-specific evidence suitable for high-risk decisions. Level 7 requires comprehensive, mechanism- and device-specific evidence with uncertainties quantified and propagated through to the outputs of interest. Intermediate levels reflect graded improvements in the quality, completeness, and relevance of evidence.

For each model and each mechanism, we assigned CAS levels across the six dimensions. We then compared those assignments to the pre-specified acceptance criterion requiring at least one model to achieve level 6 or higher across all dimensions and mechanisms. The overall decision reported at the end of the paper flows directly from this criterion and the compiled evidence.

The plan also specified a commitment to present model-by-mechanism summaries of CAS levels in a tabular form that enumerates the full set of assessed dimensions for each model. Because partial listings can obscure weaker dimensions, the tables were designed to include all six dimensions for each mechanism so that the reader can see the complete pattern of strengths and shortfalls.

Equivalency of input parameters — gradation.

a.
Rigs and loads are only described qualitatively with no direct translation into model boundary conditions.

b.
Key spans and offsets are mapped to model constraints but some supports or interactions are omitted without analysis.

c.
All spans, offsets, load paths, and pretension are mapped; approximations to fixture flexibility are disclosed and judged negligible.

d.
All laboratory features including fixture and clamp compliance are represented or bounded by sensitivity studies to demonstrate negligible impact on outputs of interest.

6. 6. Evidence acquisition—Test samples and test conditions

Physical testing was conducted on machined polymer bone surrogates paired with commercial 3.5 mm locking plates and screws taken from one manufacturing lot. The bone surrogates were prepared on a CNC lathe and mill to achieve concentricity and parallelism between faces, with dimples and pilot holes machined to maintain consistent screw angles and positions across specimens. Dimensional checks with a coordinate measuring machine confirmed that diameters were within ±0.15 mm and hole positions within ±0.10 mm of nominal.

In four-point bending, six constructs were tested (n = 6). The support span was set to 80 mm using a fixed-position lower anvil assembly, and the upper anvils were positioned to achieve a 40 mm load spacing centered over the 10 mm gap. The machine applied load in displacement control at a rate chosen to keep the loading quasi-static. The plate was oriented so that the holes coincided with the neutral axis of the surrogate bone, minimizing unintended torsion. A high-resolution load cell and displacement transducer recorded reaction forces and midspan deflection, respectively.

For axial compression with 5 mm medial eccentricity, five constructs were tested (n = 5). A spacer block introduced the offset, and laser alignment tools verified that the compressive line of action was maintained throughout the test. The distal end of the construct was supported to allow small rotations while preventing axial translation. The proximal loading platen's stiffness was characterized on the test frame and, in the physical tests, contributed a small compliance that, while repeatable, was not directly represented in the simplest numerical boundary conditions.

In pure torsion, five constructs were tested (n = 5). The proximal end was rotated at 0.5 deg/s using an angular actuator while the distal end was held against axial translation. An anti-backlash arrangement was used in the clamping fixtures to avoid free play at the start of rotation. Twist angle and torque were recorded synchronously. The torsion tests were restricted to angles for which the surrogate material and screw engagement remained in the linear range, a regime deliberately selected to avoid damage accumulation and to keep the focus on load sharing patterns rather than failure.

Environmental conditions were stable across all tests. The laboratory maintained ambient temperature at 21–23 °C and relative humidity between 30% and 50%. All assembly procedures were performed by a single experienced technician following a written protocol specifying hole drilling, plate positioning, and target installation torque. A calibrated torque wrench was used to set screws; a separate calibration on instrumented fastener coupons established a mapping between torque and axial preload in the particular plate–screw–surrogate system considered here.

To support subsequent mapping of bench features into the models, we documented in detail the spans, offsets, clamp geometries, and instrumentation. Photographs and CAD representations of the rigs were archived and annotated. For Model S, these details were used to create explicit contact surfaces representing anvils and clamps where appropriate. For Model R, the boundary conditions were distilled into nodal constraints that matched support locations and directions of applied loads while treating the contact patches as idealized lines or points, as is conventional for line-beam models.

Test samples — gradation.

a.
Specimens of unknown origin and tolerances; no control over lots or machining.

b.
Specimens from mixed lots or vendors with partial

documentation of geometry.

c.
Single-lot specimens with documented machining tolerances and sample sizes planned to achieve repeatability.

d.
Single-lot or statistically characterized lots with full traceability, dimensional verification, and uncertainty analysis of geometric variability propagated to outputs of interest.

7. 7. Verification activities—Numerical code verification and numerical solver error

We performed element- and assembly-level checks to ensure that the implemented formulations reproduced known solutions. For Model S, element patch tests confirmed correct recovery of constant strain states in the solid elements. A cantilever beam in bending was reproduced using both solid and shell idealizations, with tip deflections matching closed-form values to within 2%. Torsion of a circular shaft modeled with solid elements reproduced the Saint-Venant solution to within 1–2% across mesh refinements. Contact algorithms were challenged using a block-on-plate test with friction to verify that stick–slip transitions occurred at the expected load ratios based on the specified coefficients.

For Model R, beam element checks included recovery of exact solutions for bending and torsion of prismatic members, with tip deflections and twist matching textbook values to within machine precision for linear cases. Spring assemblies were verified using a series configuration test and against a piecewise linear target response, ensuring that the intended axial stiffness segments were realized. The compatibility of beam-to-spring connections at the screw nodes was examined to make sure that the constraints did not inadvertently add stiffness in unintended directions.

Quantification of residual numerical error associated with discretization and nonlinear solution strategies was performed through mesh refinement and step-size sensitivity studies. In Model S, the mesh was refined near screw holes and contact interfaces until per-screw axial forces changed by less than 5% in four-point bending. In eccentric compression, localized contact chattering at the near-gap screw required stabilization parameter tuning and finer meshing to reduce variability in the per-screw force predictions to approximately 8%. In torsion, at higher twist angles, convergence difficulties were encountered and per-screw predictions shifted by roughly 12% with refinement in the most sensitive configurations, a level we judged material for high-risk decision-making.

We tracked residual norms and contact status histories in Model S to ensure that apparent convergence did not mask unconverged contact states. When contact chattering occurred, we systematically explored stabilization parameters and ramped loading to manage convergence without unduly damping legitimate interface behavior. The final parameter selections were justified based on stability and minimal impact on per-screw force distribution in the bending and compression cases; however, torsion presented shape-change conditions that amplified sensitivity to these numerical choices.

In Model R, solutions were robust. Because the model avoids contact and large deformations, linear static steps produced stable results in bending and torsion. Even in the eccentric compression case, where an offset introduces coupled bending effects, linear solution steps remained stable and reproducible. Per-screw axial force fractions changed by less than 3% under mesh refinement in bending and by less than 2% in eccentric compression and torsion, reflecting the model's numerical stability. These small changes do not, by themselves, ensure that the model faithfully represents the apparatus; they only indicate that the computed responses are algorithmically robust given the model's idealizations.

Verification activities were documented with version-controlled scripts and notebooks that automate benchmark setups and result checks. This enabled repeatability of code verification and numerical error quantification whenever the models were updated.

Numerical solver error — gradation.

a.
No quantification of discretization or solver-induced error.

b.
Single mesh or step size with anecdotal convergence statement.

c.
Mesh and step-size sensitivity studies performed for representative loads with error reported on outputs of interest.

d.
Targeted refinement and solver-parameter sensitivity quantified per mechanism until changes in outputs of interest are below pre-set thresholds tied to decision risk.

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

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

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

8. 8. Input equivalency and model inputs

To translate laboratory features into the computational models, we mapped the support spans, load spacings, offsets, and clamp constraints into boundary conditions in a one-to-one fashion wherever possible. In four-point bending, the 80 mm support span and 40 mm load spacing were represented explicitly in Model S via contact between the plate and rigid anvil surfaces and, in Model R, via beam nodes constrained to the same geometric positions. For the eccentric compression mechanism, we imposed a 5 mm shift of the load path relative to the surrogate bone axis and ensured that the distal support conditions allowed the same rotational freedom as the apparatus. In torsion, Model S used rotational boundary conditions coupled with clamping contact surfaces, whereas Model R imposed end rotations about the beam axis.

The plate, screws, and surrogate bone were assigned material properties based on measurements or supplier data for the specific materials used. The stainless steel plate and screws were assigned Young's moduli of 200–210 GPa and Poisson's ratio of 0.30. The polymer surrogate bone had a Young's modulus of 2.0–2.5 GPa and Poisson's ratio of 0.35, based on tensile and compressive coupon tests performed per ASTM guidance. Friction coefficients were measured on representative material pairs: screw head–plate friction was taken as $\mu = 0.15 \pm 0.03$ and screw–bone (polymer) friction as $\mu = 0.25 \pm 0.05$, derived from inclined-plane and pin-on-disk tests. For Model S, these coefficients were assigned to the corresponding contacts; for Model R, friction effects were not represented explicitly, consistent with the beam–spring idealization.

Screw pretension was represented in the models by converting installation torque to axial preload. The calibration used instrumented coupons in which screws were tightened through plate holes into polymer surrogate material while axial force was measured. The mapping between torque and preload was then used to specify the initial bolt load in Model S for each screw. In Model R, the effect of pretension was captured indirectly in the initial stiffness of the axial spring connecting the screw to the bone, ensuring that the clamped joint stiffness was consistent with the preloaded condition observed in the coupons.

For Model R, the axial spring parameters were identified from prior testing of isolated screws under axial displacement in the same surrogate polymer. The spring law was piecewise linear with an initial high-stiffness regime reflecting the preloaded state and a subsequent stiffness appropriate to small additional relative motions during bench loading. While these calibrations were performed under axial loading, they were not re-identified for combined compression–bending or torsional regimes, which limits their fidelity when the load path departs from pure axial engagement.

Boundary conditions in both models approximated fixture compliances. In four-point bending, the stiffness of the platens and anvils was high relative to the construct and was not found to influence per-screw force fractions to a degree detectable in the sensitivity studies. In eccentric compression and torsion, some fixture flexibility exists in the apparatus; in Model S, we approximated these effects by allowing limited compliance via contact definitions, whereas in Model R these compliances were not directly representable and were, instead, neglected. The implications of these choices are reflected in the credibility assignments presented later.

This section summarizes the data provenance for material and contact parameters and the manner in which laboratory geometry and control modes were represented in each model. These pieces together establish the foundation on which the credibility judgments for parameter selection and apparatus mapping were made.

Model inputs — gradation.

a.
Model parameters are assumed without citing their origins.

b.
Parameters are taken from literature or analogous systems with limited justification for current application.

c.
Parameters are measured on representative materials or supplied by manufacturers with tolerances and are mapped to model constructs.

d.
Parameters are measured and validated under mechanism-relevant conditions; uncertainties are quantified and shown to have negligible impact on outputs of interest for the context of use.

9. 9. Use error considerations

Analysts working with both models followed a standard modeling checklist and shared example setups for each mechanism to promote consistency. A peer reviewer examined initial model builds prior to production runs, and model transfer between analysts included discussion of common pitfalls. Templates were made available to streamline the creation of new instances of each mechanism for alternative plate geometries.

Numerical solver error — gradation.

a.
No records of solver settings or convergence behavior are kept.

b.
Default solver settings with qualitative statements about convergence.

c.
Documented solver settings and residual histories; sensitivity to stabilization or step-size parameters reported for representative cases.

d.
Standardized solver configurations with quantitative convergence criteria and sensitivity results per mechanism tied to acceptance thresholds.

10. 10. Results—Credibility assessment scores

This section presents model-by-mechanism assignments on the 0–7 Credibility Attainment Scale for the six assessed dimensions. The assignments draw on the evidence described earlier. For brevity, the summaries highlight the chief strengths and the limiting shortfalls. The tables provide a concise, sortable listing of levels for each combination of model, mechanism, and dimension, along with a one-line justification that restates specific, quantitative elements also argued in the body of the paper.

In broad terms, the three-dimensional contact model (Model S) showed strong mapping of the laboratory configurations into boundary and contact conditions in four-point bending, used measured material and friction properties and applied screw preload based on a lab-derived torque-to-tension mapping, and passed standard implementation checks. Its solution variability was contained in bending but increased under eccentric compression and was most pronounced in torsion. The reduced-order beam–spring surrogate (Model R) was numerically stable and reproduced elementary beam behavior, but it did not represent the rig features as faithfully in eccentric compression and torsion, and its axial spring parametrization was not re-identified for those combined-loading regimes.

Because the predefined acceptance criterion requires CAS of 6 or higher for all six dimensions across all mechanisms for at least one model, any dimension falling below this level in any mechanism prevents acceptance. The tables below should be read with that criterion in mind.

Test conditions — gradation.

a.
Rigs, spans, and control modes are not quantitatively documented.

b.
Partial documentation of apparatus geometry and control with gaps for key parameters.

c.
Complete documentation of spans, offsets, rates, and alignment used to guide modeling.

d.
Full documentation including fixture stiffness and clamp compliance with uncertainties reported and used in modeling or sensitivity analysis.

10.1. 10.1 Credibility-assessment summary—Model S

Model S is the higher-fidelity representation with explicit solid geometry and contact interfaces. Its strengths lie in reproducing laboratory spans, offsets, and clamp interactions and in employing measured physical properties and a lab-derived mapping from torque to preload for screw pretension. Numerical solver behavior was acceptable in bending, more variable in eccentric compression due to contact chattering near the gap, and most challenging in torsion, where higher twist angles amplified sensitivity to mesh and solver parameters. The apparatus was well documented and faithfully mirrored in bending; some flexibility of fixtures under eccentric compression and torsion was not captured in detail but had modest influence on bending and more noticeable impact on the other mechanisms. Sample counts were six for bending and five for the other two modes, all from a single lot with documented machining tolerances.

Equivalency of input parameters — gradation.

a.
Qualitative description of the test setup with no direct model translation.

b.
Partial translation of key spans/offsets with unquantified fixture effects.

c.
Direct translation of spans, offsets, supports, and torque-to-tension mapping with disclosed approximations to fixture flexibility.

d.
Full translation including fixture compliance or bounding sensitivity studies showing negligible impact on outputs of interest.

10.2. 10.2 Credibility-assessment summary—Model R

Model R is a reduced-order surrogate that trades geometric and contact detail for computational speed. It used beam elements for the plate and screws and nonlinear axial springs for screw–bone engagement. In four-point bending, this representation captured the principal load path with end constraints approximating platen contacts, and its numerical behavior was stable with small changes in outputs under refinement. In the eccentric compression mechanism, the 5 mm offset was introduced as a nodal shift of the load path, but fixture flexibility was not representable and spring parameters were not re-identified for the combined compression–bending regime. In torsion, the model imposed global twist but, lacking off-axis coupling in the screw–bone engagement and warping representation, it could not reflect secondary load paths active in the screw group under twist. These limitations, while not affecting numerical stability, constrained the model's ability to mirror the apparatus faithfully and to use mechanism-specific parameters.

The following table summarizes the CAS levels for the six dimensions across the three mechanisms for Model R, accompanied by one-line justifications that restate quantitative elements detailed earlier.

Model inputs — gradation.

a.
Surrogate parameters assumed by analogy without calibration.

b.
Parameters calibrated for a different mechanism without justification for transfer.

c.
Parameters calibrated on representative coupons with mechanism limitations disclosed.

d.
Parameters re-identified per mechanism with uncertainty bounds and demonstrated robustness of outputs to parameter variation.

11. 11. Discussion

The high-risk context of use specified for these models required strong, mechanism-specific evidence across six credibility dimensions for at least one model to be considered acceptable. Our results indicate that while both models demonstrate notable strengths, neither achieves the required levels across all dimensions and mechanisms. The following discussion interprets the pattern of evidence, focusing on how the design of each model interacts with the laboratory mechanisms and where that interaction results in limitations material to the stated decision.

A central requirement is that the benches be mirrored in the simulations such that the applied forces and kinematics in the model correspond directly to those experienced by the physical construct. In this regard, Model S captured the spans, load spacing, and boundary supports of four-point bending with near-exactitude and represented screw pretension through a mapping from torque to axial preload derived from instrumented fastener coupons. In eccentric compression, that close correspondence was weakened by the absence of explicit platen flexibility, and in torsion by the lack of explicit clamp compliance modeling, both of which subtly alter the distribution of boundary reaction forces even if the global twist or offset is matched.

The parameters that give the models their physical character—elastic moduli, Poisson ratios, friction coefficients, and screw preload—were drawn from direct measurements on representative materials or supplied by the manufacturer, with tolerances stated and, for the friction coefficients, means and standard deviations estimated from repeated tests. In Model S, friction acted at the screw head–plate and screw–bone interfaces and, combined with bolt preload, controlled the onset of micro-slip. In Model R, the absence of explicit friction or off-axis coupling meant that these mechanisms influenced load sharing only through the axial spring, which was calibrated under purely axial conditions and not re-identified for combined loading.

Checks that the implemented numerical formulations reproduce known solutions formed a strong base of evidence for both models. For Model S, patch tests and classic beam and torsion benchmarks produced deflections and twists within a few percent of closed-form solutions across mesh refinements. For Model R, beam and spring assemblies recovered textbook responses essentially exactly in the linear regime that the model inhabits. These checks do not guarantee fidelity to the benches but do show that the computation itself is not the source of error in the behaviors represented by those formulations.

The magnitude of the residual variability in per-screw force predictions attributable to discretization, step size, and solver parameters was small in bending for Model S (<5% change with refinement), increased in eccentric compression (~8% change with refinement due to contact chattering near the gap), and was largest in torsion (~12% change with refinement at higher twists where convergence was most challenging). Model R exhibited minimal change in outputs with mesh or step variation (<3% in bending and <2% in eccentric compression and torsion), reflecting numerical stability but not necessarily accurate representation of the laboratory load paths in the more complex mechanisms.

The apparatus and laboratory protocols were quantitatively documented and then used to set up the simulations. Spans of 80 mm and 40 mm for supports and loads in four-point bending, a 5 mm medial offset in eccentric compression, and a twist rate of 0.5 deg/s in torsion were all mirrored. Alignment procedures, backlash mitigation in torsion, and load-control modes were matched to the extent permitted by each model's design. This documentation provides confidence that departures in predictions are not due to unknown or drifting laboratory conditions but rather to model idealizations and parameter identification choices.

The physical components used to generate the laboratory evidence came from a single production lot and were machined to common geometry, with six specimens in bending and five each in eccentric compression and torsion. Dimensional verification showed small deviations from nominal, and repeatability was fair across specimens in the two five-sample test modes. While these sample sizes and controls are reasonable for screening purposes, they fall short of the most rigorous characterization expected for high-risk decisions, where more extensive statistical characterization and propagation of geometric variability into model predictions would be desirable.

Turning to the model-by-mechanism patterns, Model S performed strongly in four-point bending. The translation of the apparatus into the model was direct, pretension mapping was grounded in laboratory calibration, material and contact parameters were measured and assigned with tolerances, and solver variability was kept within a few percent with a manageable amount of stabilization. The sample base of six constructs from a single lot with tight machining tolerances allowed the assignments at or near the high end of the scale for most dimensions. The chief limitation in bending was the need for mild stabilization in contact, which, though not materially affecting the output at the loads examined, prevented a highest-level assignment for solver-related variability.

In eccentric compression, Model S retained solid footing in terms of material and contact parameterization and implementation checks, but departures from bench fidelity emerged. The 5 mm offset was imposed, yet the omission of explicit platen compliance and sensitivity to contact chattering at the near-gap screw resulted in per-screw axial fractions that changed by approximately 8% under refinement—higher than desired for a high-risk decision. The test base of five specimens from a single lot was adequate but not extensive. Together, these factors produced mid-scale assignments for some dimensions, short of the high-risk target.

Torsion posed the greatest difficulty for Model S. While global twist and fixture geometry were matched and implementation checks on torsion benchmarks were strong, convergence issues at higher twist angles led to solver sensitivity in the per-screw axial force predictions of roughly 12%. The clamp flexibility of the physical apparatus, which shifts how twist is transmitted into shear and axial force paths at the screw group level, was not represented explicitly. These two aspects limited the assignments for apparatus mapping and solver-related variability, and the five-specimen test base, while fair, was not sufficiently rich to offset those limitations.

Model R presented a different pattern. In four-point bending, the beam–spring surrogate captured the principal geometry and loading features, and the numerical behavior was robust and reproducible. The end constraints approximated platen contacts, and while local contact patch details were idealized, the mapping of spans and load magnitudes was adequate. The axial spring stiffness was drawn from prior coupon testing of isolated screws under axial loading in the same surrogate material. These elements supported mid- to upper-level assignments for several dimensions, with the caveat that contact patch idealization prevents the highest assignment for apparatus fidelity.

In eccentric compression, however, the limitations of the reduced-order surrogate became more pronounced. The 5 mm offset was implemented as a nodal shift in the load path, but the apparatus flexibility could not be represented, and most importantly, the axial spring parameters were not re-identified for the combined compression–bending regime that elevates demand on near-gap screws through coupled bending and axial engagement. As a result, the mapping of laboratory features into the model and the selection of parameters pertinent to this mechanism were insufficient for a high-risk decision.

Under torsion, Model R imposed global twist correctly but did not include mechanisms to represent warping and secondary load paths that arise due to off-axis coupling and frictional constraints in the screw group. Without a torsion-specific parameter identification for the screw–bone engagement and absent a representation of screw head tilting or thread flank shear, the surrogate's mapping of the apparatus into model degrees of freedom and its parameterization lacked key elements for this mechanism. Although the numerical solutions were stable and reproducible, the representation gap prevented strong assignments for apparatus fidelity and parameter selection.

Taken as a whole, the evidence indicates that while Model S approaches the high-risk threshold in the most plate-dominant bending case, it falls short in mechanisms where contact nonlinearity and fixture flexibility play larger roles. Model R, on the other hand, is attractive for its speed and numerical stability but lacks sufficient mechanism-specific mapping and parameter identification in eccentric compression and torsion to support high-stakes design screening. Given the predefined criterion requiring level 6 or higher across all six dimensions and all three mechanisms for at least one model, these limitations are dispositive.

Test samples — gradation.

a.
Unknown lot origin and uncontrolled geometry.

b.
Mixed lots with partial geometry documentation.

c.
Single lot with dimensional verification and planned sample sizes adequate for screening.

d.
Statistically characterized lots with uncertainty propagation to outputs of interest supporting high-risk decisions.

12. 12. Conclusion

This study applied a risk-informed assessment to two finite element models intended to predict screw load sharing in a 3.5 mm locking plate construct under three standard bench mechanisms. The assessment was organized under a six-dimension credibility plan aligned to the needs of high-stakes design screening. A 0–7 Credibility Attainment Scale was used, and a strict acceptance criterion required that at least one model reach a level of 6 or higher on all dimensions for all mechanisms.

The three-dimensional contact model (Model S) performed well in four-point bending, with strong mapping of the apparatus into the model, measured physical parameters, sound implementation checks, and acceptable solver-related variability. However, in eccentric compression and more significantly in torsion, solver-induced variability increased to levels that we judge material for the stated high-risk context, and the omission of explicit clamp and platen compliances limited the fidelity with which boundary reactions were reproduced.

The reduced-order beam–spring surrogate (Model R) maintained strong numerical stability and reproduced basic beam behavior but did not include sufficient mechanism-specific mapping or parameter identification in eccentric compression and torsion. The axial spring parameters were not re-identified for these combined-loading regimes, and the surrogate lacks the capacity to represent warping and secondary load paths in the screw group under twist.

Given these shortfalls, the predefined acceptance criterion was not met. The final decision is Not accepted for the stated context of use. The models, as presently configured and supported by the compiled evidence, should not be used as the sole basis for down-selecting new plate configurations ahead of bench testing in this high-risk setting. Future work could address the identified gaps by representing fixture compliances more completely, tuning spring and contact parameters under mechanism-relevant conditions, and reducing solver sensitivity in torsion through targeted meshing and contact formulations.

Numerical code verification — gradation.

a.
No quantitative benchmark evidence.

b.
Limited benchmarks with qualitative agreement only.

c.
Comprehensive benchmarks with quantitative agreement within tolerance for elements and contacts used.

d.
Benchmarks plus peer-reviewed verification artifacts traceable to model versions and run conditions.

References

[1] H. Zetterberg, N. Yilmaz, K. Okonkwo, B. Halvorsen, Damping effects on flow split in computational models, J. Biomech. Eng. 67 (2005) 369–459.

[2] M. Rasmussen, M. Silva, F. Halvorsen, M. Eriksen, Wall roughness effects on peak force in computational models, Int. J. Numer. Methods 87 (2016) 98–771.

[3] C. Yilmaz, K. Ueda, Element type effects on contact pressure in computational models, Comput. Methods Biomech. 113 (2024) 207–601.

[4] A. Halvorsen, W. Pettersson, M. Bergman, G. Tanaka, T. Tanaka, Load path effects on chest compression in computational models, J. Verif. Valid. Uncertain. 43 (2022) 134–759.

[5] W. Okonkwo, P. Nakamura, Wall roughness effects on flow split in computational models, J. Biomech. Eng. 44 (2018) 100–619.

[6] T. Fontaine, K. Pettersson, Flow rate effects on rib deflection in computational models, Comput. Fluids 77 (2022) 234–407.

[7] R. Eriksen, D. Yilmaz, D. Jankowski, Material model effects on head displacement in computational models, J. Verif. Valid. Uncertain. 125 (2005) 377–504.

[8] R. Ishikawa, M. Rasmussen, W. Moreau, N. Vasquez, Load path effects on flow split in computational models, Ann. Biomed. Eng. 94 (2017) 120–887.

[9] L. Pettersson, J. Okonkwo, M. Ishikawa, R. Jankowski, Friction effects on rib deflection in computational models, Ann. Biomed. Eng. 145 (2022) 286–708.

[10] D. Grigoryan, W. Tanaka, Material model effects on contact pressure in computational models, Ann. Biomed. Eng. 100 (2014) 93–700.

[11] N. Zetterberg, K. Halvorsen, L. Yilmaz, Cement mantle effects on head displacement in computational models, Comput. Methods Biomech. 40 (2009) 275–629.

[12] S. Jankowski, J. Yilmaz, D. Moreau, S. Tanaka, Mesh density effects on hemolysis index in computational models, Med. Eng. Phys. 120 (2008) 227–539.

[13] J. Okonkwo, D. Nakamura, N. Lindqvist, Element type effects on head displacement in computational models, Med. Eng. Phys. 24 (2009) 33–442.

[14] T. Pettersson, D. Yilmaz, K. Wojcik, D. Duarte, P. Fontaine, Bone density effects on contact pressure in computational models, Ann. Biomed. Eng. 86 (2014) 20–849.

[15] J. Nakamura, F. Okonkwo, B. Okonkwo, K. Yilmaz, G. Kaur, Damping effects on rib deflection in computational models, J. Mech. Behav. Biomed. Mater. 127 (2008) 334–822.

[16] H. Halvorsen, N. Castellano, W. Ueda, N. Moreau, J. Jankowski, Time step effects on shear stress in computational models, J. Verif. Valid. Uncertain. 76 (2013) 353–608.

[17] C. Castellano, N. Zetterberg, C. Lindqvist, T. Yilmaz, Inlet profile effects on outlet velocity in computational models, Comput. Fluids 129 (2008) 359–559.

[18] L. Wojcik, J. Kaur, D. Okonkwo, H. Grigoryan, Inlet profile effects on rib deflection in computational models, Int. J. Numer. Methods 99 (2014) 80–630.

[19] H. Moreau, E. Tanaka, Cement mantle effects on cortical strain in computational models, Int. J. Numer. Methods 121 (2019) 4–699.

[20] K. Bergman, R. Halvorsen, L. Abbott, W. Halvorsen, Contact stiffness effects on hemolysis index in computational models, Ann. Biomed. Eng. 95 (2023) 16–541.

[21] A. Castellano, M. Nakamura, T. Pettersson, K. Eriksen, J. Yilmaz, Mesh density effects on contact pressure in computational models, Ann. Biomed. Eng. 138 (2021) 289–788.

[22] P. Fontaine, T. Eriksen, E. Lindqvist, Flow rate effects on peak force in computational models, Comput. Methods Biomech. 108 (2006) 89–698.

[23] M. Duarte, A. Ueda, Damping effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 32 (2015) 386–470.

[24] N. Abbott, K. Silva, J. Moreau, Flow rate effects on shear stress in computational models, Comput. Fluids 56 (2013) 25–576.

[25] W. Lindqvist, G. Wojcik, M. Grigoryan, Solver tolerance effects on stress range in computational models, J. Verif. Valid. Uncertain. 94 (2012) 44–637.

[26] M. Fontaine, C. Quintero, G. Jankowski, B. Ishikawa, Inlet profile effects on peak force in computational models, Comput. Methods Biomech. 72 (2013) 279–734.

[27] H. Jankowski, J. Ishikawa, E. Halvorsen, L. Nakamura, P. Grigoryan, Boundary condition effects on shear stress in computational models, Ann. Biomed. Eng. 136 (2023) 355–845.

[28] L. Fontaine, L. Silva, G. Jankowski, Boundary condition effects on flow split in computational models, J. Biomech. Eng. 67 (2017) 90–795.

[29] F. Pettersson, E. Zetterberg, G. Duarte, H. Ueda, Cement mantle effects on hemolysis index in computational models, Comput. Methods Biomech. 121 (2006) 46–713.

[30] T. Silva, E. Silva, Time step effects on head displacement in computational models, Med. Eng. Phys. 58 (2010) 385–466.

[31] H. Fontaine, E. Ishikawa, S. Nakamura, M. Eriksen, Solver tolerance effects on rib deflection in computational models, J. Biomech. Eng. 57 (2019) 265–491.

[32] T. Halvorsen, T. Lindqvist, Boundary condition effects on fatigue life in computational models, J. Verif. Valid. Uncertain. 65 (2015) 330–866.

[33] E. Pettersson, W. Castellano, G. Ueda, R. Yilmaz, Yield stress effects on hemolysis index in computational models, Ann. Biomed. Eng. 89 (2011) 387–562.

[34] H. Moreau, L. Tanaka, C. Vasquez, C. Castellano, M. Eriksen, Inlet profile effects on rib deflection in computational models, Ann. Biomed. Eng. 98 (2008) 391–838.

Table 1: Credibility Attainment Scale summary for Model S across three mechanisms. Levels are on a 0–7 scale.

Credibility factor	Level	Basis
Equivalency of input parameters - Model S - Four-point bending	6	80 mm support span and 40 mm load spacing were mirrored; torque-to-preload mapping from coupon calibration applied; support contacts represented.
Model inputs - Model S - Four-point bending	6	Measured E and ν for steel (200–210 GPa, 0.30) and polymer (2.0–2.5 GPa, 0.35); friction μ=0.15±0.03 (head–plate), μ=0.25±0.05 (screw–bone); preload applied as bolt load.
Numerical code verification - Model S - Four-point bending	6	Patch tests passed; beam-bending tip deflection within 2% of closed-form; contact benchmarks behaved as expected.
Numerical solver error - Model S - Four-point bending	5	Mesh refinement changed per-screw axial forces by <5%; mild contact stabilization required.
Test conditions - Model S - Four-point bending	6	Lab rig spans and load spacing documented; quasi-static rate and alignment matched; anvil geometry represented.
Test samples - Model S - Four-point bending	5	n=6 from a single lot; CNC machining with ±0.15 mm diameter tolerance and ±0.10 mm hole placement.
Equivalency of input parameters - Model S - Eccentric axial compression	5	5 mm medial offset imposed; distal support freedom matched; minor fixture compliance not modeled explicitly.
Model inputs - Model S - Eccentric axial compression	5	Same measured properties and friction; torque-to-preload mapping applied; sensitivity to μ and preload noted.
Numerical code verification - Model S - Eccentric axial compression	6	Same verified elements and contact formulations applied; linear subcases matched benchmarks.
Numerical solver error - Model S - Eccentric axial compression	4	Contact chattering at the near-gap screw; refinement reduced per-screw variability to ~8%.
Test conditions - Model S - Eccentric axial compression	5	Offset and axial control documented and reproduced; platen stiffness approximated.
Test samples - Model S - Eccentric axial compression	4	n=5; same lot; modest surrogate variability observed.
Equivalency of input parameters - Model S - Pure torsion	4	Global twist boundary replicated; clamp compliance not explicitly modeled.
Model inputs - Model S - Pure torsion	5	Material and friction values as measured; bolt preload retained during twist.
Numerical code verification - Model S - Pure torsion	6	Solid-element shaft torsion matched Saint-Venant solution within ~2% in linear regime.
Numerical solver er-	3	Convergence difficulties at

Table 2: Credibility Attainment Scale summary for Model R across three mechanisms. Levels are on a 0–7 scale.

Credibility factor	Level	Basis
Equivalency of input parameters - Model R - Four-point bending	4	Beam end constraints approximated platen contacts; springs mapped from axial engagement; local contact idealized.
Model inputs - Model R - Four-point bending	4	Axial spring stiffness from prior coupon calibration; no mechanism-specific tuning for bending beyond axial engagement.
Numerical code verification - Model R - Four-point bending	6	Beam element benchmarks and spring assembly checks matched analytic targets.
Numerical solver error - Model R - Four-point bending	6	Linear static solutions robust; per-screw axial fractions changed <3% with refinement.
Test conditions - Model R - Four-point bending	5	Global spans and load magnitudes matched; contact patches idealized to nodal supports.
Test samples - Model R - Four-point bending	5	n=6 from a single lot; geometry per CNC with documented tolerances.
Equivalency of input parameters - Model R - Eccentric axial compression	3	5 mm eccentricity introduced as nodal offset; fixture compliance not representable in the surrogate.
Model inputs - Model R - Eccentric axial compression	3	Axial spring parameters not re-identified for combined compression–bending.
Numerical code verification - Model R - Eccentric axial compression	6	Beam and spring components verified; linear subcases reproduced textbook responses.
Numerical solver error - Model R - Eccentric axial compression	6	Stable linear solutions; mesh/step sensitivity <2% on per-screw fractions.
Test conditions - Model R - Eccentric axial compression	5	Offset magnitude matched; platen–construct interactions idealized in boundary conditions.
Test samples - Model R - Eccentric axial compression	4	n=5; same lot; modest surrogate variability documented.
Equivalency of input parameters - Model R - Pure torsion	2	Global twist imposed but warping and secondary screw-group load paths not representable with axial springs alone.
Model inputs - Model R - Pure torsion	3	No torsion-specific re-identification of axial spring parameters.
Numerical code verification - Model R - Pure torsion	6	Beam torsion checks acceptable; twist per unit torque matched analytic solutions.
Numerical solver error - Model R - Pure torsion	6	Stable, linear solutions; refinement changed outputs by <2%.
Test conditions - Model R - Pure torsion	5	Twist boundary matched globally; clamp details idealized.
Test samples - Model R - Pure torsion	4	n=5; repeatability fair; single manufacturing lot.