Credibility Assessment of Explicit Finite Element Models for Total Knee Replacement Wear and Contact Mechanics Under NASASTD7009A

Alyssa R. Kemp, PhD[a], Noah J. Feldman, MS[b], Meera V. Srinivasan, PhD[c], James P. O'Connell, PhD[1]

[a]*Computational Orthopaedics Laboratory, Center for Orthopaedic Innovation, Denver, CO, USA,*
[b]*Biomechanical Simulation Group, OrthoSim LLC, Boston, MA, USA,*
[c]*Department of Mechanical Engineering, Midwestern Technical University, Chicago, IL, USA,*

Abstract

Finite element analysis (FEA) is widely used to guide the design of total knee replacements (TKRs) and to anticipate wear, contact mechanics, and stability phenomena that are difficult to probe experimentally. We report a structured, NASASTD7009A–style credibility assessment of two explicit dynamics models developed for a cemented posteriorstabilized (PS) TKR: EXDKnee v1.2, which computes transient articular and post–cam contact, and EXDWearGrow v2.0, which accumulates freesurface degradation through wear and creep with adaptive remeshing. Four clinically relevant mechanisms were assessed: (1) articular surface wear during ISOstyle level walking, (2) peak contact stress in deep flexion under combined axial–rotational loading, (3) edgeloading under varus malalignment with internal tibial rotation, and (4) post–cam impact and sliding during stair descent and sittostand. The assessment spans eight evidence dimensions: provenance of supporting data, appropriateness of the mathematical and physical representation, implementation correctness checks, numerical residuals from time integration and contact enforcement, comparison to bench and simulator measurements, sensitivitybased stability of conclusions, quantification of spread due to input variability, and representativeness of the laboratory load histories. We adopt a private 0–12 scoring scale with anchors at 3 (initially acceptable), 6 (adequate for comparative assessment), 9 (strong), and 12 (compelling). Inputs and comparators include (i) three primary laboratory campaigns comprising 36 stationhours of instrumented knee simulator testing under ISO 14243 waveforms with bovine serum at 25 g/L protein, (ii) deepflexion pressure film and instrumented post–cam rigs, and (iii) previously published, fully specified wearrate datasets for UHMWPE. Across all scenarios, the median deviation from physical tests was 9.1% for contact metrics and 11.7% for wear or cam loads; solver artifacts were contained by energy balance checks (cumulative nonphysical energy below 0.5% in gait and 0.9% in deep flexion) and by contact penetration limits (<0.02 mm). Uncertainty analyses propagated variability in material properties, friction, and loading, yielding 95% intervals of ±15% or tighter for gait wear and post–cam impact, and broader envelopes up to ±22% under rim loading. One credibility matrix is presented per model (32 rows each; 64 rows total), with each row scored for one evidence dimension under one mechanism. We conclude that both models are strong to compelling for levelwalking and post–cam questions, and adequate to strong for highflexion and edgeloading questions, provided users remain within stated testcondition envelopes and acknowledge wider uncertainty for rimcontact predictions.

Keywords: total knee replacement, explicit dynamics, credibility

1. Introduction

Longterm performance of posteriorstabilized total knee replacements depends on a coupled set of mechanisms spanning smooth articular sliding, cornercase rim loading, and post–cam engagement. Ultrahighmolecularweight polyethylene (UHMWPE) inserts experience multidirectional sliding, crossshear, and timedependent deformation. Transient contact with the femoral component's condyles and with the tibial insert's post can generate high local stresses and impact loads that influence fatigue, wear, and risk of instability. Because reproducing all clinically relevant maneuvers under controlled bench conditions is impractical, explicit finite element analysis (FEA) has become a key planning and design tool.

This paper evaluates, under a NASASTD7009A–style framework, the credibility of two explicit FEA models for a cemented PS knee: EXDKnee v1.2, which focuses on highfidelity transient contact mechanics, and EXDWearGrow v2.0, which augments explicit dynamics with cyclebycycle surface evolution via wear and creep using adaptive meshing. We target four mechanisms: articular surface wear during ISOstyle level walking, peak contact stress in deep flexion under combined axial–rotational loading, edgeloading under varus malalignment with internal tibial rotation, and post–cam impact and sliding during stair descent and sittostand.

The credibility dimensions considered here follow the widely used, evidencebased structure of NASASTD7009A, adopting a private 0–12 scale to summarize the strength of evidence by model and by mechanism. In keeping with best practices, the assessment is transparent about the contents and provenance of inputs and comparators, the mathematical and algorithmic choices, numerical controls, agreement with experiments, sensitivity and uncertainty, and the representativeness of the tests used to bound intended use.

The experimental datasets used to anchor inputs and check results were drawn primarily from dedicated instrumentedrig campaigns with traceable calibration and complete raw waveforms, accounting for 82% of all comparisons; the remainder came from peerreviewed simulator studies with full procedural reporting and sufficient metadata to reconstruct loads and kinematics.

Output comparison — gradation.

a.
Level 1: Only qualitative agreement (e.g., shape or trend) reported; no common metrics or aligned conditions.

b.
Level 2: Quantitative metrics reported under nonidentical conditions, or partial alignment of boundary conditions; error bars absent.

c.
Level 3: Direct metrictometric comparison under matched conditions with point estimates; at least one independent dataset.

d.
Level 4: Multiple independent datasets under matched conditions; prespecified acceptance criteria; residuals analyzed with uncertainty bounds.

2. Device Configuration and Materials

The assessed construct is a cemented posteriorstabilized total knee replacement. The femoral component is CoCrMo (ASTM F1537) with condylar radii of 24.5 mm (medial) and 21.8 mm (lateral) at 0–30° flexion with progressive flattening to 18.0 and 16.5 mm beyond 60°. The trochlear groove is not engaged in the evaluated tasks. The tibial baseplate is Ti6Al4V (ASTM F136) with a UHMWPE insert (GUR 1020based, highly crosslinked, remelted) mechanically locked by dovetail. The insert's tibial post is centered anteroposteriorly, with a rectangular planform of 10.8 mm transverse width and 13.2 mm anteroposterior width, chamfered at 1.2 mm radii; the post face is nearly vertical (3° anterior tilt), and the femoral cam has a 14.5 mm radius with a 1.0 mm corner blend.

Linear elastic properties are assigned for the metals (CoCrMo: E = 210 GPa, ν = 0.30; Ti6Al4V: E = 110 GPa, ν = 0.33). The UHMWPE is represented with a pressuresensitive, ratedependent constitutive model: a smallstrain instantaneous modulus Einst = 1.05 GPa at 1 s1 decreasing to 0.82 GPa at 103 s1; Poisson's ratio ν = 0.46; and a Drucker–Prager cap for volumetric creep with a longterm creep compliance J(t) approximated by a Findley law, ϵc(t) = k σ t n with k = 5.1×105 MPa1 and n = 0.18 under 0.5–12 MPa mean stress. Plastic flow is capped by an effective yield stress of 22 ± 3 MPa (mean ± SD), based on ringondisc data. UHMWPE density is 0.935 g/cm3 for wear volume conversion.

Contact is enforced via a penalty method with an augmented Lagrange update; the surface friction coefficient for CoCrMoUHMWPE in bovine serum at 37 °C is modeled as Coulomb with μ = 0.05 (nominal), varied between 0.03 and 0.08 in sensitivity and uncertainty studies. The target maximum normal penetration at equilibrium is 0.02 mm under ISO 14243 peak axial loads; the nominal penalty stiffness per contact pair is tuned to limit the peak interpenetration to 0.012–0.018 mm under those conditions. Lubrication is not resolved; the Coulomb law and crosssheardependent wear rule for EXDWearGrow v2.0 absorb the effective boundary friction.

Geometric and material values were independently checked against the manufacturer's drawings and certificates of compliance. Dimensional metrology of three asmanufactured inserts confirmed post widths to within 0.05 mm and articular sagittal radii to within 0.12 mm of drawings; this precision supports narrow confidence intervals during contact verification. The laboratory load histories reproduced task kinetics to within 3% RMS and matched joint angles within 1.5°, providing confidence that the simulated scenarios closely reflect intended clinical maneuvers for the four mechanisms.

Test conditions — gradation.

a.
Level 1: Conditions poorly matched to intended use; key kinematics or environments absent.

b.
Level 2: Partial overlap with intended use; at least one primary boundary condition matched (e.g., axial load) but others not (e.g., torque).

c.
Level 3: Boundary conditions fully matched to a recognized standard or welldocumented maneuver; environment approximated (e.g., serum surrogate).

d.
Level 4: Boundary conditions matched and verified by independent instrumentation with traceable calibration; environment controlled and reported with tolerances.

3. Computational Models and Preprocessing

Two models are assessed. EXDKnee v1.2 is an explicit dynamics contact model targeting transient contact pressure, area, shear traction, crossshear index, subsurface stress in UHMWPE, and instantaneous post–cam impact loads. It employs centraldifference time integration with selective mass scaling limited to a factor of 1.02 to satisfy the Courant condition over the finest UHMWPE elements. EXDWearGrow v2.0 extends the explicit dynamics kernel with a cyclebycycle update of the free surface via an Archardtype wear rule generalized for crossshear and counterface hardness, and a creep update driven by mean stress using a Findleytype law; adaptive remeshing maintains element quality as the surface evolves.

For both models, the articular surfaces are meshed with hexahedral elements where possible. The insert uses a structured mesh with 0.5 mm elements through the thickness in the first 2.0 mm beneath the surface, gradating to 1.0–1.5 mm in the core; inplane articular elements are 0.5 mm nominal, clustering to 0.35 mm in regions of high curvature (post corners, cam blend). The femoral condyles are meshed with 0.5–0.8 mm triangular prisms mapped along the contact normal to resolve pressure gradients. A typical EXDKnee v1.2 model contains

[redacted]

Figure 1: Overview of the assessment flow: models and preprocessing feed simulated mechanisms; each mechanism is compared to targeted laboratory tests and simulator data; evidence is synthesized along eight dimensions and summarized by permodel matrices on a 0–12 scale.

[redacted]

Figure 2: Schematic of the posteriorstabilized knee construct showing femoral cam and tibial post geometry; highlighted regions show condylar contact patches during level walking and the post–cam interface during stair descent.

approximately 1.2 million elements and 1.35 million nodes; EXDWearGrow v2.0 begins with ~1.5 million elements due to a denser surface discretization and evolves to 1.6–1.8 million elements with weardriven refinement. Surfaces are smoothed by a curvaturepreserving Laplacian operator constrained to <0.01 mm displacement, preventing artificial flattening.

Boundary conditions follow taskspecific prescriptions described later. All joints are modeled with 6DOF relative motion enforced by applying the ISO 14243 tibiofemoral kinematics (for gait) or by imposing specific flexion angles, axial loads, and torques (for deep flexion and edgeloading). The cam–post interface is engaged by driving the femur into the tibial post at prescribed flexion for stair descent and sittostand, with posterior tibial translation triggered by quadriceps and hamstrings surrogates in the load control. Mesh generation, contact pair definitions, and load scripts are validated by automated checks that confirm surface normals, nodeset completeness, and timehistory consistency with specified waveforms.

With production settings, residual numerical artifacts were small: peak element Courant limits were enforced with a 0.35 microsecond step, selective mass scaling was kept below 1.02 across all tasks, and cumulative nonphysical energy stayed under 0.5% during gait, 0.9% in deep flexion, and 0.7% for post–cam events.

Numerical solver error — gradation.

a. Level 1: No quantitative reporting of time step, stability, or energy balance; contact penetration uncontrolled.

b. Level 2: Reported time step or stability metric; limited checks (e.g., energy balance) for a subset of cases.

c. Level 3: Full reporting of stability controls (time step, mass scaling), energy and hourglass budgets, and penetration; acceptance thresholds defined and met for most cases.

d. Level 4: Same as Level 3 with formal error budgets per mechanism, parameter sweeps demonstrating invariance, and automatic monitors integrated into regression tests.

4. Mechanisms and Loading Profiles

Four mechanisms reflecting core TKR performance questions were simulated. For level walking, ISO 142431/3–style profiles were used: flexion from 0 to 58° with a midstance valley at 17°; axial load peaking at 2,600 N; anterior–posterior (AP) shear ±110 N; and internal–external (IE) torque ±5.0 N·m. The gait cycle was sampled at 1,001 points for EXDKnee v1.2 contact simulations and aggregated into 100 representative subcycles for EXDWearGrow v2.0 wear and creep updates. The lubricant surrogate was bovine serum at 25 g/L protein maintained at 37 ± 1 °C.

For peak contact stress in deep flexion, a quasistaticplustransient protocol was used: flexion of 120°, axial load stepped to 2,800 N over 25 ms, followed by a 15 ms ramp of an internal tibial rotation to 15° with a superimposed 18 N·m torque, maintaining small AP shear (<60 N). This sequence was designed to capture the peak engagement of the posterior condyles with the insert's posterior lip without cam–post contact.

Edgeloading scenarios introduced a 3° varus malalignment of the tibial insert relative to the baseplate and a 5° internal tibial

Table 1: Device geometry and material properties used in the simulations.

Credibility factor	Level	Basis
Femoral component	—	CoCrMo; condylar radii 24.5/21.8 mm (0–30°), flattening to 18.0/16.5 mm (>60°)
Tibial baseplate	—	Ti6Al4V; insert dovetail lock
UHMWPE insert	—	GUR 1020, highly crosslinked; Einst 1.050.82 GPa (1 s1103 s1), ν 0.46; yield 22±3 MPa
Post and cam	—	Post 10.8×13.2 mm with 1.2 mm chamfers; cam radius 14.5 mm with 1.0 mm blend
Friction	—	Coulomb µ = 0.05 nominal (0.03–0.08 in studies)
Contact control	—	Penalty with augmented Lagrange; peak penetration targeted <0.02 mm under ISO loads
Lubricant	—	Bovine serum 25 g/L, 37±1 °C; not explicitly modeled hydrodynamics

rotation; the axial load ramped to 2,400 N over 10 ms at 30° flexion and was sustained for 80 ms while a 3 mm medial shift of the load line induced rim contact. This setup is consistent with prior rimloading experiments using pressure film and dye to confirm the contact footprint.

Post–cam impact and sliding were simulated for stair descent and sittostand maneuvers. For stair descent, flexion transitioned from 40° to 72° while AP tibial translation shifted posteriorly by 8 mm; the cam engaged the post at ~46° with a relative approach speed of 0.28–0.32 m/s and kinetic energy of 0.42–0.55 J depending on step timing; repeated sliding followed over 40 ms with 3.5–5.0 N·m IE torque. For sittostand, a slower cam engagement at 0.12–0.16 m/s was used with 2,100 N axial load and 2.5–3.0 N·m IE torque, producing smaller but sustained post contact forces.

The profiles above are chosen to overlap standard practice and to isolate effects central to the mechanisms of interest. Local parameter sweeps showed output changes less than 5% for ±10% changes in friction and UHMWPE modulus for level walking and post–cam engagement, whereas rimloading peaks varied by up to 11% under the same input perturbations, indicating reduced stability of conclusions for the edgeloading case.
Results robustness — gradation.

a.
Level 1: No sensitivity analysis; single configuration reported.

b.
Level 2: Oneatatime parameter variations with qualitative commentary.

c.
Level 3: Structured sensitivity study across key parameters with quantitative effect sizes; assessment of mesh/timestep independence.

d.
Level 4: Global sensitivity (e.g., variancebased) or designed experiments with interaction effects; robustness quantified against prestated thresholds.

5. Data Pedigree and Test Conditions

Supporting data comprised three primary laboratory campaigns conducted inhouse on an AMTI sixstation knee simulator and two specialty rigs (deepflexion pressure mapping and instrumented post–cam). The simulator executed ISO 142431/3 gait cycles at 1 Hz using bovine serum (25 g/L protein) maintained at 37 ± 1 °C with viscosity checks at the start and end of each run (1.4 ± 0.1 mPa·s at 37 °C). Load cells and encoders were calibrated within six weeks of testing and verified by daily drift checks (<0.3% fullscale over 8 hours). The simulator generated six million cycles (Mc) per station across three stations (18 Mc total), with wear debris collected, filtered, and gravimetrically quantified (ASTM F1714) before conversion to volume via UHMWPE density 0.935 g/cm3.

Pressure film (Fuji Prescale Super Low and Medium) was employed in deep flexion to map peak contact; films were preconditioned and calibrated per manufacturer instructions and digitized at 600 dpi with a 10bit scanner; a conversion curve was generated by platen compression (R2 = 0.997). For rimloading, dyed contact revealing methods confirmed footprint position relative to the tibial insert rim. Post–cam loads were measured via a custom instrumented cam with a uniaxial highbandwidth force transducer (10 kHz sampling), in series with highspeed video (2,000 fps) to track relative cam–post approach velocity and engagement timing; the rig delivered impacts with 0.1–0.6 J kinetic energy.

External data included two peerreviewed simulator studies reporting PS knee wear rates under ISO conditions with sufficient metadata to reconstruct boundary conditions and surface finishes. Both studies provided volumetric wear per Mc (mean 33 ± 6 mm3/Mc and 41 ± 9 mm3/Mc) and crossshear indices computed from contact patch kinematics. These were used to supplement our inhouse datasets and to verify generality of the observed trends.

All raw analog data streams (loads, torques, displacements, temperatures) were archived at 2 kHz with synchronized timestamps; processed datasets (wear volume over cycles, pressure maps, and impact forces) were stored with versioned analysis scripts to enable independent reprocessing. The combined dataset spans 36 stationhours of simulator operation and 14 hours of specialty rig time. The integrity of acquisition and environmental control, along with complete metadata, was the basis for subsequently drawing quantitative comparisons to the explicit models without further tuning.
Data pedigree — gradation.

a.
Level 1: Data sources unspecified or secondary compilations; incomplete metadata.

b.
Level 2: Mix of primary and secondary sources with partial metadata; limited calibration detail.

c.
Level 3: Predominantly primary data with traceable calibration; raw data available for a subset of measurements.

d.
Level 4: Entirely primary, traceable datasets with complete raw data and metadata; independent replication or interlab checks.

6. Model Form Rationale and Assumptions

EXDKnee v1.2 emphasizes highfidelity transient contact and stress fields using explicit time integration. This choice is wellsuited to capturing post–cam impacts and rapid transitions during gait stance–swing changes. The UHMWPE constitutive description combines ratedependent elasticity with a capped plastic flow and creep component to represent shortterm load redistribution without fully resolving longterm viscoelasticity; this is sufficient for the short windows (milliseconds to seconds) in EXDKnee. Surface wear is not accumulated in EXDKnee; instead, drivers such as contact pressure, area, and crossshear are reported.

EXDWearGrow v2.0 adds a wear–creep growth loop around explicit dynamics: within each subcycle of the applied task, contact tractions are used to compute local wear depth via an Archardtype law with a crossshear multiplier MCS = 1 + α ξ, where ξ is the crossshear index derived from the directionality of sliding and α = 1.8 calibrated from ringondisc tests. The wear coefficient k is set at 2.8×107 mm3·N1·m1 (nominal); sensitivity analyses span (2.0–3.8)×107. Creep evolves by the Findley relation using the mean stress integrated over the dwell time of subcycles. After each cycle block (100 subcycles), the surface is updated and adaptively remeshed to maintain element quality (Jacobian > 0.7; aspect ratio < 5). This approach enables capture of evolving conformity and its feedback on contact tractions.

Assumptions shared by both models include Coulomb friction with a constant μ within a maneuver, neglect of thermal effects from frictional heating (temperature changes < 1 °C estimated over the simulated windows), and neglect of fluid film lubrication beyond the effective friction input. The tibial baseplate and femoral component are assumed rigid relative to the UHMWPE insert for stress computations (a standard simplification that slightly overpredicts local polymer stress by <5% in comparative elastic models that include metal compliance). Cement layer and bone are not modeled; their effect on articular and post–cam contact over the evaluated time windows is negligible.

As a consolidated finding, the model form is sufficient to capture highspeed contact transitions and the gradual topography changes from wear and creep that are central to the four mechanisms, with constitutive features and update laws aligned to the time scales probed in each task.
Model form — gradation.

a.
Level 1: Phenomena of interest omitted or misrepresented; key physics absent.

b.
Level 2: Core physics represented but with simplifications that limit applicability to target scenarios.

c.
Level 3: Physics and constitutive choices aligned to scenarios; acknowledged limitations outside intended range.

d.
Level 4: Comprehensive representation with couplings essential to mechanisms; alternative forms evaluated and justified.

7. Numerical Code Verification

Implementation correctness was assessed through a suite of independent checks. For linear elasticity, manufactured solutions with polynomial displacement fields were prescribed on hexahedral and prismatic meshes; the L2 displacement error decreased with secondorder convergence and reached 0.08% at 0.5 mm elements. A patch test verified stress equilibration across element boundaries to within 0.12%. Contact algorithms were verified against an axisymmetric Hertzian sphereonflat problem at 1,000 N: the predicted contact radius was within 3.1% of the analytic solution, and the pressure distribution had a 4.2% L2norm error. For frictional sliding, a ringondisc benchmark reproduced tangential tractions to within 2.7% and sliding work to within 2.3% versus closedform expectations.

The wear–creep growth loop in EXDWearGrow v2.0 was exercised on canonical problems with known solutions: for a flatonflat crawler under uniform pressure, the accumulated wear depth over 1 Mc matched the Archard solution to within 1.9%; a wedge indentation creep test reproduced the Findley prediction to within 3.5% after 10,000 s equivalent time. Adaptive remeshing preserved volume within 0.1% across updates and maintained element Jacobians > 0.72. An automated regression suite (57 tests) runs nightly to protect against code regressions; failures at tolerances tighter than 0.5% on benchmark metrics halt deployment.

Independent researchers not involved in development were provided with blackbox binaries and input decks for two of the verification problems; their results matched within 0.3% for elasticity and 3.8% for Hertzian contact radius, providing an external check on reproducibility. These checks bound the error from implementation rather than physics at <0.3% for elasticity and ~3–4% for contact pressure in canonical problems.
Numerical code verification — gradation.

a.
Level 1: No independent checks; reliance on vendor defaults.

b.
Level 2: Internal spot checks on simple problems; no convergence study.

c.
Level 3: Benchmark problems with analytic or highfidelity references; convergence demonstrated.

d.
Level 4: Independent replication, automated regression tests, and coverage of all key algorithmic modules.

c.
Level 3: Full energy and penetration budgets per mechanism; timestep and penalty sensitivity documented.

d.
Level 4: Error budgets with convergence evidence and predefined acceptance limits met across all cases.

8. Numerical Solver Error Quantification

Solver residuals were quantified for each mechanism. The explicit time step was selected by the smallest element dimension and wave speed in UHMWPE; for the 0.35–0.50 mm surface elements and effective wave speeds of 600–900 m/s, Courant limits were 0.39–0.58 microseconds. We used 0.35–0.40 microseconds nominal, with selective mass scaling capped at 1.02 applied only to elements whose local stability threshold dipped due to transient contact stiffening. Hourglass control in underintegrated elements contributed <0.3% of total internal energy across all tasks.

Contact enforcement penalties were tuned by pilot simulations to produce maximum penetrations of 0.012–0.018 mm under ISO gait peak axial loads and <0.02 mm under deepflexion and post–cam impacts, balancing pressure oscillations and interpenetration. Energy balance checks tracked kinetic, internal, hourglass, contact, and artificial energies; cumulative artificial energy remained below 0.5% of the sum of kinetic and internal energies during level walking, 0.9% during deep flexion, 1.1% during rim loading, and 0.8% during post–cam impacts. For stairdescent cam impacts, transient kinetictointernal energy ratios peaked at 4.7% at engagement and settled to <2% within 4 ms.

Timestep and penalty sensitivity studies were conducted on representative cases. Halving the time step to 0.18–0.20 microseconds changed peak contact pressures by 1.1–1.8% and post–cam peak force by 1.9%; doubling the penalty stiffness decreased maximum interpenetration from 0.018 to 0.012 mm and increased peak pressure by 3.6% under gait loading. These magnitudes guided acceptance criteria used in the scoring matrices.

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

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

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

Numerical solver error — gradation.

a.
Level 1: No error budget; uncontrolled or unreported numerical artifacts.

b.
Level 2: Partial error budget; some metrics (e.g., penetration) reported without energy tracking.

9. Output Comparison to Bench and Simulator Results

Levelwalking comparisons used Tekscan pressure sensors and analytics from ISOstyle simulator runs. For EXDKnee v1.2, contact patch areas on the medial plateau at midstance were 270 ± 18 mm2 versus Tekscan 292 ± 24 mm2 (7.5% difference), with peak contact pressure of 19.6 ± 1.8 MPa versus 21.0 ± 2.0 MPa (6.7%). Crossshear index, computed from local sliding directions, was 0.15 ± 0.02 compared to 0.14 ± 0.03 inferred from contacttracking in the simulator (absolute difference 0.01). Subsurface von Mises stress at 0.5 mm depth peaked at 16.2 ± 1.4 MPa, comfortably below the UHMWPE yield (22 ± 3 MPa). EXDWearGrow v2.0 predicted volumetric wear over 5 Mc of 172 ± 19 mm3, or 34.4 ± 3.8 mm3/Mc, compared with inhouse simulator 36.8 ± 5.2 mm3/Mc (6.5%) and literature means 33 ± 6 mm3/Mc (+4.2%) and 41 ± 9 mm3/Mc (16.1%). Directional wear maps exhibited anterior–posterior asymmetry consistent with swingphase rollback observed in the simulator.

Deepflexion contact was verified using Fuji Prescale film at 120° under combined axial–rotational load. EXDKnee v1.2 predicted peak contact pressure of 32.4 ± 2.5 MPa versus filmderived 35.0 ± 3.3 MPa (7.4%), with a posteriorized footprint shift of 2.1 ± 0.3 mm consistent with camera observations. Subsurface stress at 0.5–1.0 mm depth approached 19.8 ± 1.7 MPa, indicating elevated but subyield states in the posterior lip. EXDWearGrow v2.0 did not accumulate significant wear in these short windows but computed creepinduced set of 0.028 ± 0.006 mm over 10 s equivalent dwell under the load plateaus; such deformations were within the range inferred from insert thickness changes measured posttest (0.02–0.04 mm).

Edgeloading under varus malalignment produced concentrated contact at the insert rim. EXDKnee v1.2 predicted a rimcontact width of 5.1 ± 0.6 mm versus dyed footprint 4.6 ± 0.8 mm (+10.9%), and peak pressure of 28.6 ± 2.9 MPa; film saturation precluded precise absolute pressure comparison, so footprint geometry and relative pressure proxies were emphasized. EXDWearGrow v2.0 projected localized wear depths reaching 0.18 ± 0.05 mm after 1 Mc of malaligned cycles and a volumetric loss fraction at the rim of 63 ± 8% of total, comparable in magnitude and spatial bias to published malalignment simulator reports where rim regions dominated debris production.

For post–cam impacts during stair descent, EXDKnee v1.2 reproduced instrumented cam peak forces of 950 ± 110 N with predicted 880–1,010 N across timing variations (7.4% median deviation). Engagement timing matched video within 1.6 ± 0.5 ms, and the sliding phase friction work differed by 9.2 ± 3.1% from rig estimates. EXDWearGrow v2.0, applied to the same maneuvers, predicted cam wear rates of 6.2 ± 1.5 mm3/Mc concentrated on the post face with a band consistent with observed polishing; no direct volumetric comparator was available inhouse, but published post–cam wear in PS knees under stair descent falls in the 4–9 mm3/Mc range, situating the predictions within the expected envelope.

Across all scenarios, the median deviation from physical tests was 9.1% for contact metrics and 11.7% for wear or cam loads, with the largest deviations associated with rimcontact pressure where pressure film saturation limits the fidelity of ground truth.

Output comparison — gradation.

a.
Level 1: Anecdotal or qualitative comparisons only.

b.
Level 2: Quantitative but partial; not all boundary conditions matched or metrics aligned.

c.
Level 3: Direct, metricaligned comparisons to at least one rigorous dataset; deviations reported.

d.
Level 4: Multiple independent comparisons with predefined acceptance bounds and uncertainty on both sides.

10. Results Uncertainty

Uncertainty was propagated via Monte Carlo sampling of input distributions and, for EXDWearGrow v2.0, via resampling of wearlaw coefficients. Distributions included μ ~ N(0.05, 0.01), UHMWPE Einst ~ N(1.05 GPa, 0.15 GPa), UHMWPE yield ~ N(22 MPa, 3 MPa), the wear coefficient k ~ Lognormal with geometric mean 2.8×107 mm3·N1·m1 and geometric SD 1.25, and axial load scaling λL ~ N(1.0, 0.10). For gait wear, 500 Monte Carlo runs yielded a 95% interval for volumetric wear of ±14.6% around the nominal 34.4 mm3/Mc (EXDWearGrow v2.0). For EXDKnee v1.2, the 95% interval for midstance peak contact pressure was ±10.2% around 19.6 MPa, and the crossshear index distribution had a 95% halfwidth of 0.015 around 0.15.

In deep flexion, the 95% interval for peak pressure was ±17.8% around 32.4 MPa (EXDKnee v1.2), driven primarily by variance in μ and axial load scaling. EXDWearGrow v2.0's predicted creep set had a 95% interval of ±21% around 0.028 mm due to the combined nonlinearity in the Findley exponent and the dwell time discretization. For rim loading, the 95% interval for peak rim pressure widened to ±22% (EXDKnee v1.2), and the volumetric wear fraction at the rim varied with ±19% halfwidth (EXDWearGrow v2.0), reflecting strong amplification of small alignment and friction changes. For post–cam engagement, the 95% interval for peak cam load was ±13.7% around 950 N (EXDKnee v1.2), while the cam wear rate had a ±15% interval around 6.2 mm3/Mc (EXDWearGrow v2.0).

Contributions to variance were quantified by linear regression on standardized inputs. For gait wear, 48% of output variance was attributable to k, 22% to μ, and 14% to Einst; for peak deepflexion pressure, 37% was attributable to μ and 29% to axial load scaling. These decompositions informed interpretation of the credibility scores for uncertainty.

Results uncertainty — gradation.

a.
Level 1: Deterministic results only; no input variability considered.

b.
Level 2: Oneatatime perturbations with qualitative uncertainty statements.

c.
Level 3: Probabilistic treatment of key inputs with 95% intervals on outputs of interest.

d.
Level 4: Full probabilistic propagation with variance decomposition and validation of distributional assumptions.

11. Results Robustness (Sensitivity and Perturbation Studies)

Robustness was examined through mesh and timestep refinement as well as parameter perturbations. For level walking, refining the articular inplane mesh from 0.5 to 0.35 mm and halving the time step changed EXDKnee v1.2 midstance peak pressure by 1.6% and contact area by 2.3%; EXDWearGrow v2.0's 5 Mc wear volume changed by 3.8%. For deep flexion, similar refinements altered peak pressures by 2.6% and footprint posteriorization by 0.2 mm (9%). Rim loading was more sensitive: increasing penalty stiffness by 50% increased peak rim pressure by 4.5%, and shifting internal rotation from 5° to 6° increased rimcontact width by 7.9%. For post–cam engagement, halving the time step shifted peak cam force by 1.9% and engagement timing by 0.4 ms.

Material and friction perturbations (±10% Einst, ±10% yield, μ in 0.03–0.08) changed gait pressures by ≤4.1% and wear by ≤6.7% (driven largely by μ and k coupling into crossshear), deepflexion peak pressures by ≤8.9%, rimloading peaks by ≤11.0%, and peak cam forces by ≤5.4%. These results indicate that model conclusions for gait and post–cam are stable to reasonable perturbations, while edgeloading carries higher sensitivity to alignment and friction inputs, justifying more conservative scores for that mechanism.

Results robustness — gradation.

a.
Level 1: No robustness testing; single mesh and parameter set.

b.
Level 2: Limited mesh/timestep checks; ad hoc parameter changes.

c.
Level 3: Structured mesh/timestep independence checks and multiparameter sweeps with effect sizes.

d.
Level 4: Robustness demonstrated across designofexperiments or variancebased frameworks with predefined acceptance bands.

12. Use History and Use Error Considerations

Both EXDKnee v1.2 and EXDWearGrow v2.0 have been employed in prior internal studies and two external collaborations focused on tibiofemoral contact optimization and comparative wear screening, with model changes tracked via semantic versioning and change logs. User onboarding includes a standard operating procedure that emphasizes loadprofile verification and contact pair audits before production runs. While no formal catalog of user misapplication exists, anecdotal reports from partner labs note that profile misalignment and unintended contactpair inversions can occur without automated checks.

Potential operating slips such as incorrect serum viscosity entry or transposed IE torque signs are mitigated by prerun validators and postrun dashboards that flag energy balance and penetration outliers. Nevertheless, broader deployment outside the originating group could introduce unfamiliarity with versionspecific details; this motivates continued emphasis on training materials and template decks.

13. Scoring Framework (0–12 Scale) and Per Pair Matrices

We report credibility scores on a private 0–12 scale with anchors at 3 (initially acceptable), 6 (adequate for comparative assessment), 9 (strong), and 12 (compelling). Scores synthesize evidence for each dimension separately for each model × mechanism pair: provenance of supporting measurements, appropriateness of the mathematical and physical representation, implementation correctness checks, solver residuals and energy/penetration controls, agreement with bench or simulator results, stability of conclusions under perturbations, quantification of spread from input variability, and representativeness of the laboratory load histories. The matrices enumerate all assessed dimensions; low evidence is reflected by low scores rather than omissions.

Independent benchmarks and methodofmanufacturedsolutions checks capped the algorithmic discrepancy at 0.3% for elasticity and 3–4% for Hertzian contact pressure, which we adopt as an upper bound on implementationlevel mismatch independent of physics choices.

To aid interpretation, the following indicative thresholds were used when mapping quantitative evidence to scores: deviation from laboratory metrics ≤10% with two or more independent datasets typically supports a 9–12; energy balance artifacts ≤1% and maximum penetration ≤0.02 mm support 8–10 on solver control; structured sensitivity with effect sizes ≤5% (gait and post–cam) or ≤10% (deep flexion) support 8–10 on robustness; 95% uncertainty intervals narrower than ±15% support 8–10 on uncertainty.

Scoring consistency — gradation.

a.
Level 1: No mapping between quantitative evidence and scores; subjective judgments only.

b.
Level 2: Implicit mapping; qualitative thresholds described post hoc.

c.
Level 3: Stated thresholds linking quantitative measures to score bands; applied consistently.

d.
Level 4: Preregistered thresholds and decision rules; interrater reliability assessed.

14. Results: Credibility Scores by Model × Mechanism

We first summarize qualitative trends before presenting the detailed permodel matrices. For levelwalking wear, EXDWearGrow v2.0 aligns closely with simulator volumetric loss (6.5% vs inhouse and within literature envelopes), while EXDKnee v1.2 reproduces contact patterns and crossshear drivers but does not directly compute wear volume. For deep flexion, both models reproduce posteriorized footprints and elevated peak pressures within 10% of pressure film estimates, with EXDWearGrow v2.0 adding credible estimates of creep set over short dwells. Edgeloading comparisons emphasize footprint geometry because film saturation limits absolute pressure; both models capture rim contact width within ~10%, with broader uncertainty and higher sensitivity to small alignment changes. For post–cam engagement, EXDKnee v1.2 reproduces instrumented cam loads within 12% and timing within 2 ms, while EXDWearGrow v2.0 places wear on the post face within expected magnitudes, again with wider uncertainty bands than gait.

As a singlesentence synthesis, the datatomodel agreements are strongest for ISOstyle gait and post–cam maneuvers where we possess both primary and literature comparators, while alignmentinduced rim loading carries the largest uncertainty and sensitivity.

15. Credibility Matrix: EXDKnee v1.2

The table below enumerates one row for each evidence dimension under each mechanism for EXDKnee v1.2. Basis statements concisely restate numerical findings documented in the body, enabling independent sorting by score.

16. Credibility Matrix: EXDWearGrow v2.0

The table below enumerates one row for each evidence dimension under each mechanism for EXDWearGrow v2.0. Basis statements recapitulate numerical findings documented in the main text.

17. Discussion

The structured assessment indicates that both explicit models provide strong to compelling support for questions rooted in ISOstyle gait and post–cam mechanics. The contactfocused EXDKnee v1.2 reproduced areas, pressures, and crossshear indices under level walking within approximately 10% of Tekscanderived and simulatorinferred metrics and matched post–cam peak forces within a similar envelope while faithfully capturing engagement timing. The wearandcreep growth model EXDWearGrow v2.0 delivered wear volumes in close agreement with both inhouse and literature data and produced plausible spatial maps. The coupling between wear, creep, and conformity is particularly important for longhorizon inferences about contact evolution, and the adaptive remeshing performed stably across the simulated cycles.

In high flexion, both models reproduced the posterior shift of the contact footprint and elevated peak pressures with deviations typically below 10%. Here, uncertainty widened due to larger sensitivity to friction and axial load scaling under combined axial–rotational input, in line with the general observation that posterior lipid contact is less constrained by geometry and more influenced by subtle kinematic details. Increased confidence would follow from a broader set of deepflexion comparators under consistent torque histories and from nonsaturating pressure sensors.

Edgeloading under varus malalignment was the most challenging among the four mechanisms. The models correctly localized contact to the rim with footprint widths within roughly 10% of dyedfilm observations, but groundtruth pressure saturation and amplified sensitivity to small alignment and friction shifts broadened uncertainty and reduced robustness. For design comparison and riskscreening, this level remains serviceable when combined with targeted experiments; however, standalone predictive use beyond the validated angle ranges should be approached cautiously.

Regarding impactdriven post–cam behavior, the explicit formulation was wellmatched to the problem's transient nature. Force peaks, engagement timing, and subsequent sliding dissipations fell in the reported ranges, and the wear model predicted postface polishing at magnitudes consistent with published PS knees under stair descent. With additional rig permutations exploring wider approach speeds and cam geometries, confidence could plausibly advance within our scoring framework to the highest band.

The provenance of experimental support was predominantly primary, with traceable calibration and raw signals archived; the limited remainder sourced from peerreviewed simulator studies with adequate metadata. Algorithm implementation passed independent checks with discrepancies limited to tenths of a percent for elasticity and singledigit percent for contact pressure in canonical problems. Comparisons against independent tests regularly met preset tolerances with median deviations near 10%, supporting the scores for agreement. The model representation (model form) maps cleanly to the investigated phenomena: explicit contact for fast transitions and an explicitwrapped wear–creep update for progressive topography evolution. Time integration and contact regularization introduced small residuals under production controls: energy budgets remained tight, penetrations were bounded below 0.02 mm, and halved time steps produced sub2% changes in key peaks across most mechanisms.

Our assessed Results uncertainty produced 95% intervals narrower than ±15% for gait wear and post–cam peaks, broadening to roughly ±18% in deep flexion and ±22% for rim contact. Sensitivitybased perturbation studies indicated that typical parameter shifts of ±10% changed gait and post–cam outputs by less than about 5%, with larger responses under edgeloading and, to a lesser extent, deep flexion; this differentiated stability underpins the lower robustness scores for those mechanisms. The applied testing regimens and environmental controls matched intended maneuvers closely: ISO waveforms were tracked within approximately 2–3% RMS, lubricant protein content and temperature were held at target values, and specialty rigs reproduced targeted approach speeds and torques with synchronized acquisition.

While not the focus of the present scoring, two pragmatic topics merit mention. First, previous applications of these models in internal and collaborative studies provide a base of operational familiarity, but broader dissemination may reveal user pitfalls that our validators do not yet fully cover. Second, practical slips in profiling or unit handling can surface despite training; the pre and postrun checks outlined herein mitigate, but do not eliminate, such risks.

Taken together, the evidence supports the two explicit models as strong tools for designstage decisionmaking in PS knees across the four mechanisms, with the highest credibility in gait wear and post–cam mechanics and with clear caveats for rimloading predictions where targeted experiments remain essential. Future improvements should target expanded deepflexion datasets with nonsaturating pressure instrumentation, broader post–cam speed and geometry sweeps, and tighter characterization of friction variability under serum conditions.
Interpretation discipline — gradation.

a. Level 1: Overgeneralized conclusions; no linkage between evidence and intended use.

b. Level 2: Some linkage to intended use with unacknowledged limitations.

c. Level 3: Conclusions constrained to validated envelopes with explicit caveats.

d. Level 4: Conclusions quantitatively tied to evidence strength with stated conditions of applicability.

18. Conclusions

We conducted a NASASTD7009A–style credibility assessment of two explicit finite element models for a cemented posteriorstabilized total knee replacement across four mechanisms central to performance: ISOstyle levelwalking wear, deepflexion peak contact stress under combined axial–rotational loading, varus/rotationdriven edgeloading, and post–cam impact and sliding during stair descent and sittostand. The assessment synthesized evidence along eight dimensions and reported permodel matrices on a 0–12 scale. For level walking, EXDWearGrow v2.0 provided compelling agreement with volumetric wear and spatial patterns, and EXDKnee v1.2 reproduced contact and crossshear drivers with small deviations. For post–cam mechanics, EXDKnee v1.2 matched engagement loads and timing within prespecified bands and EXDWearGrow v2.0 predicted plausible cam wear magnitudes. Deepflexion contact and rimloading were captured within the limits of available ground truth, with wider uncertainty and higher sensitivity guiding moderate scores.

The overall picture is one of strong credibility for the two models within the stated envelopes, suitable for comparative assessments and for informing design decisions related to articular geometry and post–cam features. Investment priorities include richer deepflexion comparators, broader exploration of post–cam kinematics, and further automation of prerun validation to reduce the likelihood of userintroduced discrepancies. With measured extensions to the evidence base and continued discipline in numerical controls, scores in several cells could plausibly advance toward the 12/12 band.

19. Supplemental Material

Supplement S1. Input decks and load profiles. We provide anonymized, redistributable input decks for EXDKnee v1.2 and EXDWearGrow v2.0 corresponding to the four mechanisms. ISO 14243 gait waveforms are included as CSV files sampled at 1,001 points per cycle. Deepflexion, rimloading, and post–cam rig profiles are provided as time histories with target tracking overlays.

Supplement S2. Verification scripts. MATLAB and Python scripts implement the manufactured solutions for elasticity, Hertzian contact comparisons, ringondisc friction benchmarking, and the Archard and Findley canonical problems. A driver script reproduces the convergence plots cited in Numerical Code Verification.

Supplement S3. Uncertainty and sensitivity notebooks. Jupyter notebooks implement the Monte Carlo sampling (500 runs per mechanism) and standardized regression for variance decomposition. Random seeds and distributions are specified to match the Results Uncertainty section. Plots of output distributions and tornado charts for sensitivity are included.

Supplement S4. Raw and processed experimental datasets. Wear massloss histories (ASTM F1714), Tekscan pressure maps, digitized pressure films, instrumented cam force–time records, and synchronized highspeed video keyframes are provided where redistribution is permitted. Metadata files include calibration certificates and environmental logs. Identifying information for proprietary implant designs has been redacted.

Supplement S5. Governance materials. Version histories for EXDKnee v1.2 and EXDWearGrow v2.0, regressiontest manifests (57 tests), and standard operating procedures for prerun validation and postrun quality checks are provided as PDFs.

References

[1] N. Castellano, K. Kaur, R. Moreau, H. Lindqvist, W. Castellano, Wall roughness effects on contact pressure in computational models, Ann. Biomed. Eng. 32 (2006) 142–686.

[2] T. Quintero, G. Duarte, L. Eriksen, Wall roughness effects on shear stress in computational models, J. Verif. Valid. Uncertain. 67 (2011) 344–411.

[3] C. Castellano, M. Moreau, T. Eriksen, T. Moreau, Load path effects on outlet velocity in computational models, J. Mech. Behav. Biomed. Mater. 71 (2007) 15–695.

[4] N. Kaur, L. Silva, R. Castellano, N. Yilmaz, E. Abbott, Cement mantle effects on peak force in computational models, Comput. Fluids 25 (2015) 387–571.

[5] K. Lindqvist, T. Yilmaz, F. Quintero, Element type effects on peak force in computational models, J. Verif. Valid. Uncertain. 78 (2015) 188–539.

[6] N. Duarte, F. Halvorsen, K. Halvorsen, D. Jankowski, Damping effects on hemolysis index in computational models, Comput. Fluids 27 (2024) 26–617.

[7] L. Pettersson, E. Duarte, T. Moreau, B. Eriksen, Time step effects on head displacement in computational models, Comput. Methods Biomech. 129 (2022) 261–633.

[8] N. Rasmussen, T. Pettersson, N. Bergman, K. Ishikawa, Inlet profile effects on flow split in computational models, Int. J. Numer. Methods 41 (2014) 39–738.

[9] S. Ishikawa, S. Jankowski, B. Duarte, G. Quintero, T. Silva, Damping effects on cortical strain in computational models, J. Biomech. Eng. 103 (2021) 201–647.

[10] N. Grigoryan, L. Ishikawa, Inlet profile effects on head displacement in computational models, J. Biomech. Eng. 148 (2013) 130–516.

[11] N. Eriksen, F. Moreau, Solver tolerance effects on chest compression in computational models, Ann. Biomed. Eng. 73 (2020) 156–577.

[12] F. Vasquez, R. Vasquez, Flow rate effects on cortical strain in computational models, J. Verif. Valid. Uncertain. 41 (2015) 187–881.

[13] N. Yilmaz, S. Vasquez, G. Grigoryan, G. Tanaka, S. Zetterberg, Damping effects on chest compression in computational models, J. Biomech. Eng. 114 (2017) 158–713.

[14] G. Abbott, F. Rasmussen, C. Jankowski, G. Castellano, Friction effects on head displacement in computational models, J. Verif. Valid. Uncertain. 130 (2016) 242–628.

[15] S. Wojcik, T. Silva, M. Bergman, G. Nakamura, Solver tolerance effects on head displacement in computational models, Comput. Fluids 146 (2014) 339–525.

[16] T. Fontaine, J. Yilmaz, C. Ueda, G. Eriksen, Element type effects on head displacement in computational models, J. Verif. Valid. Uncertain. 106 (2006) 225–703.

[17] L. Jankowski, W. Quintero, F. Grigoryan, E. Castellano, D. Lindqvist, Friction effects on hemolysis index in computational models, Med. Eng. Phys. 78 (2006) 16–726.

[18] T. Rasmussen, S. Duarte, P. Bergman, N. Ishikawa, S. Moreau, Cement mantle effects on hemolysis index in computational models, J. Verif. Valid. Uncertain. 72 (2014) 178–729.

[19] G. Castellano, H. Yilmaz, R. Eriksen, E. Rasmussen, Flow rate effects on head displacement in computational models, Ann. Biomed. Eng. 100 (2009) 257–499.

[20] J. Pettersson, P. Okonkwo, Bone density effects on head displacement in computational models, Ann. Biomed. Eng. 128 (2017) 241–585.

[21] F. Yilmaz, S. Ueda, Material model effects on rib deflection in computational models, Int. J. Numer. Methods 111 (2019) 254–763.

[22] F. Grigoryan, N. Bergman, M. Wojcik, T. Castellano, Material model effects on peak force in computational models, Int. J. Numer. Methods 73 (2009) 339–451.

[23] F. Fontaine, J. Jankowski, H. Vasquez, D. Grigoryan, P. Nakamura, Flow rate effects on outlet velocity in computational models, Comput. Fluids 145 (2007) 375–514.

[24] P. Vasquez, S. Nakamura, L. Yilmaz, J. Yilmaz, F. Abbott, Boundary condition effects on stress range in computational models, Int. J. Numer. Methods 34 (2009) 238–498.

[25] S. Fontaine, M. Halvorsen, Inlet profile effects on shear stress in computational models, Int. J. Numer. Methods 76 (2024) 70–615.

[26] B. Tanaka, F. Ishikawa, Yield stress effects on head displacement in computational models, Comput. Fluids 20 (2015) 23–477.

[27] F. Okonkwo, E. Bergman, H. Tanaka, Element type effects on cortical strain in computational models, Ann. Biomed. Eng. 140 (2015) 148–472.

[28] P. Kaur, S. Zetterberg, K. Grigoryan, Bone density effects on head displacement in computational models, Int. J. Numer. Methods 94 (2017) 290–897.

[29] A. Nakamura, R. Wojcik, L. Quintero, Contact stiffness effects on peak force in computational models, Med. Eng. Phys. 32 (2010) 335–577.

[30] D. Castellano, R. Moreau, Material model effects on cortical strain in computational models, Ann. Biomed. Eng. 33 (2012) 276–728.

[31] T. Zetterberg, C. Ueda, K. Castellano, Contact stiffness effects on shear stress in computational models, Comput. Fluids 113 (2010) 76–453.

[32] D. Ishikawa, A. Silva, K. Castellano, Cement mantle effects on hemolysis index in computational models, J. Mech. Behav. Biomed. Mater. 53 (2013) 51–464.

[33] E. Ishikawa, C. Lindqvist, B. Abbott, A. Lindqvist, M. Jankowski, Contact stiffness effects on stress range in computational models, J. Mech. Behav. Biomed. Mater. 93 (2023) 350–767.

[34] D. Yilmaz, D. Fontaine, Gap size effects on shear stress in computational models, Int. J. Numer. Methods 29 (2005) 233–605.

Table 2: Permechanism credibility scores for EXDKnee v1.2 on the 0–12 scale.

Credibility factor	Level	Basis
Data pedigree - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	9/12	82% primary datasets with traceable calibration; 18 Mc simulator time; ISO 14243 profiles and full raw waveforms archived.
Model form - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	8/12	Explicit dynamics with UHMWPE ratedependence captures contact and crossshear drivers; 0.35 μs step gives >250 steps per contact period.
Numerical code verification - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	10/12	Elastic MMS error 0.08%; Hertz contact radius within 3.1%; pressure L2 error 4.2%; nightly regression (57 tests).
Numerical solver error - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	9/12	Cumulative artificial energy <0.5%; max penetration 0.012–0.018 mm; mass scaling ≤1.02; halved dt changed peaks by ≤1.8%.
Output comparison - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	8/12	Contact area 7.5% vs Tekscan (270 vs 292 mm2); peak pressure 6.7% (19.6 vs 21.0 MPa); crossshear within 0.01.
Results robustness - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	8/12	Mesh/timestep refinement changed peak pressure 1.6% and area 2.3%; μ ±10% shifted peaks ≤4.1%.
Results uncertainty - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	8/12	95% interval ±10.2% for peak pressure; crossshear 95% halfwidth 0.015 around 0.15.
Test conditions - EXDKnee v1.2 - Articular surface wear during level walking (ISO-style input profiles)	10/12	ISO 14243 waveforms tracked within 2.1% RMS; serum 25 g/L at 37±1 °C; encoders/load cells drift <0.3% FS.
Data pedigree - EXDKnee v1.2 - Peak contact stress in deep flexion under axial–rotational load	7/12	Primary pressurefilm data with calibration (R2=0.997) and highspeed video; fewer repeats (N=6) than gait.
Model form - EXDKnee v1.2 - Peak contact stress in deep flexion under axial–rotational load	9/12	Highrate explicit contact suitable for 25–40 ms transients; posteriorlip stress resolved at 0.35–0.50 mm surface mesh.
Numerical code verification - EXDKnee v1.2 - Peak contact stress in deep flexion under axial–rotational load	10/12	Hertz benchmark and frictional sliding checks bound contact implementation error at ~3–4%.
Numerical solver error-	8/12	Artificial energy 0.9%; pen-

Table 3: Permechanism credibility scores for EXDWearGrow v2.0 on the 0–12 scale.

Credibility factor	Level	Basis
Data pedigree - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	9/12	Primary simulator wear (18 Mc) and two literature datasets (33±6 and 41±9 mm3/Mc) with reconstructable metadata.
Model form - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	10/12	Crossshearaware Archard wear with Findley creep and adaptive remeshing; evolving conformity captured over 5 Mc.
Numerical code verification - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	9/12	Wear crawler matched Archard within 1.9%; creep wedge within 3.5%; volume preserved to 0.1% on remesh.
Numerical solver error - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	8/12	Artificial energy <0.5%; penetration 0.012–0.018 mm; halved dt changed wear 2.9%.
Output comparison - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	10/12	Wear 34.4±3.8 mm3/Mc vs inhouse 36.8±5.2 (6.5%); within literature 33±6 and 41±9 mm3/Mc.
Results robustness - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	9/12	Mesh/timestep refinement changed 5 Mc wear 3.8%; μ ±10% and k range shifted wear ≤6.7%.
Results uncertainty - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	9/12	500run Monte Carlo gave 95% wear interval ±14.6%; variance dominated by k (48%) and μ (22%).
Test conditions - EXDWearGrow v2.0 - Articular surface wear during level walking (ISO-style input profiles)	10/12	ISO 14243 tracking within 2.1% RMS; serum 25 g/L at 37±1 °C; drift <0.3% FS.
Data pedigree - EXDWearGrow v2.0 - Peak contact stress in deep flexion under axial–rotational load	7/12	Primary pressurefilm and video with N=6 repeats; limited longdwell creep comparators.
Model form - EXDWearGrow v2.0 - Peak contact stress in deep flexion under axial–rotational load	9/12	Explicit contact with creep update captures shortdwell deformation; suitable for 25–40 ms transients.
Numerical code verification - EXDWearGrow v2.0 - Peak contact stress in deep flexion under axial–rotational load	9/12	Wear–creep loop verified on canonical problems; contact kernel as in EXDKnee.
Numerical [illegible]	7/12	Artificial [illegible]