Credibility assessment of explicit-dynamics head protection models under ASME V&V40 for skull and brain kinematics

Elena Markham, PhD[a], Ravi Prakash, PhD[b], Marta C. Ionescu, MS[c], Daniel K. Huang, PhD[d]

[a] *Impact Biomechanics Laboratory, Department of Mechanical Engineering, North River University, USA,*
[b] *Computational Mechanics Group, Advanced Safety Systems Inc., USA,*
[c] *Human Injury Research Institute, Center for Neuroprotection, USA,*
[d] *Department of Biomedical Engineering, Eastport Technical University, USA,*

Abstract

Objective: To assess, under the ASME V&V40 risk-informed framework, the credibility of two explicit-dynamics finite element analysis (FEA) models of a cranial impact protection system for predicting skull and brain kinematics across normal and oblique impact mechanisms relevant to falls and sports collisions. Methods: Two computational models were evaluated. The Helmeted Rigid-Skull Kinematics (HRK) model uses a rigid cranium surrogate to quantify center-of-gravity linear and angular accelerations, force time histories, and contact metrics. The Skull–Brain Dynamics (SBD) model includes an anatomically detailed skull, meninges, cerebrospinal fluid (CSF) layer, and brain to estimate skull kinematics, brain bulk motion, and tissue-level strain. Mechanisms included normal drop and oblique impact with tangential loading. Activities encompassed numerical code checks and benchmark problems, sensitivity to spatial and temporal resolution, rationale for contact/material/boundary conditions, alignment of validation scenarios to the Context of Use (COU), documentation and traceability of software processes, and test articles used for comparison. A private 0–8 Credibility Evidence Scale (CES) with half-point increments was applied to each factor for each (model × mechanism) pair. Predefined acceptance thresholds were set by Model Risk Level (High) and the COU. Results: For HRK, normal drop yielded CES scores (numerical code: 7.5; mesh/time step: 7.0; representation choices: 6.5; COU alignment: 6.5; process control: 7.0; empirical comparators: 6.0). Oblique yielded (7.5; 6.5; 5.5; 7.0; 7.0; 6.0). For SBD, normal drop yielded (7.5; 6.5; 6.0; 6.5; 7.0; 5.5). Oblique yielded (7.5; 6.0; 5.0; 7.0; 7.0; 5.5). Mesh refinement reduced changes in kinematic outputs to ≤3% and in 95th-percentile brain strain to ≤6% at the chosen resolutions. Explicit solver conservation, benchmark comparisons, and hourglass control metrics remained within pre-set tolerances. The selected material, contact, and boundary condition ideals captured gross trends in force–time and kinematic responses, while late-time rotational rebound remained underpredicted in oblique cases. Validation scenarios and metrics matched the COU outputs and covered the majority of exposure directions and severities targeted for pre-market design selection. Repository-based software management with automated regression suites supported reproducibility. The physical comparators included NOCSAE-style drop tests (n = 48), an oblique anvil rig at 30° and 45° (n = 36), and five post-mortem human subject (PMHS) oblique impacts. Conclusions: Evidence strength is sufficient to use HRK as the primary comparative model for design screening across the two mechanisms, and to use SBD as a supportive model for screening and hypothesis generation on rotational mitigation and strain-based endpoints. The principal improvement opportunities are increased high-severity oblique testing with matched input characterization, additional damping and slip-interface calibration to capture rotational rebound, and extended convergence demonstrations for tissue-level metrics.

Keywords: head protection, verification and validation, explicit dynamics

1. Introduction

Protective headgear is tasked with moderating the exchange of momentum and energy between the human head and the environment, suppressing peak forces and attenuating linear and angular kinematics associated with skull fracture, focal contusions, and diffuse brain injury. As product variants multiply—new shell geometries, liner foams, slip interfaces—designers increasingly rely on simulation to compare alternatives prior to and in parallel with physical testing. Because these models feed decisions that can influence safety-critical choices, their credibility must be established in a risk-informed manner and aligned with the decisions they are intended to support.

The ASME V&V40 standard provides a structured approach that ties verification, validation, and uncertainty considerations to a Context of Use (COU) and an assigned Model Risk Level. Under V&V40, evidence is marshalled across lines of inquiry—accuracy of the computational method, appropriateness of modeling choices for the phenomena of interest, congruence between validation scenarios and decision needs, stability of the implementation, and quality of the data used for comparison—and scored against acceptance levels commensurate with the consequences of error. Here we implement this approach for two explicit-dynamics finite element models of a cranial impact protection system intended to attenuate skull and brain kinematics during representative impacts.

The present work focuses on two mechanistic loading pathways: a normal drop impact and an oblique impact with tangential loading. The former challenges axial crushing and load spreading to control linear acceleration and peak force; the latter engages shear and bending in the shell–liner–head assembly and provokes head rotation through coupled normal–tangential interface forces. To support design selection and test planning, we assess two computational models. The Helmeted Rigid-Skull Kinematics (HRK) model represents the head as a single rigid body to isolate interactions in the helmet assembly and provides center-of-gravity (CoG) accelerations and impact forces. The Skull–Brain Dynamics (SBD) model resolves skull, meninges, CSF, and brain dynamics to produce both skull kinematics and tissue-level strain metrics in the cerebrum, which are increasingly used as surrogates for diffuse injury risk.

We employ a private Credibility Evidence Scale (CES) ranging from 0 to 8 to grade the strength of evidence for each factor per (model × mechanism) combination. Scores are paired with text that records the basis for the grade and with numbers traceable to methods and results presented below. The analysis proceeds in the style of V&V40, with acceptance thresholds dictated by a High Model Risk Level appropriate to pre-market design decisions that could otherwise underrate rotational risk. The goal is to make transparent, for each model and mechanism, where evidence is strong, where it is weak or incomplete, and how that alignment affects whether, and how, the model may be used.

This paper is structured as follows. We summarize the device and modeling approaches, define the COU and risk level, and describe the two mechanisms. We then present the computational formulations and the activities bearing on method accuracy, spatial–temporal resolution, and the rationale for the constitutive, contact, and boundary choices. We outline how the validation activities relate to the outputs and decision questions, detail the comparators used for model-to-data evaluation, and describe software process controls. With a scoring method and acceptance thresholds specified, we then report factor-by-factor CES scores for each model and mechanism and synthesize the implications for the COU.

Context of Use clarity — gradation.

a.
Decision not clearly stated; audience and outputs undefined; model risk level absent.

b.
Decision stated but outputs and metrics only loosely mapped; preliminary risk rationale.

c.
Decision, outputs, and metrics explicitly mapped; risk rationale documented with traceability to stakeholders.

d.
Decision, outputs, metrics, and model limitations co-defined with stakeholders; risk-based acceptance thresholds pre-specified and justified.

2. Device and explicit-dynamics modelling overview

The device under evaluation is a cranial impact protection system comprising an outer thermoplastic shell, an energy-absorbing liner made of a multi-density expanded polystyrene (EPS) foam, a low-friction slip layer, comfort padding, and a retention system. The intent is to reduce peak CoG linear acceleration in normal impacts and to moderate angular velocity and angular acceleration in oblique scenarios by combining crushable foam deformation with controlled sliding at the head–helmet interface. The two models share CAD geometry for the shell and liner and differ primarily in their head representations and the scope of predicted outputs.

Both models are implemented in an explicit-dynamics FEA code widely used for impact and crash applications. Elements are predominantly hexahedral for the foam and tetrahedral for shell features with complex curvature. The HRK model treats the headform as a rigid body with mass and inertia matching a 50th percentile anthropomorphic head (mass 4.5 kg, principal moments of inertia about CoG of 0.016, 0.018, 0.020 kg·m^2). The SBD model includes an anatomically detailed skull segmented from a representative CT dataset, a dura–arachnoid complex, a CSF layer modeled as a nearly incompressible viscoelastic medium via an ALE formulation interfaced to the Lagrangian brain, and a brain parenchyma modeled as transversely isotropic visco-hyperelastic tissue with nominal axonal fiber alignment in the hemispheres derived from diffusion tensor imaging templates.

For the shell, an elastic–plastic orthotropic material with strain-rate sensitivity is used; coupon tests at nominal strain rates of 50, 200, and 800 s1 were conducted to calibrate the Cowper–Symonds parameters and directional yield stresses (S0 = 45–58 MPa depending on direction; C = 1600 s1; p = 4.5). The EPS liner is represented by a crushable foam model with volumetric compaction; uniaxial compression curves were measured at 5, 50, and 250 s1 to define the plateau stress and densification onset as functions of nominal density (28, 40, and 55 kg/m^3). For the slip layer, we used a rate- and pressure-dependent Coulomb friction with a baseline friction coefficient of 0.25 and a limiting static peak of 0.32 under low slip speed and low interface pressure, based on ring-on-disk tests using a polyurethane scalp surrogate against the fabric laminate of the slip layer.

Loading and boundary conditions match, to the extent feasible, standard impact test setups. The normal drop impact is implemented by releasing the helmeted headform from heights corresponding to impact energies of 5.9, 9.8, and 14.7 J (nominal drop heights 0.15, 0.25, and 0.375 m for the 4.5 kg mass) onto a rigid flat anvil. The oblique impact uses impact speeds of 4.0, 6.0, and 7.5 m/s onto abrasive-coated anvils set at 30° and 45° from the horizontal, with the headform CG path aligned to contact the anvil at a target location over the parietal-temporal region. In both mechanisms, gravity is included and headform neck constraints are omitted to reflect head-only configurations used in the corresponding rigs.

The HRK model outputs include CoG linear acceleration time histories (resolved components and resultant), angular

Figure 1: Overview of the cranial impact protection geometry, load cases and the mechanisms assessed, shown across the full page width.

acceleration and velocity computed by integrating contact moments divided by the mass moments of inertia, contact force distribution, and interface slip distance. The SBD model outputs include the same skull-level kinematics plus brain bulk angular velocity and acceleration relative to skull, and tissue-level maximum principal strain (MPS) reported at 95th-percentile across the cerebrum using element-wise Green–Lagrange strain. Postprocessing includes standard filters (CFC 1000 for accelerations) and rigid-body kinematics extraction from nodal data at the headform CoG.

Model scope adequacy — gradation.

a.
Key structures or physics absent; outputs cannot address decision.

b.
Major structures present; significant physics approximated by surrogates; outputs cover primary metrics.

c.
All critical structures and interactions included; secondary physics approximated; outputs cover primary and secondary metrics.

d.
Comprehensive multiphysics with validated couplings; outputs directly trace to decision metrics with quantified margins.

3. Context of Use: pre-market comparative performance evaluation to inform design selection and test planning; Model Risk Level: High (device safety decision influence is direct; erroneous predictions could underrate rotational risk)

The COU is to compare alternative helmet design variants—shell geometry, liner density grading and vent patterns, and slip-layer material/coefficient changes—before committing to tooling and extensive certification testing. The questions to be answered by simulation are: (1) do variants A, B, and C reduce peak resultant CoG linear acceleration under normal drops relative to a baseline; (2) do they reduce peak CoG angular acceleration and moderate peak angular velocity in oblique impacts; and (3) do modifications materially affect brain tissue strain predictions under oblique loading? These answers direct which physical tests to prioritize and which design variants to carry forward.

The outputs supporting these decisions are peak resultant linear acceleration, peak resultant angular acceleration, time-to-peak for both kinematic measures, peak contact force, and 95th-percentile MPS in the brain for oblique impacts. The range of intended impacts reflects sports-play and fall conditions catalogued in prior epidemiological and instrumented-head datasets: normal impact speeds spanning 2.5–4.0 m/s and oblique resultant speeds spanning 4–8 m/s at incidence angles of 30° to 60°. The tested mechanisms cover the more severe portion of this range.

Given that incorrect ranking or masking of high angular acceleration could compromise safety, the Model Risk Level is High. Misclassification consequences include advancing a design with inferior rotational mitigation or discounting a design that would have materially improved rotational protection. Therefore, strong evidence is required for those factors most directly bearing on predicted kinematic magnitudes and time histories, and moderate evidence is sufficient for supporting process-oriented factors that primarily affect traceability and reproducibility rather than physical fidelity.

Risk-informed acceptance — gradation.

a.
No linkage between model risk and acceptance levels.

b.
Qualitative linkage; acceptance levels not tied to decision impact.

c.
Quantitative acceptance thresholds for key outputs matched to risk level.

d.
Tiered acceptance thresholds with sensitivity to consequences and mitigations, reviewed with stakeholders.

4. Mechanisms: normal drop impact and oblique impact with tangential loading

Normal drop impacts represent a vertical transfer of kinetic energy into the helmet–anvil system. Liner crushing under the crown or parietal contact is the dominant mode, with shell flexural contributions affecting how load is distributed. The key observables in this mechanism are peak resultant linear acceleration and peak normal force at the contact, with angular kinematics playing a secondary role due to minor asymmetries or off-axis placement.

Oblique impacts add a tangential component at the contact with the anvil, which couples to shell bending and interface shear to produce head rotation. The interaction of the slip layer with the skin surrogate or scalp governs the transition from stick to sliding, and the rate dependence of both friction and foam response shapes the torque–time history on the head. In this regime, both peak angular acceleration (which is sensitive to early-time stick) and peak angular velocity (shaped by the integrated torque minus damping) are critical. Brain tissue strain is driven by relative motion of the skull and brain modulated by the CSF layer's ability to shear and by dura–arachnoid tethering.
Mechanism representativeness — gradation.

a.
Mechanism does not capture dominant real-world loading pathways.

b.
Mechanism captures a subset of relevant pathways; key couplings absent.

c.
Mechanism captures dominant pathways and key couplings; secondary pathways approximated.

d.
Mechanism comprehensively captures primary and secondary pathways with empirically supported couplings.

5. Computational Models: HRK and SBD formulations and outputs

Helmeted Rigid-Skull Kinematics (HRK) model: The headform is instantiated as a single rigid body with mass and inertia matched to a 50th percentile anthropometric head. The helmet assembly is meshed with 2.0–3.0 mm shell elements in areas of high curvature and 3.5–4.0 mm elsewhere; liner foam uses 2.5–3.5 mm hexahedra with through-thickness refinement to resolve compaction fronts. Contact between shell and anvil is modeled with penalty enforcement and a calibrated friction coefficient of 0.5 against abrasive surfaces in oblique tests; slip-layer-to-head contact uses the measured friction model with pressure and rate dependence. Outputs are extracted at 10 µs sampling and filtered per CFC 1000.

Skull–Brain Dynamics (SBD) model: The skull is a 1.5 mm cortical layer backed by trabecular bone, meshed with 1.0–2.0 mm tetrahedral elements in the cranial vault and refined to 0.8–1.0 mm near foramina and sutures. The CSF is represented in an ALE domain with nominal thickness of 2.0 mm over the hemispheres and narrower under the skull base; a viscosity of 0.7 mPa·s is used, and density is 1000 kg/m^3. The dura–arachnoid is a 0.4 mm membrane with orthotropic elastic properties. The brain parenchyma is meshed with 1.5–2.0 mm hexahedra in the cerebrum and 1.0–1.5 mm tetrahedra in the cerebellum and brainstem where geometry forces transitions. The visco-hyperelastic constitutive law uses a neo-Hookean base (G0 = 1.2 kPa) with Prony series terms (G1 = 0.8 kPa, τ1 = 8 ms; G2 = 0.6 kPa, τ2 = 45 ms) and an anisotropic reinforcement along white matter directions with a small-fiber modulus of 0.3 kPa. Coupling between ALE CSF and Lagrangian brain uses a penalty-based FSI with hourglass control and bulk viscosity tuned to avoid spurious oscillations. Outputs include skull CoG kinematics and 95th-percentile MPS derived from Green–Lagrange strains transformed into the laboratory frame.

Common solver settings: A global explicit time step is computed by the smallest element dimension and material wave speed; typical stable time increments are 0.2–0.4 µs, and mass scaling is not used. Numerical damping is limited to a linear bulk viscosity with Q1 = 0.05 and Q2 = 1.0 to stabilize shocks. Hourglass control uses a stiffness formulation for reduced-integration elements with coefficient set such that hourglass energy remains below 0.5% of total internal energy throughout the event. Contact penalty stiffnesses are calibrated to avoid excessive interpenetration (<0.5 mm at peak load) without driving instability.

Output processing: Linear and angular accelerations are calculated and filtered to CFC 1000; peak values and times to peak are extracted from each signal. For SBD, MPS is computed element-wise; we report the 95th-percentile across the cerebrum and time to peak MPS. Brain bulk rotation is measured as the angular velocity of the brain mass about the skull-fixed coordinate frame to capture skull–brain relative motion.
Output metric traceability — gradation.

a.
Outputs are not defined or not linked to measurable quantities.

b.
Outputs defined but transformations/filters undocumented; traceability partial.

c.
Outputs defined with documented filters and transformations; traceable to standards.

d.
Outputs defined, standardized, and cross-checked against multiple measurement modalities; traceability audited.

6. Numerical code verification (explicit solver stability checks and benchmark problems)

We exercised a suite of problems to probe algorithmic correctness of the explicit solver used for both models. First, element-level checks included linear elastic patch tests in tension and shear for 8-node hexahedra and 4-node shells, with numerical solutions matching analytical stress fields within 0.15% L2-norm error at single-element and multi-element configurations. A distorted-mesh patch test (skew angles up to 35°) produced errors below 0.35%. Second, energy conservation was tested by a free vibration of a constrained shell plate initialized with kinetic energy but no external work; total energy remained within ±0.2% over 10 ms with bulk viscosity disabled. Third, a one-dimensional shock tube test (Sod problem) using the ALE formulation reproduced the analytic solution for density and velocity profiles with mean absolute error below 1.2% at a Courant number of 0.5. Fourth, the Taylor anvil problem was simulated with copper material at 200 m/s impact; final length and mushroom radius matched experimental correlations within 1.1% and 2.0%, respectively.

To probe contact algorithms, a sphere-on-plate impact was run at 3.0 m/s with penalty contact and a range of friction coefficients; coefficient calibration recovered known restitution trends, and normal impulses matched momentum-theory predictions within 0.9%. For hourglass control, reduced-integration foam elements were compressed uniaxially under velocity-controlled boundary conditions; hourglass energy never exceeded 0.4% of total internal energy. With contact, foam crushing under a rigid indenter at 1.5 m/s produced force–displacement curves that matched measured quasi-one-dimensional compaction data within 2.8% RMS error after rate correction.

Solver stability under severe contact and shear was tested via a helmet shell segment impacting an inclined rigid plane at 7.5 m/s using the same contact and friction settings as the oblique rig; no non-physical interpenetrations exceeding 0.3 mm occurred, and the explicit time step remained above 0.18 µs without mass scaling. A suite of 112 regression cases covering these verifications and representative helmeted impacts was configured to run on continuous integration (CI) infrastructure. For each run, we tracked hourglass energy fraction, energy balance (kinetic + internal external work), contact penetrations, and maximum stable time step; flagged thresholds were 0.5%, 1.0%, 0.5 mm, and 0.15 µs, respectively.

Numerical code verification — gradation.

a.
No method-of-manufacture checks; solver results unbenchmarked.

b.
Basic element and conservation checks; limited comparison to analytics.

c.
Comprehensive patch, conservation, and canonical impact benchmarks; tolerances met and monitored.

d.
Extended verification suite with automated regression, provenance capture, and independent replication.

7. Discretization error (mesh convergence for helmet shell, skull, and brain; time-step sensitivity)

We undertook spatial and temporal resolution studies for both models and both mechanisms using uniform and targeted refinements. For HRK in normal drops, three mesh levels were used for the shell (element edge length 4.0, 3.0, 2.2 mm) and liner (3.5, 2.8, 2.2 mm). Peak resultant CoG linear acceleration at the 9.8 J condition decreased from 99.6 g to 98.5 g to 97.0 g across the three levels, indicating a 1.5% difference between the latter two. Peak normal force decreased from 5.34 kN to 5.25 kN to 5.18 kN (1.3% change between the last two). Halving the time step from 0.32 to 0.16 µs at the middle mesh changed peak linear acceleration by 0.8% and peak force by 0.6%. For HRK oblique impacts at 6.0 m/s, the mesh levels were as above; peak resultant angular acceleration decreased from 4,950 rad/s^2 to 4,820 rad/s^2 to 4,710 rad/s^2 (2.3% between the last two), while peak angular velocity shifted from 31.2 rad/s to 30.7 rad/s to 30.3 rad/s (1.3% change for the last pair). Time-step halving changed angular acceleration by 1.1%.

For SBD, spatial refinements targeted the skull base, falx–tentorium interfaces, and temporal lobes. Normal drops at 9.8 J used brain element sizes of 2.0, 1.6, and 1.3 mm in cerebrum (and proportional reductions elsewhere). Peak resultant skull linear acceleration decreased from 98.9 g to 97.6 g to 96.9 g (0.7% between the last two), while the 95th-percentile MPS was 0.184, 0.179, and 0.175 (2.2% between the last two). Time-step halving changed peak MPS by 1.6%. For oblique impacts at 6.0 m/s and 30°, peak skull angular acceleration changed from 4,930 to 4,780 to 4,690 rad/s^2 (1.9% between the latter two), while 95th-percentile MPS changed from 0.238 to 0.224 to 0.214 (4.5% between the latter two). Time-step halving at the middle mesh changed the latter by 2.9%. We observed that peak MPS is slightly more sensitive to mesh refinement than kinematic outputs, particularly in regions of high strain gradients near the temporal cortex and corpus callosum.

To further probe temporal integration accuracy, we ran a subset of cases with the bulk viscosity reduced by 50% and monitored the stability of MPS evolution; the timing of the MPS peak shifted less than 0.5 ms, and its amplitude changed less than 2.0% for both mechanisms at the middle and fine meshes. For HRK, reducing the contact penalty stiffness by 20% at the fine mesh increased peak penetration to 0.42 mm and changed peak angular acceleration by 0.7%, indicating robustness to moderate variations in contact stiffness at the chosen resolutions.

We also performed Richardson extrapolation on peak linear acceleration for normal drop at the HRK fine meshes, assuming first-order convergence due to contact nonlinearity, resulting in an extrapolated value of 96.2 g and a Grid Convergence Index (GCI) of 1.9% for the fine mesh at a safety factor of 1.25. For the SBD 95th-percentile MPS in oblique impact, a similar approach with an assumed order of 1.0 yielded an extrapolated MPS of 0.208 and a GCI of 2.9% for the fine mesh.

Spatial–temporal resolution sufficiency — gradation.

a.
Single mesh/time step; no sensitivity results.

b.
Two levels of refinement; qualitative or partial sensitivity results.

c.
Three or more levels; quantitative sensitivity and GCI-like reporting; tolerances met for key outputs.

d.
Targeted local refinements and time-step studies with extrapolation; error budgets allocated to outputs.

8. Model form (contact, material models, and boundary condition rationale per mechanism)

Material and contact ideals were chosen to represent the helmet assembly's dominant responses under the two mechanisms. For the shell, an elastic–plastic, orthotropic, rate-sensitive model captures bending and membrane behavior under short-duration loads. Calibration against coupon tests and a shell-panel impact at 5 m/s yielded stress–strain predictions within 5% RMS error across rates and directions. The EPS liner crushable foam law includes a plateau region and densification at volumetric strains beyond 0.6, reproducing energy absorption without stiffness overshoot; quasi-one-dimensional compaction and spherical-indenter tests supported the parameterization. The slip layer relies on a pressure- and rate-dependent Coulomb model that reproduces stick–slip transitions; ring-on-disk and pin-on-disk tests with the scalp surrogate supplied the friction map over slip rates of 0.05–2.0 m/s and interface pressures of 5–20 kPa.

For the HRK model, these choices emphasize accurate helmet–anvil–headforce interactions and allow the rigid head surrogate to transmit resulting moments and forces to the CoG. For the SBD model, skull stiffness and inertia are represented anatomically, while CSF and brain laws enable relative motion and strain localization. A visco-hyperelastic brain law with modest anisotropy was selected to represent the short-time response needed for 0–30 ms impacts; parameter values were drawn from shear and compression measurements at 1–200 s1 and from inverse identification against PMHS head rotation data. The CSF representation as a viscous fluid in an ALE domain allows shear and squeeze to mediate skull–brain motion transfer while avoiding locking.

Boundary conditions replicate standard rigs. The normal drop constrains the anvil with a rigid support and sets initial conditions via gravitational potential and an initial drop height. The oblique impact sets an anvil at 30° or 45° and assigns initial translational velocity to the headform/helmet with rotational velocity zero, consistent with launch carriage techniques; the abrasive surface on the anvil enforces significant tangential force buildup before sliding commences at the slip interface. In both mechanisms, neck effects are not modeled, as the validation data use head-only fixtures; for field relevance beyond the rigs, these would need to be added and validated but are outside the present COU.

Calibration against system-level tests included a bare-helmet (no slip layer) oblique rig series at 6.0 m/s, used to tune shell and liner interface friction (0.5 ± 0.05). The slip layer map was not altered after its independent characterization. Damping was kept minimal to avoid artificially suppressing rotational peaks; residual damping was set to reproduce the decay of the late-time tail in shell panel impacts. The choices reflect a trade-off: emphasizing fidelity to early-time torque buildup at the cost of some late-time underdamping at the slip interface. This was accepted given the COU focus on peak values and times to peak.

To interrogate model structure effects, we conducted alternative runs with a purely plastic foam formulation (no volumetric compaction law) and observed overprediction of peak linear acceleration by 8–10% in normal drop; similarly, replacing the rate- and pressure-dependent friction with a constant 0.25 resulted in a 12–18% increase in peak angular acceleration at 45° oblique impacts due to prolonged early-time stick. These sensitivity tests suggest that selected form choices are material to the outputs of interest.

Form selection and justification — gradation.

a.
Key constitutive and contact forms are undocumented or ad hoc.

b.
Form choices partially justified with limited calibration against components.

c.
Form choices justified with component and subsystem tests; sensitivities explored.

d.
Form choices justified against multi-scale evidence; alternates explored and rejected with quantitative rationale.

9. Relevance of the validation activities to the COU (metrics and scenarios aligned to skull and brain kinematics targets)

Validation activities were chosen to connect directly to the kinematic outputs and decision questions in the COU. For normal drops, we executed drop tower tests with the helmet on a NOCSAE-style headform across nominal drop heights that matched the simulation energies of 5.9, 9.8, and 14.7 J. We instrumented the headform with tri-axial accelerometers at the CoG, processed via CFC 1000, and extracted peak resultant linear acceleration and time to peak. For oblique impacts, we used a pendulum-based rig with a rigid anvil set at 30° and 45°, abrasive surface treatment, and high-speed video to track head rotation; angular velocity histories were obtained from video tracking and integrated to angular acceleration using smoothing splines, providing an independent measure against the IMU-based estimates at the headform CoG.

To address brain-level validation targets, we used published PMHS datasets with reflective targets implanted into the brain parenchyma and skull to quantify skull–brain relative motion under oblique impact. Specifically, five impacts with resultant speeds 5.5–7.0 m/s and incidence angles 30–45° were used to extract peak brain bulk angular velocity and an approximate 95th-percentile MPS derived from inverse dynamic modeling in the original studies. While these reconstructions have uncertainty, they provide a directionally consistent check on whether the SBD model captures both the timing and magnitude of strain-based endpoints.

The range of loading conditions in the test program maps to the higher-frequency, moderate-to-high-severity slice of the targeted exposure distribution for the intended use activities. The outputs gathered from the tests—peak resultant linear acceleration, angular acceleration, angular velocity, and MPS-related surrogates—correspond to the COU metrics. To enhance comparability, we used the same impact locations and helmet sizes across simulation and experiment to the extent permitted by manufacturing tolerances, and processed signals with identical filters and definitions.

COU-to-validation alignment — gradation.

a.
Validation scenarios and metrics unrelated to decision metrics.

b.
Partial overlap between validation and decision metrics; limited scenario overlap.

c.
Direct overlap of metrics and good coverage of scenarios; processing harmonized.

d.
Comprehensive overlap with traceable mapping of scenarios, metrics, and tolerances to decision questions.

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

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

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

10. Test samples (physical test articles, headforms, and cadaveric datasets used for comparison)

We assembled a set of physical comparators designed to span the mechanisms and outputs of interest. For normal drops, forty-eight tests were executed using a drop tower with a rigid flat anvil and a NOCSAE-style headform. Twelve tests per energy level (5.9, 9.8, 14.7 J) were conducted, split across three helmet sizes (S, M, L), with four repeats each. Accelerometers at the headform CoG were sampled at 20 kHz and filtered at CFC 1000; peaks and time to peaks were recorded along with contact force from a load cell under the anvil. The helmets were conditioned at 21 ± 2 °C and 50 ± 10% relative humidity for 24 hours prior to testing to control environmental effects on foam response.

For oblique impacts, thirty-six tests were executed on a pendulum-based rig with a rigid anvil at 30° and 45°, 18 per angle, with impact speeds of 4.0, 6.0, and 7.5 m/s. Headform instrumentation included a tri-axial gyro and tri-axial accelerometer package (IMU) at the CoG; high-speed video at 10,000 fps captured head rotation to cross-check angular velocity from the IMU. The anvil surface used abrasive paper with a mean roughness of 120 grit to ensure significant tangential contact forces. Tests were split across sizes S, M, L with matched helmet units to the extent possible.

To inform brain-level validation, we accessed five PMHS oblique impact cases with head-only fixtures and implanted targets to measure skull–brain relative motion. Impact speeds were 5.5–7.0 m/s at 30°–45°, and output metrics included brain bulk angular velocity and reconstructed MPS measures from the original analyses. The uncertainty in these reconstructions was acknowledged; rather than strict numeric matching, we sought consistency in timing and magnitude ranges to support directional conclusions for the SBD model.
Empirical sample strength — gradation.

a.
Few or no tests; unrepresentative conditions; poor instrumentation.

b.
Moderate number of tests; partial coverage of conditions; standard instrumentation.

c.
Substantial test count with coverage across sizes and severities; redundant instrumentation.

d.
Extensive test matrix with environmental control and multi-modal measurements tailored to decision metrics.

11. Software quality assurance (version control, defect tracking, regression tests)

Model assets, solver decks, and postprocessing scripts were managed in a Git repository with branching via a feature-branch workflow and protected main branch. All changes underwent peer review and automated style/lint checks prior to merge. A test harness executed 112 regression cases, including element patch tests, shock tube, Taylor anvil, helmet panel impacts, and representative normal and oblique helmeted impacts. For each case, numeric tolerances were defined for peak outputs and energy and hourglass metrics. Build artifacts were stored with cryptographic checksums to ensure traceability, and configuration files specified solver versions and platform libraries to support repeatability.

Defect tracking used an issue tracker with severity coding and linkage to commits and pull requests. Two critical solver-interface defects were identified during development: one related to a friction law option that did not properly update the rate dependence at high slip speeds, and another concerning an hourglass control parameter that could be overwritten by a contact initialization routine. Both were corrected and verified before the analysis baseline was frozen. Releases were tagged, and batch runs were executed only on tagged states; logs and outputs were archived in a write-once storage and mirrored to a secondary server.

Environment control included containerized solver environments to minimize configuration drift. A nightly build re-ran a reduced subset of regression cases and compared key metrics against baselines. Deviations beyond tolerance triggered alerts and were investigated before proceeding with production runs.
Software process control — gradation.

a.
Ad hoc scripts and files; no version control or tests.

b.
Basic version control; limited peer review; a few spot checks.

c.
Structured version control with peer review, issue tracking, and automated regression tests.

d.
Full lifecycle controls with traceability, containerized environments, and auditable releases.

12. Equivalency of input parameters (nominal material and friction values and their mapping to test articles)

Nominal material properties and interface coefficients were mapped from laboratory measurements to the models and test articles through shared conditioning and processing protocols. EPS density and rate curves were taken from batches used to mold the liners for testing, and slip-interface friction maps were measured using fabric lots that matched the installed slip layer. For the shell, coupon lots were culled from the same sheet runs used to thermoform the shell halves. Helmet size-specific geometries and mass properties were adjusted in the models and were selected to match the tested units. The pressure and rate regimes probed in friction testing included those expected in both normal and oblique mechanisms. These mappings ensured consistent provenance of inputs across simulation and test while acknowledging that within-lot variability and aging effects may still influence both.
Input mapping transparency — gradation.

a.
Inputs drawn from literature or guesswork; no mapping to test articles.

b.
Partial mapping of inputs; some shared provenance; gaps in conditions.

c.
Systematic mapping from measured inputs with shared conditioning; parameter ranges cover scenarios.

d.
Formal equivalency demonstrations with uncertainty bounds propagated into predictions.

13. Credibility scoring method (0–8 CES) and acceptance thresholds

We applied a private Credibility Evidence Scale (CES) that assigns a score from 0 to 8 to each credibility factor for each (model × mechanism) combination. A score of 0 denotes an absence of evidence; 2 indicates minimal, preliminary evidence; 4 indicates moderate, partially convergent evidence; 6 indicates substantial, convergent evidence from multiple lines; and 8 indicates convergent, multi-source evidence with residual uncertainties small relative to decision tolerances. Half-point scores are permitted to reflect intermediate states. The scale is ordinal; absolute distances are not linear in risk reduction but provide a consistent relative gauge across factors.

For this COU and a High Model Risk Level, we specify acceptance thresholds for the primary outputs: for HRK and SBD, the scores for the solver-method checks and for spatial–temporal resolution must be at least 6.5 to proceed with design comparisons; structure-and-idealization choices must be at least 5.5 for HRK and 5.5 for SBD to ensure the modeled interactions reflect the mechanisms of interest; alignment to the COU must be at least 6.0 for both models to ensure scenarios and metrics are representative; process control must be at least 6.5 to guarantee reproducibility; and the strength of empirical comparators must be at least 5.5 to reduce the risk of overfitting or misinterpretation. These thresholds are set to limit the likelihood of accepting a design whose predicted kinematic mitigation is illusory due to numerical or structural model weaknesses.

We weighted factors equally within each model–mechanism evaluation for transparency, but we emphasized in the narrative synthesis where a factor creates a gating limitation for the COU. For example, if rotational metrics in oblique impacts depend heavily on interface slip behavior, then the structural representation will dominate the interpretation, even if other factors score high. Conversely, high process control cannot compensate for poor mapping to the COU.
Evidence scale calibration — gradation.

a.
Scale not defined; thresholds absent.

b.
Scale defined; thresholds qualitative.

c.
Scale defined with quantitative thresholds per risk level and output.

d.
Scale defined, calibrated against historical decisions, and sensitivity-tested.

14. Results: factor scores for HRK under normal drop impact (per factor on 0–8 CES)

We report here the graded evidence and supporting numbers for the HRK model under normal drop impacts. For checks of the computational method, the explicit solver exhibited energy balance within ±0.2% in free vibration and resolved the shock tube and Taylor anvil with errors below 1.2% and 2.0%, respectively; in helmeted normal-drop regressions, hourglass energy remained below 0.4% and contact penetrations below 0.4 mm. This body of results supports a high score for method checks in this mechanism.

Resolution sensitivity for HRK normal drops shows small changes between the fine and finest meshes: peak resultant CoG linear acceleration at 9.8 J decreased from 98.5 g to 97.0 g (1.5%); peak normal force decreased from 5.25 kN to 5.18 kN (1.3%); and halving the time increment from 0.32 µs to 0.16 µs changed peaks by less than 0.8%. At 5.9 and 14.7 J, the percentage changes were similar (1.1–2.0%). A Richardson extrapolation at 9.8 J gave a GCI of 1.9% for peak linear acceleration at the fine mesh with a safety factor of 1.25.

Representation choices for materials, contact, and boundary conditions were exercised through a comparison to system-level normal drops at 5.9, 9.8, and 14.7 J. The model predicted peak resultant linear acceleration with mean absolute error of 4.3 g across energies compared to experiments (simulation 68.9 ± 4.1 g at 5.9 J vs test 72.1 ± 5.0 g; 97.6 ± 5.8 g vs 101.2 ± 6.6 g at 9.8 J; 132.4 ± 7.3 g vs 138.5 ± 8.0 g at 14.7 J). Force–time curves showed similar shapes with a slight underprediction of the early rising slope and an RMSE of 7.8% relative to the experimental force waveforms normalized by peak force. Interface slip was minimal in normal drops, and friction coefficients had negligible influence on outputs at these conditions.

Alignment to the COU is supported by the test matrix mapping to the three energies of interest and by direct correspondence of outputs and processing (CFC 1000, peak extraction). The range 5.9–14.7 J falls within the expected severity for the target activities; the comparison of peak CoG linear acceleration and force is exactly the set of metrics used for the design screening questions. The location of impact (crown and parietal) and head-only fixture are consistent with the rigs used in both simulation and experiment.

Process control is strong: all HRK normal-drop runs are traceable to tagged repository versions; the regression suite recorded peak linear acceleration, peak force, energy metrics, and penetration with tolerances. No regression case exceeded tolerance during the production period; nightly runs flagged two minor deviations tied to a platform library update, which were reverted before proceeding. The versioned datasets and CI logs permit reproduction of all reported results.

The empirical comparators consist of forty-eight normal drop tests spanning the three energies and three helmet sizes. Instrumentation and processing match the model outputs. The sample size supports statistical estimates of mean and variance; however, aging effects and headform skin variability were not exhaustively probed. The comparators are therefore sufficient to capture gross trends and energy scaling but leave minor environmental and material-variability effects underexplored.

Evidence integration for HRK-normal — gradation.

a.
Sparse, inconsistent evidence; major gaps; low decision utility.

b.
Moderate, partly convergent evidence; specific gaps qualify use.

c.
Substantial, convergent evidence on key outputs; residual gaps documented.

d.
Comprehensive, multi-source convergence; small residuals relative to tolerances.

15. Results: factor scores for HRK under oblique impact (per factor on 0–8 CES)

For the HRK model under oblique impact with tangential loading, method checks remain consistent with those described above; the contact-heavy validations were stable at 7.5 m/s, with penetrations below 0.3 mm and hourglass energy below 0.4% of internal energy. No solver instabilities were observed during the oblique production runs.

Resolution sensitivity for oblique HRK runs showed peak resultant angular acceleration at 6.0 m/s and 30° decreasing from 4,950 rad/s^2 to 4,820 rad/s^2 to 4,710 rad/s^2 across coarse, fine, and finest meshes, a 2.3% change between the fine and finest. Peak angular velocity decreased from 31.2 to 30.7 to 30.3 rad/s (1.3% between the latter two). Halving the time step at the fine mesh changed peak angular acceleration by 1.1% and peak angular velocity by 0.7%. These changes are within preset tolerances for kinematic outputs (≤5%).

The structural and idealization choices for materials, contacts, and boundary conditions captured early-time torque buildup but underpredicted rotational rebound in the 45° impacts. In the 6.0 m/s, 45° condition, peak angular acceleration was 4,980 ± 420 rad/s^2 in tests and 4,560 ± 260 rad/s^2 in simulation (8.4% bias); peak angular velocity was 35.2 ± 2.1 rad/s in tests and 32.9 ± 1.8 rad/s in simulation (6.5% bias). The timing to peak angular acceleration matched within 0.9 ms. The underprediction at late time suggests that damping and slip-interface representation could be improved to recover rebound dynamics more faithfully, though the early peaks relevant to design screening are close.

Alignment to the COU is favorable: the selected angles (30°, 45°) and speeds (4.0, 6.0, 7.5 m/s) cover the oblique severities emphasized in the COU, and the outputs (peak angular acceleration and velocity) and processing (CFC 1000, video cross-checks) directly match the decision metrics. Spatial impact location and head-only fixture mirror the validation rigs and simulations.

Process control for the HRK oblique suite mirrors the normal-drop suite. A friction-law rate dependence defect identified early in development was corrected and verified with unit tests before oblique runs were launched. All runs are traced to tagged baselines and reproduced with containerized environments; CI tolerances held across the production runs.

The oblique empirical comparators include thirty-six tests across two angles and three speeds, with redundant measurements (IMU and high-speed video) and controlled abrasive surfaces. The dataset spans three helmet sizes; however, there is limited coverage of environmental temperature and humidity extremes, and the headform skin condition was not varied systematically. Consequently, while the comparators are strong for the primary outputs and conditions, they may not fully bracket interface variability.

Evidence integration for HRK-oblique — gradation.

a.
Sparse, inconsistent evidence; major gaps in oblique behavior.

b.
Moderate evidence; early-time torque captured but gaps in rebound modeling.

c.
Substantial evidence on kinematic peaks; residual late-time underprediction documented.

d.
Comprehensive oblique evidence with multi-physics calibration including rebound phases.

16. Results: factor scores for SBD under normal drop impact (per factor on 0–8 CES)

For the SBD model in normal drop impacts, method checks are identical to those documented earlier. ALE–Lagrange coupling in the shock tube and contact-rich helmeted cases remained stable, and hourglass energy was controlled under 0.4% of internal energy throughout. The explicit solver's conservation and benchmark performance supports a high score for method verification in this mechanism as well.

Resolution sensitivity for SBD normal drops shows low sensitivity for skull-level kinematics and slightly higher sensitivity for tissue-level strain. At 9.8 J, peak resultant skull CoG linear acceleration changed from 98.9 g (coarse) to 97.6 g (fine) to 96.9 g (finest), a 0.7% change between the latter two. The 95th-percentile MPS decreased from 0.184 to 0.179 to 0.175, a 2.2% change between the latter two. Halving the time increment at the fine mesh altered peak MPS by 1.6%. Richardson extrapolation suggested a GCI of 1.9% for peak acceleration and 2.9% for MPS at the fine mesh, using a safety factor of 1.25.

The choices of brain visco-hyperelastic law and CSF representation have limited consequence under normal drops, where linear acceleration dominates and rotational demands are small. Predicted peak resultant linear acceleration across energies matched tests with mean absolute error within 4.6 g (simulation 69.3 ± 4.5 g vs 72.1 ± 5.0 g at 5.9 J; 97.1 ± 5.3 g vs 101.2 ± 6.6 g at 9.8 J; 131.9 ± 7.2 g vs 138.5 ± 8.0 g at 14.7 J). The 95th-percentile MPS remained modest (0.12–0.18 across energies), consistent with expectations. The system-level fit indicates that the additional internal degrees of freedom do not compromise fidelity at the skull level in this mechanism.

Alignment to the COU is satisfied through direct matching of energies, impact locations, and outputs; the skull-level kinematic outputs and processing are identical to HRK. While brain-level MPS is included for completeness, it is not a COU decision driver in this mechanism, and we do not emphasize it in rankings.

Process control is the same as for HRK, as both models share the same repository and CI infrastructure; SBD-specific regressions include CSF–brain coupling tests and brain-element patch tests to guard against inadvertent changes in coupling parameters or material laws. No regressions failed during production runs.

The empirical comparators for normal drops are the same set of forty-eight tests. The addition of brain-level validation is limited in this mechanism; we did not have PMHS data for pure vertical drops with brain instrumentation. As a result, while skull-level comparisons are strong, tissue-level validation in this mechanism remains indirect.

Evidence integration for SBD-normal — gradation.

a.
Sparse skull-level match; no tissue-level checks.

b.
Good skull-level match; limited tissue-level surrogates.

c.
Strong skull-level match; tissue-level outputs stable and within expectations.

d.
Strong skull- and tissue-level convergence with targeted validation.

17. Results: factor scores for SBD under oblique impact (per factor on 0–8 CES)

For the SBD model under oblique impact with tangential loading, the explicit solver's stability and energy checks are consistent with those reported earlier. ALE–Lagrange coupling remained robust, and contact penetrations were below 0.3 mm in production runs. Hourglass energy was less than 0.4% of internal energy throughout.

Resolution sensitivity for SBD oblique runs shows a higher strain sensitivity than for skull-level kinematics, as anticipated. At 6.0 m/s and 30°, 95th-percentile MPS decreased from 0.238 (coarse) to 0.224 (fine) to 0.214 (finest), a 4.5% change between the latter two; halving the time step at the fine mesh changed peak MPS by 2.9%. Skull-level peak angular acceleration changed from 4,930 to 4,780 to 4,690 rad/s^2 (1.9% between the latter two). These values remain within preset tolerances (≤10% for MPS, ≤5% for kinematics) at the selected meshes.

The material and contact choices captured early-time torque and bulk brain rotation but tended to underpredict late-time rotational rebound and the tail of the angular velocity curve. At 6.0 m/s, 45° impacts, peak brain bulk angular velocity was 38.2 ± 2.4 rad/s in the PMHS reconstructions and 36.1 ± 2.2 rad/s in simulation (5.5% bias), with peak timing within 1.2 ms. The 95th-percentile MPS at the same condition was 0.235 ± 0.018 in reconstructions and 0.214 ± 0.016 in simulation (9.0% bias). In 30° impacts at 4.0 m/s, biases were smaller (3.1% in ω and 4.2% in MPS). The structure appears adequate to rank design variants by relative mitigation but may not yield tight absolute estimates of MPS without additional tuning.

Mapping to the COU is strong for this mechanism: the angles and speeds span the targeted oblique severities, and the outputs—peak angular acceleration and velocity at the skull and 95th-percentile MPS—are directly aligned. The data sources include both headform tests and PMHS directional checks for brain motion and strain proxies, which reinforces relevance to the decision questions.

Process control is equivalent to that of the HRK model, with SBD-specific regression cases including CSF–brain coupling and tissue response checks. All runs are traceable and reproducible via tagged repositories and containerized environments. No regression cases failed during oblique production runs.

The empirical comparators for this mechanism include thirty-six headform oblique tests and five PMHS cases with brain instrumentation and reconstructions, providing both skull-level and brain-level anchors. However, the PMHS data have uncertainties from marker tracking and inverse estimation, and environmental sweeps are limited. The comparators are therefore supportive for directional validation but not definitive for absolute MPS magnitudes.

Evidence integration for SBD-oblique — gradation.

a.
Sparse oblique tissue-level evidence; major uncertainties.

b.
Moderate skull-level match; limited brain-level checks.

c.
Substantial skull-level match and directional brain-level consistency; residual underprediction at late time.

d.
Comprehensive skull- and brain-level validation with quantified uncertainties and correction factors.

18. Synthesis across model–mechanism pairs: strengths, gaps, and implications for the stated COU

Across both models and both mechanisms, several patterns emerge that bear on how the models should be used for the COU. The rigid-head HRK model consistently predicts skull-level kinematic peaks with small bias in both normal and oblique impacts, particularly at early times. The SBD model preserves this skull-level fidelity and adds the ability to interrogate brain bulk motion and tissue strain, with slightly larger sensitivity to spatial–temporal resolution and a modest tendency to underpredict late-time rotational response in oblique impacts. For comparative design ranking on kinematic peaks, HRK is well-suited as a primary screening tool; for insights into how design variants change tissue-level strain under oblique loading, SBD is informative, with caveats regarding absolute magnitudes.

The test program aligns well with the COU metrics and mechanisms, supporting transfer of validation insights to the decision questions. Normal drops cover energy levels used in screening, and oblique conditions span angles and speeds most relevant to rotational mitigation. The oblique tests include redundant measurements (IMU and high-speed video), increasing confidence in the skull-level outputs. The PMHS oblique dataset, while limited, provides an anchor for brain bulk motion and strain timing and scale.

From a numerical standpoint, solver performance is robust across cases of interest; energy balances and hourglass controls are within tolerance, and contact algorithms function without pathology. Spatial and temporal resolution studies show that skull-level outputs are relatively insensitive at the selected meshes, with changes under 3% for kinematic peaks; tissue-level strains in SBD show greater sensitivity, with changes under 6% between the fine and finest levels in oblique cases. These levels are below the pre-specified tolerances for the COU (≤5% for kinematic peaks; ≤10% for MPS).

At the level of structural representation, the models reproduce the phase and magnitude of early impact dynamics well but tend to mute the rebound phase of rotation in oblique impacts, consistent with residual damping at the slip interface and potential underrepresentation of interfacial compliance. This affects peak angular velocity more than peak angular acceleration and has a greater effect on SBD tissue-level strain than on skull-level kinematics. For comparative ranking, these tendencies reduce the risk of false positives (claiming large rotational mitigation when none exists) but could conceal small improvements in designs that mainly influence rebound phases.

Process controls and traceability are solid, enabling reproducibility and auditability of results. The regression suite, tagged releases, and containerized environments limit the risk of spurious differences due to software drift. The empirical comparators are adequate in count and coverage for the COU's primary aims, though they could be expanded to explore environmental and aging effects and to increase PMHS coverage at the higher end of oblique severities.

Considering the above, we judge that the HRK model can be used as the primary comparative tool for ranking design variants by their ability to reduce peak CoG linear acceleration in normal drops and peak CoG angular acceleration in oblique impacts; the SBD model can be used alongside HRK to evaluate directional changes in brain bulk motion and MPS across designs, with recognition that absolute MPS values carry additional uncertainty and should be corroborated by targeted experiments when feasible.
Use recommendation clarity — gradation.

a.
No explicit use recommendation; implications unclear.

b.
General recommendation; conditions and caveats not specified.

c.
Specific recommendation tied to evidence strengths and gaps; caveats documented.

d.
Recommendation optimized to COU with explicit boundaries, triggers for re-assessment, and monitoring plan.

19. Discussion

The evidence assembled for both models satisfies, to varying degrees, the needs of the COU under a High Model Risk Level. The HRK model's parsimony is advantageous: fewer internal degrees of freedom simplify the interpretation of skull-level kinematics and reduce susceptibility to compounding uncertainties. Its limitations surface most clearly in oblique rebounds, where interfacial damping and compliance, if imperfectly represented, can cause underprediction of angular velocity tails. Nonetheless, early-time peaks—which are central to the decision questions—are recovered well enough for comparative ranking.

The SBD model, by contrast, offers insights into phenomena that underlie diffuse injury mechanisms, such as the distribution of strain within the cerebrum and the timing of brain bulk rotation relative to skull motion. The balance between brain constitutive modeling and CSF shear representation is delicate; modest changes in viscosity or anisotropy can alter how relative motion translates into tissue deformation. Our sensitivity and convergence studies indicate that at the chosen mesh and time-step settings, brain-level metrics are numerically stable within the tolerances we set for the COU. The larger residual uncertainty is structural: the degree to which the simplified CSF and dural tethering abstractions faithfully capture skull–brain coupling under rapidly applied tangential loads.

These observations suggest several avenues for improvement. First, the representation of the slip layer could be enriched by allowing a compliance layer and a more detailed state-dependent friction law, calibrated to multi-axis shear tests that sweep pressure and sliding history. Second, damping allocation between bulk materials and interfaces should be revisited, perhaps using system identification against instrumented oblique tests that emphasize the rebound phase. Third, additional PMHS tests at 7.5–9.0 m/s and 30°–60°, with improved brain instrumentation or noninvasive imaging, would strengthen tissue-level validation, especially for the high end of the severity range.

Decisions in product development must often proceed with incomplete information. The practical question is whether the residual uncertainties are understood, bounded, and small enough relative to the expected design deltas. Here, expected design-driven changes in peak angular acceleration between variants typically exceed 10–15%, while the numerical and structural uncertainties we observed for that metric are on the order of 3–8%; therefore, the models are likely to resolve rank ordering with appropriate confidence. For 95th-percentile MPS, expected design-driven changes can be smaller (5–10%); given residual underprediction and sensitivity, we advise using SBD MPS as a directional indicator to identify promising design trends that should then be tested physically.

Finally, the interpretation of comparative results relies on the assumption that inputs in simulation and tests are commensurate. While we have mapped material properties and interface coefficients from shared batches and conditioning protocols, the equivalence of installed properties in assembled helmets and headforms remains an area for continuous attention. Differences in adhesive thickness, pad fit, and scalp surrogate wear can modulate effective interface behavior. Ongoing characterization within the test program should be used to update and bound the input space used in simulations.

In summary, the modeling program offers credible support to the COU's design-screening questions while highlighting where additional investment would most efficiently reduce residual risk: enhanced interface modeling to capture rebound, extended oblique PMHS datasets for tissue-level checks, and broader environmental sweeps in the test program to quantify context-dependent variability.

Improvement prioritization — gradation.

a. No priorities identified; improvements ad hoc.

b. General areas for improvement listed; priorities unclear.

c. Specific improvements prioritized by impact on decision risk.

d. Prioritized improvements with estimated risk reduction and a plan for re-assessment.

20. Conclusions

Under the ASME V&V40 framework and a High Model Risk Level, we evaluated two explicit-dynamics head protection models against a suite of activities spanning method checks, spatial–temporal resolution, representation choices, alignment to the COU, process control, and empirical comparators. Using a 0–8 Credibility Evidence Scale, we found that the HRK model attains substantial, convergent evidence for normal and oblique kinematic peaks and is well-suited for comparative design screening. The SBD model complements HRK by providing tissue-level strain estimates in oblique impacts; its evidence is sufficient for directional use in design decisions, with recognition that absolute MPS values require cautious interpretation.

Across models and mechanisms, discretization error remained below 3% for kinematic peaks and below 6% for 95th-percentile strain at the selected meshes and time steps, which are within the pre-set tolerances for the COU. The selected contact, material, and boundary idealizations captured early dynamics and magnitude trends in both mechanisms, while underpredicting rotational rebound in oblique cases, indicating that interface damping and compliance could be refined. Numerical code verification showed energy conservation within 0.2%, benchmark comparisons within 1–2%, and controlled hourglass energy below 0.5% of internal energy, indicating stable and accurate implementation of the explicit methods over the regimes of interest. The validation scenarios and metrics matched the COU's outputs and covered angles, speeds, and energies that account for the majority of exposures targeted by the decision context, supporting the use of the models for pre-market design selection and test planning. Process control was maintained through versioned repositories, automated regression tests (112 cases), and containerized environments, with two critical defects identified and corrected before baseline freeze and all production runs traceable by checksum. The experimental comparators comprised forty-eight normal drops and thirty-six oblique tests across sizes and severities, supplemented by five PMHS oblique cases for brain motion and strain surrogates; while robust for primary outputs, the samples left environmental variability and aging effects as secondary areas to expand.

Conclusion robustness — gradation.

a. Conclusions not tied to evidence; claims exceed data.

b. Conclusions loosely tied to evidence; scope creep present.

c. Conclusions supported by evidence; scope and caveats clear.

d. Conclusions fully supported, with quantified residual risk and explicit conditions of validity.

References

[1] L. Jankowski, J. Yilmaz, S. Abbott, D. Tanaka, Wall roughness effects on hemolysis index in computational models, Int. J. Numer. Methods 84 (2009) 292–833.

[2] W. Moreau, T. Castellano, T. Abbott, R. Ishikawa, C. Abbott, Material model effects on cortical strain in computational models, J. Mech. Behav. Biomed. Mater. 92 (2006) 136–820.

[3] M. Wojcik, J. Abbott, F. Jankowski, Gap size effects on contact pressure in computational models, J. Biomech. Eng. 56 (2012) 365–864.

[4] R. Quintero, J. Ishikawa, E. Pettersson, Boundary condition effects on contact pressure in computational models, Ann. Biomed. Eng. 118 (2005) 286–876.

[5] R. Vasquez, D. Vasquez, Time step effects on hemolysis index in computational models, Med. Eng. Phys. 28 (2013) 132–565.

[6] B. Rasmussen, W. Fontaine, B. Pettersson, Inlet profile effects on contact pressure in computational models, Int. J. Numer. Methods 93 (2022) 397–837.

[7] H. Kaur, D. Castellano, C. Nakamura, C. Castellano, F. Eriksen, Element type effects on peak force in computational models, Comput. Fluids 61 (2012) 190–811.

[8] P. Nakamura, B. Castellano, L. Lindqvist, F. Abbott, E. Wojcik, Inlet profile effects on fatigue life in computational models, J. Verif. Valid. Uncertain. 144 (2013) 375–651.

[9] W. Zetterberg, J. Pettersson, F. Lindqvist, F. Ueda, N. Duarte, Boundary condition effects on outlet velocity in computational models, Med. Eng. Phys. 101 (2013) 362–595.

[10] J. Halvorsen, D. Jankowski, C. Quintero, Solver tolerance effects on chest compression in computational models, Ann. Biomed. Eng. 36 (2012) 378–707.

[11] G. Ueda, F. Wojcik, R. Vasquez, H. Jankowski, H. Moreau, Friction effects on shear stress in computational models, J. Mech. Behav. Biomed. Mater. 109 (2013) 31–603.

[12] F. Lindqvist, S. Kaur, R. Eriksen, Material model effects on chest compression in computational models, Ann. Biomed. Eng. 55 (2008) 218–868.

[13] R. Rasmussen, H. Yilmaz, Yield stress effects on peak force in computational models, Med. Eng. Phys. 35 (2018) 311–819.

[14] K. Lindqvist, N. Fontaine, S. Ueda, Element type effects on head displacement in computational models, Comput. Methods Biomech. 89 (2014) 175–744.

[15] A. Tanaka, R. Nakamura, L. Bergman, Bone density effects on cortical strain in computational models, J. Verif. Valid. Uncertain. 54 (2020) 138–632.

[16] T. Vasquez, T. Silva, Material model effects on chest compression in computational models, J. Verif. Valid. Uncertain. 142 (2008) 106–577.

[17] K. Ishikawa, R. Tanaka, R. Ishikawa, A. Eriksen, Element type effects on stress range in computational models, J. Verif. Valid. Uncertain. 105 (2023) 400–861.

[18] E. Silva, A. Abbott, N. Wojcik, L. Quintero, K. Okonkwo, Friction effects on shear stress in computational models, Comput. Fluids 130 (2009) 30–502.

[19] K. Rasmussen, M. Bergman, D. Grigoryan, Damping effects on rib deflection in computational models, J. Biomech. Eng. 85 (2011) 3–447.

[20] C. Halvorsen, N. Halvorsen, N. Grigoryan, M. Kaur, R. Vasquez, Material model effects on chest compression in computational models, Comput. Methods Biomech. 129 (2005) 354–882.

[21] P. Yilmaz, M. Jankowski, K. Ishikawa, B. Vasquez, M. Vasquez, Load path effects on flow split in computational models, Ann. Biomed. Eng. 110 (2017) 158–619.

[22] R. Nakamura, K. Okonkwo, G. Tanaka, W. Bergman, Contact stiffness effects on stress range in computational models, Comput. Methods Biomech. 114 (2020) 271–444.

[23] J. Grigoryan, M. Vasquez, Element type effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 92 (2017) 36–442.

[24] J. Bergman, C. Moreau, Yield stress effects on rib deflection in computational models, J. Biomech. Eng. 46 (2012) 260–418.

[25] D. Grigoryan, E. Vasquez, L. Ueda, F. Grigoryan, Flow rate effects on chest compression in computational models, J. Verif. Valid. Uncertain. 107 (2023) 265–614.

[26] B. Eriksen, S. Lindqvist, Cement mantle effects on chest compression in computational models, Int. J. Numer. Methods 33 (2010) 15–482.

[27] B. Vasquez, F. Jankowski, Friction effects on stress range in computational models, J. Mech. Behav. Biomed. Mater. 115 (2018) 272–501.

[28] M. Okonkwo, T. Wojcik, G. Lindqvist, Bone density effects on flow split in computational models, Int. J. Numer. Methods 82 (2007) 271–426.

[29] S. Pettersson, A. Silva, R. Quintero, R. Eriksen, Solver tolerance effects on cortical strain in computational models, Comput. Methods Biomech. 70 (2005) 335–431.

[30] E. Nakamura, B. Pettersson, H. Eriksen, N. Lindqvist, Mesh density effects on chest compression in computational models, J. Verif. Valid. Uncertain. 93 (2021) 67–631.

[31] D. Eriksen, J. Fontaine, C. Rasmussen, Damping effects on stress range in computational models, J. Mech. Behav. Biomed. Mater. 104 (2006) 393–799.

[32] T. Kaur, D. Quintero, W. Wojcik, Solver tolerance effects on rib deflection in computational models, J. Mech. Behav. Biomed. Mater. 80 (2013) 349–764.

[33] N. Pettersson, C. Abbott, W. Zetterberg, N. Wojcik, D. Lindqvist, Material model effects on fatigue life in computational models, Comput. Methods Biomech. 130 (2017) 201–865.

[34] A. Fontaine, R. Tanaka, Friction effects on head displacement in computational models, Med. Eng. Phys. 42 (2006) 180–520.

Table 1: Credibility assessment summary table for HRK model across mechanisms (0–8 CES). Each row reports the factor scope, the score, and the one-line basis with numbers traceable to the prose.

Credibility factor	Level	Basis
Numerical code verification - HRK - normal drop impact	7.5	Energy balance within ±0.2%; shock tube MAE <1.2%; Taylor anvil length error 1.1%; hourglass energy <0.4%.
Discretization error - HRK - normal drop impact	7.0	Peak CoG linear acceleration 98.5 g to 97.0 g (1.5%) between fine and finest; Δt halving changes <0.8%.
Model form - HRK - normal drop impact	6.5	Force–time RMSE 7.8%; peak linear acceleration MAE 4.3 g across 5.9, 9.8, 14.7 J; minimal slip influence.
Relevance of the validation activities to the COU - HRK - normal drop impact	6.5	Energies 5.9–14.7 J and outputs (peak CoG g, force) match COU; identical CFC 1000 processing.
Software quality assurance - HRK - normal drop impact	7.0	112-case CI; tagged releases; no tolerance breaches; checksums and logs enable full reproduction.
Test samples - HRK - normal drop impact	6.0	n=48 tests across 3 energies and 3 sizes; standard instrumentation; limited aging/skin variability coverage.
Numerical code verification - HRK - oblique impact with tangential loading	7.5	Contact regressions at 7.5 m/s stable; penetrations <0.3 mm; hourglass energy <0.4%; benchmarks within 2%.
Discretization error - HRK - oblique impact with tangential loading	6.5	Peak angular accel 4,820 to 4,710 rad/s^2 (2.3%); peak angular vel 30.7 to 30.3 rad/s (1.3%); Δt effect 1.1%.
Model form - HRK - oblique impact with tangential loading	5.5	Early torque captured but late rotational rebound underpredicted; peak angular accel 8% vs test at 45°.
Relevance of the validation activities to the COU - HRK - oblique impact with tangential loading	7.0	Angles 30°, 45° at 4.0–7.5 m/s; outputs (peak α, ω) and processing match COU; video cross-checks used.
Software quality assurance - HRK - oblique impact with tangential loading	7.0	Traceable runs; friction-law defect corrected pre-freeze; regression tolerances held across oblique suite.
Test samples - HRK - oblique impact with tangential loading	6.0	n=36 tests across 2 angles and 3 speeds; IMU and video; abrasive surface controlled; limited environment sweep.

Table 2: Credibility assessment summary table for SBD model across mechanisms (0–8 CES). Each row reports the factor scope, the score, and the one-line basis with numbers traceable to the prose.

Credibility factor	Level	Basis
Numerical code verification - SBD - normal drop impact	7.5	Patch tests ≤0.35% error; shock tube MAE <1.2%; Taylor anvil within 2.0%; hourglass energy <0.4%.
Discretization error - SBD - normal drop impact	6.5	Skull g: 97.696.9 g (0.7%); 95th MPS: 0.1790.175 (2.2%); Δt halving changes MPS 1.6%.
Model form - SBD - normal drop impact	6.0	Peak linear g MAE ≤4.6 across energies; brain MPS modest (0.12–0.18); internal DOFs did not degrade skull fit.
Relevance of the validation activities to the COU - SBD - normal drop impact	6.5	Energies and outputs match COU; skull-level metrics emphasized; brain MPS included but not a decision driver.
Software quality assurance - SBD - normal drop impact	7.0	Shared CI with HRK; SBD-specific coupling tests; no regression failures during production.
Test samples - SBD - normal drop impact	5.5	n=48 skull-level tests; no PMHS brain instrumentation for vertical drops; tissue-level validation indirect.
Numerical code verification - SBD - oblique impact with tangential loading	7.5	Oblique regressions stable; penetrations <0.3 mm; conservation within 1%; ALE–FSI coupling robust.
Discretization error - SBD - oblique impact with tangential loading	6.0	95th MPS: 0.2240.214 (4.5%); skull α: 4,7804,690 rad/s^2 (1.9%); Δt halving changes MPS 2.9%.
Model form - SBD - oblique impact with tangential loading	5.0	Early torque captured; late-time rotational rebound low; brain bulk ω 36.1 vs 38.2 rad/s (5.5%).
Relevance of the validation activities to the COU - SBD - oblique impact with tangential loading	7.0	Angles/speeds match COU; PMHS data for brain bulk motion and MPS surrogates used for directional checks.
Software quality assurance - SBD - oblique impact with tangential loading	7.0	Tagged runs; containerized environments; CI tolerances held; traceable friction-law fix applied pre-freeze.
Test samples - SBD - oblique impact with tangential loading	5.5	n=36 oblique tests plus 5 PMHS cases; limited environment sweep; PMHS reconstructions carry uncertainty.