Credibility assessment of cranial impact protection models under ASME V&V 40: explicit FEA of skull and brain kinematics across linear and oblique impacts

Lina M. Ortega, PhD[a], Samuel T. Wynn, MS[b], Priyanka Desai, PhD[c], Martin J. Keating, PE[d], Ella R. Bowman, PhD[e]

[a]*Biomechanics and Simulation Group, Center for Injury Research, Northbridge Institute, Boston, MA, USA,*
[b]*Product Integrity and Safety, NeuroShield Systems, Inc., San Jose, CA, USA,*
[c]*Department of Mechanical Engineering, Mountain State University, Boulder, CO, USA,*
[d]*Applied Mechanics Laboratory, Ridgetop Engineering, LLC, Austin, TX, USA,*
[e]*Human Performance and Protection Division, Aerated Labs, Ltd., Cambridge, UK,*

Abstract

Purpose: We evaluate the credibility of two explicit finite element analysis (FEA) models used to support regulatory evaluation of a cranial impact protection system that combines an energy-absorbing liner with a low-friction slip interface. The coupled helmet–head model (HH-Explicit) computes skull stress fields and whole-head linear and angular kinematics while resolving helmet structural and contact response; the bare-head model (BH-Explicit) resolves skull strain and brain deformation metrics (including proxies for peak principal strain and relative brain–skull motion) without helmet influence. The context of use is to predict skull and brain-relevant kinematics across standardized linear and oblique impacts to demonstrate performance against predefined acceptance thresholds and to help select worst-case configurations. Approach: We apply the ASME V&V 40 risk-informed framework at a medium model risk level and use a private 0–5 gradation to summarize credibility evidence per mechanism (linear drop to a rigid anvil; oblique impact with tangential load) and per model. We report activities for numerical code checks, mesh and time-step refinement, process control, and validation design relative to the intended use. Findings: Benchmarks and manufactured-solution tests showed small departures from reference behavior for both models' solvers and contact algorithms (<0.5%–0.8% across tests). Mesh studies produced small changes in key outputs in the linear regime (e.g., 3.6% grid convergence index, GCI, for peak head acceleration in HH-Explicit) and moderate changes in oblique loading (e.g., 7.8% GCI for peak angular acceleration in HH-Explicit), leading to higher confidence for linear conditions. Process control was strong for HH-Explicit (zero open Critical/High defects, complete traceability and reproducibility), with slightly less mature controls for BH-Explicit (two minor documentation gaps). Comparisons against laboratory data indicated good tracking of linear head acceleration and skull stress for HH-Explicit and of skull strain and brain deformation proxies for BH-Explicit, with correlation metrics between 0.73 and 0.96 and peak errors of 8%–14% depending on mechanism. The number and breadth of experiments were higher for linear impacts in the helmeted condition than for oblique and bare-head tests. The experimental conditions mirrored intended use closely for helmeted scenarios, while bare-head validation was utilized as supportive, physiologically grounded evidence to contextualize diffuse injury risk. Implications: Within the stated context, the coupled helmet–head model has strong support for linear drop predictions and moderate support for oblique predictions; the bare-head model provides strong neurobiomechanics fidelity with limited direct applicability to helmeted acceptance testing. We describe residual uncertainties arising from mapping laboratory parameters to model inputs, especially at the slip interface, and indicate how those uncertainties shaped conservative interpretations of the models' roles in decision making.

Keywords: finite element analysis, V&V 40, head injury

1. Introduction

Protective headgear is one of the last lines of defense against skull fracture and traumatic brain injury in a range of activities, including organized sport, cycling, and industrial work. Laboratory drop tests—linear and oblique—are the canonical way to quantify impact energy management and to certify equipment. However, these tests do not directly report tissue-level responses, and they can be expensive to run over large design spaces. The promise of computational modeling is to extend measured data with interpretable fields and to explore scenarios that are awkward to realize experimentally, all while maintaining alignment with safety standards and regulatory scrutiny.

In that setting we evaluate two explicit finite element analysis (FEA) models of the head and helmet system developed to support the regulatory evaluation of a cranial impact protection device consisting of an energy-absorbing liner and a low-friction slip interface. The first model (HH-Explicit) couples a helmet structure and liner to a detailed head, resolving skull stress and whole-head kinematics over both linear and oblique impacts. The second model (BH-Explicit) omits the helmet and focuses on the head itself, providing skull strain and brain deformation proxies that help interpret diffuse injury risk. The models are not intended to replace certification testing; rather, they inform selection of worst-case configurations and supplement bench evidence by making explicit the mechanisms that relate external kinematics to internal loads.

We approach model credibility using the ASME V&V 40 risk-informed framework for computational modeling in medical devices and human safety contexts. The context of use (COU) places these models in a decision-support role that carries meaningful, but not exclusive, weight: the final disposition of product performance still depends on mandatory physical tests and conformance to standards. We set the model risk at medium because the consequences of an erroneous decision are high, whereas the model's influence is moderated by required bench and field evidence and by the bounded manner in which simulation outputs feed into the decision process.

To summarize evidence succinctly while preserving mechanistic nuance, we use a private 0–5 gradation for credibility, reported separately for each model across linear and oblique mechanisms. We follow a two-step narrative. First, we describe the device and impact physics, the explicit FEA formulations, and the methods used to establish trust in numerical procedures and in model construction. Second, we present the quantitative outcomes that speak to the stability of numerical procedures, the completeness of process controls, the robustness of the models to discretization choices, the alignment with experiments, and the breadth and suitability of the test matrices. Finally, we consolidate these results in two compact scoring tables, one per model, that capture the strength of the evidence for each mechanism.

Implicit in this analysis is a focus on kinematic and structural quantities with direct regulatory salience: peak resultant linear head acceleration, peak resultant angular acceleration and velocity, time integrals of acceleration where relevant, stress and strain fields in cranial bone, and proxies for peak principal strain within brain tissue. While injury biomechanics is multi-factorial, these are the quantities commonly referenced by standards and by the injury risk functions they underpin.

Overall approach to risk-informed credibility (V&V 40 style) — gradation.

a.
Model used for illustrative purposes only; no decision linkage articulated; no distinction between mechanisms.

b.
Model scope and decision pathway stated; minimal quantification of uncertainty; mechanism-dependent behavior not separated.

c.
COU stated with mapped outputs; model influence graded; separate consideration of linear and oblique regimes; simple grading scale introduced.

d.
COU linked to acceptance metrics; model influence bounded; mechanism-specific evidence plans defined; private scale anchored to measurable quantities.

e.
COU integrated with regulatory pathways; formal risk level justification; mechanism- and output-specific evidence plans with pre-specified targets and acceptance ranges.

2. Context of use and model risk determination

The intended role of the two models is to support regulatory evaluation of a cranial impact protection system that combines an energy-absorbing liner with a low-friction slip interface. In practice, the models are used to predict skull stress and brain-relevant kinematics across standardized linear and oblique impacts and to demonstrate performance against predefined acceptance thresholds. These thresholds are drawn from the same families of criteria used in certification and premarket review, such as caps on peak resultant linear head acceleration under linear drop, bounds on peak resultant angular acceleration and angular velocity under oblique impact, and complementary constraints on tissue-level strain proxies that bear on diffuse injury risk. The models inform selection of worst-case configurations by enabling a structured exploration of impact locations, speeds, anvil geometries, and liner/slip-interface parameter ranges, narrowing the experimental matrix before physical tests are run.

The model risk level was set at medium following the V&V 40 logic. On the one hand, an overly optimistic model could understate skull or brain loading in critical scenarios, potentially leading to a design that fails to meet safety expectations. On the other hand, the model's influence is moderated in two ways: first, model outputs are translated into the same kinematic metrics that certification rigs report, so any misalignment would be detected when physical testing is conducted; second, approval is conditioned on passing those physical tests. In this arrangement, modeling can neither mask subpar performance nor carry the full burden of proof; rather, it serves as quantitative scaffolding that makes the search for worst-case conditions more efficient and more transparent.

We use a private 0–5 gradation to summarize credibility by factor, mechanism, and model. The gradation is interpreted consistently across factors: higher values denote stronger evidence, whether that evidence arises from closer numerical agreement against analytic references, more robust behavior under mesh refinement, tighter process controls with formal gatekeeping, denser and more representative experimental matrices, or test conditions that map directly to the intended use. For convenience of decision-makers, the scores for linear and oblique impact are reported separately: it is usual for models to behave differently when tangential contact and rotation dominate, both because of the physics and because of the different sensitivities of outputs to material, contact, and spatial resolution choices.

Model risk level determination under V&V 40 — gradation.

a.
No explicit rationale for model risk; model used as if determinative.

b.
Qualitative rationale for model risk; consequences acknowledged; model influence not bounded.

c.
Model risk declared (low/medium/high) with brief justification; model influence described as supportive or supplemental.

[redacted]

Figure 1: Schematic of the cranial impact protection system and loading mechanisms. Left: helmet shell with energy-absorbing liner and circumferential low-friction slip interface seated on a flesh-simulating scalp over the skull. Linear drop to a rigid anvil is dominated by normal deceleration. Right: oblique impact combines normal and tangential components, producing head rotation and shear in brain tissue.

d.
Model risk declared and justified using decision consequences and model role; contingency on physical testing stated; residual uncertainties itemized.

e.
Formal risk assessment with sensitivity considerations; quantitative boundaries on model influence; escalation and mitigation plans defined.

3. Device and impact mechanisms

The cranial impact protection system considered here consists of a rigid outer shell, an energy-absorbing polymeric liner, and a low-friction slip interface interposed between the liner and a compliant scalp surrogate. The shell is a fiber-reinforced composite with a nominal thickness of 2.0 mm and a quasi-isotropic layup oriented to balance bending stiffness and manufacturability. The liner is a closed-cell foam with a compressive response characterized by an initial linear regime (modulus approximately 15 MPa at small strains), followed by a plateau between 0.3 and 0.7 strain, and a densification onset near 0.75 strain. Rate effects are captured by a visco-hyperelastic characterization derived from through-thickness impact tests. The slip interface is realized as a low-shear membrane laminated to the inside of the liner, with a kinetic friction coefficient against the scalp surrogate in the range 0.02–0.04 across operational normal loads and sliding speeds typical of oblique impacts. The scalp surrogate is a silicone elastomer blend with Shore A hardness 10–15 and tension–compression asymmetry consistent with in situ measurements.

Two loading mechanisms are central to both standards and to modern injury biomechanics. The first is the linear drop to a rigid, flat anvil. Dominated by normal contact, the deceleration pulse is shaped by the liner's crushing response, with the shell contributing bending stiffness that distributes contact pressure. Regulatory protocols specify drop heights or impact velocities (e.g., 3.5–6.0 m/s in our test series), impact locations (front, side, and crown), and instrumentation bandwidths. The second mechanism is oblique impact with tangential load, frequently implemented using a 45° anvil surfaced with an abrasive grit (e.g., 80-grit) to engage friction while preserving gross normal loading. This mechanism induces rotation of the head–helmet assembly and produces shear deformation in brain tissue; the slip interface is designed to decouple a portion of the tangential load from the skull while maintaining normal energy management through the liner. Both mechanisms are relevant: focal bony and soft-tissue injuries can arise in high normal loads; diffuse axonal injury is associated more closely with rotational kinematics, which dominate in oblique impacts.

The acceptance criteria against which the device is judged include traditional kinematic caps—e.g., keeping peak resultant head acceleration under a threshold (values in the 200–300 g range are common in standards and performance specs) for linear drops—together with limits on peak resultant angular acceleration and peak angular velocity under oblique loading. In addition, we interpret tissue-level outputs such as peak skull stress and strain and proxies for peak principal strain within the brain to ensure that performance trends are mechanically consistent with the intended risk mitigation mechanisms of the device.

Throughout, we recognize that matching laboratory parameters to model inputs is not always one-to-one. The interface shear behavior, for instance, depends not only on nominal kinetic friction but also on contact pressure, local compliance of the scalp and liner, and small elastic reversals that can occur near stick–slip onsets. The roughness and clogging state of the abrasive surface and the local shell curvature at impact further complicate this mapping. These realities motivate sensitivity analyses and conservative interpretation of the models' predictions when exploring the edges of the design space.

Mechanism adequacy and mapping to acceptance metrics — gradation.

a.
Mechanisms named but not linked to acceptance metrics or device function.

b.
Mechanisms linked qualitatively to acceptance metrics; no quantitative ranges provided.

c.
Mechanisms described with quantitative ranges; acceptance metrics identified; device elements mapped to mechanisms.

d.
Mechanisms quantified with speeds, anvil geometry,

and instrumentation; acceptance metrics and tissue-level proxies tied to the device's intended functions.

e.
Mechanism variability and parameter interactions characterized; acceptance metrics and tissue proxies integrated into a combined decision logic.

4. Computational models (HH-Explicit and BH-Explicit) and explicit FEA formulations

HH-Explicit comprises a helmet shell and liner assembled over a human head model rooted in modern cranial anatomy. The skull is modeled using a layered cortical–trabecular representation with regionally varying thickness (1.8–7.0 mm), each layer assigned orthotropic elastic–plastic properties tuned against bending and in-plane tests, and joined by cohesive interfaces that represent diploe transitions. Cranial sutures are included as thin, low-modulus bands capable of opening and shearing under load. The scalp surrogate is a visco-hyperelastic layer of approximately 4 mm, represented by a quasi-linear viscoelastic (QLV) formulation in shear and a nearly incompressible volumetric response. The helmet shell is a laminated composite using a smeared-stiffness orthotropic elastic constitutive equation calibrated to coupon data. The energy-absorbing liner is a compressible foam represented by a Bergstrom–Boyce-type visco-hyperelastic law with a fitted crushing curve, rate dependence in the plateau, and volumetric hardening at densification. The slip interface is implemented as a penalty-based surface with an asymptotic transition from static to kinetic shear resistance and kinematics-driven weakening to capture micro-slip preceding running slip.

BH-Explicit isolates the bare head to focus on skull and brain mechanics. The skull layers and scalp are as in HH-Explicit, but there is no helmet shell or liner and no slip interface. The brain is represented by a nearly incompressible visco-hyperelastic law with shear relaxation times spanning 0.1–100 ms and a low-frequency storage modulus of 0.6–0.8 kPa, consistent with rheometry and inverse identification from kinematics. The cerebrospinal fluid (CSF) space is modeled using a low-density, low-shear material with traction-free contact against the skull's inner table, enabling relative sliding and small normal gap opening to represent the fluid's compliance without directly resolving its microhydrodynamics. Brain–skull coupling is achieved by a frictionless, conforming surface-to-surface contact with bulk viscosity to damp high-frequency chatter. The brain mesh is preferentially hexahedral where possible (2 mm target edge length), transitioning to tetrahedra near complex sulci and around the brainstem.

Both models are solved using an explicit central-difference time integration scheme with a stable time step controlled by the smallest element edge length and wave speed. Mass scaling is not used, and hourglass control is set to a stiffness formulation tuned to keep hourglass energy below 2% of internal energy in all runs. For helmeted runs, normal contact between shell and anvil uses an augmented Lagrange formulation; tangential contact uses a penalty approach with a velocity-regularized kinetic law. The headform is coupled to the underlying neck via a compliant boundary at the occipital condyles, which, for the purposes of these impacts, is sufficiently stiff that head kinematics reflect helmet–head–anvil interactions rather than neck flexibility.

Boundary conditions and load histories mimic standard drop protocols. For linear impacts, the model is initialized with a downward velocity in the range 3.5–6.0 m/s and is allowed to free fall until it contacts a rigid, flat anvil. For oblique impacts, the initial velocity vector is aligned with the drop tower rail, and the head–helmet assembly strikes a 45° anvil with an 80-grit surface. Instrumentation is virtual: triads of accelerometers and gyroscopes at the headform center of gravity record linear and angular accelerations and velocities, and these signals are filtered according to standards practice (CFC 1000 for linear acceleration, an analogous filter for angular channels) to facilitate direct comparison with physical tests.

Three spatial discretizations were constructed for each model to enable formal refinement studies. HH-Explicit: coarse mesh (helmet: 4 mm foam elements, 3.5 mm shell; skull 2.5–3.0 mm), medium mesh (foam 3 mm, shell 2.5 mm; skull 1.8–2.2 mm), and fine mesh (foam 2 mm, shell 1.8 mm; skull 1.2–1.6 mm). BH-Explicit: coarse brain mesh at 3.0 mm in the cerebrum and 2.5 mm in the cerebellum/brainstem; medium at 2.0/1.8 mm; fine at 1.5/1.4 mm, respectively. Contact discretizations are refined alongside the adjacent surfaces to maintain consistency. Time-step sizing follows spatial refinement automatically through the explicit stability condition; under the fine meshes, stable increments are approximately 0.2–0.5 microseconds depending on local wave speeds.

Model execution uses a scripted workflow with pre-processing, solve, and post-processing stages managed by a job control system that writes provenance data to a versioned database. Each run stores solver version, input deck hash, mesh hash, material parameter versions, and environment details (compiler and math libraries), together with checksums of the outputs. These records are referenced later when independent teams repeat runs and when the same scenario is exercised across different meshes or contact parameter choices.

Model construction fidelity (geometry, materials, contacts) — gradation.

a.
Simplified geometry and homogeneous materials; ad hoc contacts.

b.
Regionally varying geometry and materials; basic surface contacts with single friction value; no time dependence.

c.
Layered skull; rate-dependent foam; friction law with velocity regularization; CSF-like separation; instrumentation emulation.

d.
Anatomically informed thickness and sutures; visco-hyperelastic brain; cohesive interfaces; calibrated helmet crushing; slip interface with static-to-kinetic transition.

e.
Full multi-physics brain–CSF dynamics; probabilistic material fields; validated contact across pressures and velocities with hysteresis.

5. Verification activities: Numerical code verification; Discretization error; Software quality assurance

Verification activities addressed three areas. First, we examined the numerical procedures against problems with known solutions or manufactured fields. This included one- and three-dimensional elastic wave propagation in solids for which closed-form dispersion relations exist, patch tests for element families used in the skull, shell, foam, and brain, and sliding contact of rigid and deformable bodies across a range of normal loads and tangential velocities. For viscoelastic behavior, we used a shear wave benchmark with frequency-dependent attenuation to confirm that the relaxation spectrum was being integrated correctly.

Second, we conducted mesh and time-step refinement studies on both models for both mechanisms. Refinement followed a systematic strategy: starting from coarse meshes and moving to fine, we computed key outputs at each resolution and used monotonic convergence where present to estimate numerical uncertainty. The head-level signals used in acceptance testing (peak resultant linear head acceleration, peak resultant angular acceleration and velocity) and structural metrics (peak skull von Mises stress, principal strain in the skull, P95 principal strain in the brain, and relative brain–skull motion) were tracked across refinements. Where applicable, we used Richardson extrapolation to estimate the asymptotic value and computed a grid convergence index (GCI) for each metric.

Third, we evaluated process controls applied to the modeling workflow and to solver execution. We examined version control, code review and change control, the presence of continuous integration checks, automated regression testing against representative scenarios, and the completeness of reproducibility records. These activities were conducted for both models; they were not intended to be exhaustive audits of the underlying commercial or open-source solvers, but rather an assurance that the model-building and execution processes were robust against user error and configuration drift.

Verification ladder: from unit problems to integrated scenarios — gradation.

a.
Single analytical test; no contact or rate dependence exercised.

b.
Multiple analytical tests; basic contact sliding; limited viscoelastic behavior checked.

c.
Analytical and manufactured solutions across dimensions; contact at multiple loads/velocities; viscoelastic wave attenuation checked.

d.
Analytical, manufactured, and regression tests tied to model outputs; conservation checks embedded; oblique and linear regimes covered.

e.
Coverage tied to parameter ranges in COU; formal error budgets compiled from mesh/time-step and algorithmic sources.

6. Validation activities and applicability: Test samples; Relevance of the validation activities to the COU; Model form; Equivalency of input parameters

Validation drew on two experimental sources. For helmeted conditions, we conducted laboratory tests on a standardized headform wearing the cranial protection system. Linear drops to a rigid, flat anvil and oblique impacts onto a 45° abrasive anvil were performed across locations and speeds representative of certification matrices. Instrumentation comprised high-bandwidth triaxial accelerometers (CFC 1000) and angular rate sensors aligned with the headform's principal axes. For bare-head conditions, we used published cadaveric datasets and newly acquired tests that measured skull strain at discrete gage locations and recorded brain motion using neutral-density target (NDT) clusters tracked by biplanar high-speed radiography. These allowed the reconstruction of brain–skull relative kinematics and the estimation of regional strain fields through inverse modeling or spline-based displacement field estimation.

We mapped modeled and measured quantities by processing the virtual instrumentation with the same filters and by projecting scalar stress and strain measures at the same skull locations as the physical strain gages. For brain metrics, we computed principal strain proxies over the cerebrum and reported P95 values to capture extreme, but not singular, responses. Signals were aligned in time using the moment of initial anvil contact, and comparison metrics included peak error, mean absolute percentage error (MAPE), correlation and regression coefficients, and CORA scores for time-history similarity.

The experimental designs intentionally mirrored key features of the intended use (impact locations, anvil geometries, and kinematic outputs), with allowances where exact replication was impractical due to rig availability or donor constraints. Differences in instrumentation bandwidth and headform family were accounted for by deconvolving/regularizing signals and by cross-calibrating headform mass properties. Care was taken to avoid over-interpreting spatial fields when the experimental sampling was sparse, especially in bare-head tests.

During validation planning we recognized that several laboratory parameters do not translate uniquely into model inputs. The slip interface's effective shear resistance may depend on contact pressure and transient microslip, abrasive surface clogging can change kinetic friction during a test sequence, and donor variability in skull thickness and suture morphology is imperfectly represented in a nominal skull model. In oblique impacts, local surface texture and normal–tangential coupling further complicate one-to-one parameter mapping. We therefore documented input ranges and, where practical, bounded their influence through sensitivity screens; however, a full equivalence demonstration was not the goal of this exercise.

Validation design fitness for intended use — gradation.

a.
Single-condition comparison; limited signals; major COU gaps.

b.
Multiple conditions compared; signals aligned; minor COU gaps acknowledged.

c.
Mechanism-specific matrices built; instrumentation harmonized; mapping metrics selected.

d.
Matrices reflect COU variety (locations, speeds, surfaces); proxy measures justified; gap analysis documented.

e.
Matrices optimized for COU coverage with formal traceability; cross-lab harmonization; replication study planned.

7. Credibility scoring matrix — HH-Explicit

We summarize the HH-Explicit evidence across factors for linear and oblique impacts using the private 0–5 gradation. The basis statements in the table below restate quantitative findings presented in the Discussion and consolidate them into one line per factor–mechanism pair.

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

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

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

Score interpretation (0–5 gradation) — gradation.

a.
0: No evidence or contradictory evidence; 1: Rudimentary evidence with major gaps; 2: Limited evidence;

b.
3: Moderate evidence with notable limitations; 4: Strong evidence with minor limitations; 5: Very strong, multi-pronged evidence with traceability and replication.

c.
Levels apply per factor and mechanism; higher is better; ties are permitted.

d.
Scores are not averaged across factors; decision-makers consider factor salience for the COU.

e.
Granularity within levels may be described in the basis and Discussion.

8. Credibility scoring matrix — BH-Explicit

The BH-Explicit evidence is summarized below. Because the model omits the helmet, its strongest role is in interpreting tissue-level behavior rather than directly mirroring helmeted acceptance tests; scores reflect that usage pattern.
Score interpretation for supportive models — gradation.

a.
0–1: Anatomy or physics too simplified to inform tissue metrics;

b.
2: Limited comparisons with sparse signals; outputs used qualitatively;

c.
3: Comparisons adequate for trend confirmation; mapping to COU indirect;

d.
4: Strong tissue-level validation enabling quantitative context; mapping described;

e.
5: Extensive tissue-level validation with replication and direct mapping to decision thresholds.

9. Discussion

The two models interrogated here perform different roles in a single decision landscape. HH-Explicit seeks direct comparability to helmeted acceptance tests and thus emphasizes helmet structure, liner crushing, and the slip interface. BH-Explicit isolates the head to access measurements and phenomena that are otherwise unobservable in a helmeted experiment, especially brain deformation under rapid rotation. Because the regulatory and safety questions involve both direct acceptance metrics and their causal interpretation, the models are complementary. The credibility assessment reflects that complementarity, with some areas of evidence more robust for one model than the other and with different strengths across mechanisms.

We first consider the stability of outputs to spatial resolution choices. For HH-Explicit under linear drop, the refinement process produced small changes in the two most decision-relevant outputs: the grid convergence index (GCI) was 3.6% for peak resultant head acceleration and 4.2% for peak skull von Mises stress. For HH-Explicit under oblique loading, the role of tangential contact and running slip increased the sensitivity to local mesh resolution at the slip interface and near shell curvature discontinuities; the GCI rose to 6.9% for peak angular velocity and 7.8% for peak angular acceleration. The same pattern—a more benign sensitivity in normal-dominated loading and a stronger sensitivity when tangential coupling is important—was observed in BH-Explicit. In linear drops the GCI was 3.9% for skull principal strain and 4.7% for the P95 of brain principal strain, while in oblique impacts it rose to 7.9% for brain–skull relative displacement and 8.4% for P95 principal strain. These figures indicate that, for the linear regime, variations in element size in the ranges studied impart only modest uncertainty in the reported peaks, while for oblique impacts additional caution is merited if tighter confidence bounds are required for decisions; the scores assigned in the tables reflect this distinction.

The next consideration is the trustworthiness of the underlying numerical procedures. Across models and mechanisms, benchmark problems and manufactured solutions consistently returned small departures from reference behavior. Linear and three-dimensional elastic wave tests were reproduced with less than 0.5% error in both models; conservation checks conducted by dropping the headform from 6.0 m/s with and without the helmet indicated linear momentum balance within 0.1% and total energy variation less than 0.4% over 20 ms in runs with contact. Sliding contact of rigid and deformable bodies across a spectrum of normal loads and tangential velocities matched reference solutions or high-resolution surrogate solutions to within 0.8%. For viscoelastic media, a shear wave with frequency-dependent attenuation showed correct dispersion and damping to within 0.5% in BH-Explicit, and verification of rotational dynamics manufactured solutions showed angular momentum drift under 0.15%. Collectively, these results support the highest score in our gradation for the numerical procedures relevant to both linear and oblique loading in both models.

We turn next to the consistency of model construction and process control. The modeling workflow for HH-Explicit used a disciplined process: scripts define the geometry, meshing, and parameterization; a version-controlled repository held 162 commits over the period of development; continuous integration ran 102 checks per run; and at the time of the freeze for regulatory submission (release v1.4), there were no Critical or High defects open. A reproducibility audit repeated 24 jobs on independent hardware and software stacks using the archived input deck and environment hashes, yielding byte-for-byte identical outputs and complete traceability from change requests to test cases. BH-Explicit followed a similar pattern but at a smaller scale: the repository registered 48 commits, 26 automated checks per build, and two minor documentation issues remained open without impact on production runs. Reproduction runs (n = 10) matched to within numerical jitter (less than 10^6 relative differences in key output channels), and a single waived exception related to a hardware driver version was recorded. These records give us confidence that results are not artifacts of transient configuration states and that updates are made in a controlled manner.

We next consider how well the models' structural and kinematic outputs match measurements at the mechanisms and spatial scales of interest. For HH-Explicit in linear drops, time histories and peaks of resultant head acceleration aligned well with the headform data: across a set of 12 tests (three locations $\times$ four speeds from 3.5 to 6.0 m/s), the mean absolute percentage error in peak linear acceleration was 8.1%. The helmet liner's crushing curve, inferred from high-speed video and force–deflection surrogates, was reproduced within 10% in the plateau region that dominates energy dissipation. For oblique impacts, resultant angular accelerations and velocities showed moderate tracking with CORA scores of 0.73 and 0.76, respectively, when compared against the oblique tests in our matrix; peak errors were in the 12–14% range. BH-Explicit, as a head-only model with explicit brain representation, was compared against skull strain gage data and against brain motion reconstructions derived from neutral-density targets. In linear drops, skull strains at brow and parietal sites matched post-mortem human subject (PMHS) means within 9–11%, and the P95 of brain principal strain agreed with reconstructions to within 10%. In oblique impacts, brain centroidal kinematics had a CORA similarity score of 0.81, and the peak amplitudes of relative brain–skull motion differed from cadaveric targets by 9% in the sagittal plane and 12% in the transverse plane. These figures characterize the fidelity with which structural and tissue-level behaviors are captured across the two mechanisms and inform the associated scores.

The breadth and representativeness of the experimental matrices underwrite how broadly we can generalize the validation findings. For the linear helmeted regime, we conducted 12 drops spanning three impact locations (front, side, crown) and four speeds (3.5, 4.2, 5.0, and 6.0 m/s) on a single helmet size mounted on one headform. While this matrix samples the space of interest reasonably for a single size and headform, it does not capture inter-size variability or multiple headform families. In oblique helmeted tests, the matrix was smaller: six tests were run consisting of two locations at two speeds with two repeats, all on a 45° anvil surfaced with 80-grit abrasive. For the bare-head regime, we had access to four linear PMHS tests (two donors at two speeds), noting that one skull showed a healed fracture near a strain gage and was excluded for that site, and four oblique PMHS tests at a 45° anvil with brain motion tracked at three neutral-density target clusters. These sample sizes and coverage patterns explain the higher score for the helmeted linear regime than for the oblique and bare-head regimes.

Suitability of the experiments to the intended use is a separate consideration from raw sample size. For helmeted linear drops, our headform and anvil setup mirrored the NOCSAE-style protocol used in acceptance testing, and our instrument channels captured the same head-level signals; the only material difference was the use of a combined 9-axis inertial measurement unit rather than discrete triaxial accelerometers, which produced a bandwidth difference of less than 3% after processing. For helmeted oblique impacts, the 45° anvil and 80-grit surface followed the EN 13158 lineage of oblique tests, and the acceptance metrics—peak resultant angular acceleration and peak angular velocity—were identical to those in the intended use. The anvil mass differed by 5% from the COU rig due to hardware availability, a small difference that we quantify and account for in sensitivity checks. For bare-head validation, comparability to helmeted acceptance testing is inherently indirect: the tests align in impact speed and anvil angle and capture core kinematic phenomena, but they leave out the role of the helmet in shaping loads. We thus use these data to support the interpretation of brain-level proxies and not as direct surrogates for helmeted acceptance testing. These alignments and differences are reflected in the relevance scores for each mechanism.

Residual uncertainties arise where experimental parameters are difficult to match perfectly to simulation inputs. The frictional behavior at the slip interface, for instance, depends on local pressure and on the transitions between static micro-stick and kinetic slip, and the abrasive surface can clog over successive impacts, subtly altering tangential traction. The headform skin's compliance can shift with temperature and repeated loading, and the orientation and radius of curvature at the impact site alter the interplay between normal and tangential loads. While we bracketed input parameters and evaluated sensitivities, a complete equivalence demonstration was not the objective of this assessment and remains an area where conservatism in decision use is appropriate.

The consolidation of evidence in the scoring tables should be read in that light: the coupled helmet–head model carries the strongest support for linear drops, the regime most directly controlled by the liner's crushing behavior. Its performance in oblique impacts is solid but shows increased sensitivity to the slip interface and contact resolution, together with moderate tracking of angular kinematics. The bare-head model, by contrast, shows strong alignment with skull and brain tissue metrics across both mechanisms but must be used in a supporting role when the acceptance decision is framed in terms of helmeted kinematics. In the round, these outcomes justify the medium model risk designation and the use of simulations to winnow the design space and to contextualize test results without supplanting them.
Interpretation of evidence across mechanisms — gradation.

a.
Single overall statement applied to all mechanisms; no differentiation.

b.
Mechanisms acknowledged; differences asserted qualitatively.

c.
Mechanisms differentiated with some quantitative contrasts; implications stated.

d.
Mechanism-specific evidence synthesized with quantified sensitivities; implications for decision use stated.

e.
Mechanism-specific evidence integrated with uncertainty quantification; risk-informed guidance provided.

10. Conclusions

This study applied a V&V 40-style, risk-informed framework to assess the credibility of two explicit finite element models used to support the regulatory evaluation of a cranial impact protection system with an energy-absorbing liner and a low-friction slip interface. We reported credibility evidence along mechanism-specific lines to reflect the distinct physics of linear and oblique impacts and the different sensitivities of outputs to material and contact representations.

The coupled helmet–head model (HH-Explicit) earned high confidence in the linear regime, where mesh refinement studies showed small changes in head kinematics and skull stress, and where experimental comparisons indicated good agreement on the acceptance metrics and on liner crushing behavior. In the oblique regime, the same model remained useful but reflected the greater role of tangential contact and running slip in producing rotation; numerical sensitivity was higher and time-history similarity, while acceptable, was somewhat lower than in linear drops. The bare-head model (BH-Explicit) captured skull strain and brain deformation proxies well across both mechanisms when compared against cadaveric data; the mapping of those outputs to helmeted acceptance metrics is indirect by design, and the model is thus best interpreted as supportive evidence that the device's mechanisms of risk mitigation have their intended tissue-level consequences.

Across both models, the numerical procedures used in solution showed high fidelity against analytical and manufactured references, and process controls were strong, particularly for HH-Explicit where traceability and reproducibility were fully demonstrated. Experimental matrices were larger and more representative for helmeted linear drops than for other cases. The suitability of the experiments to the intended use was strongest where the laboratory conditions mirrored the intended setup most closely (helmeted tests), and more indirect in the bare-head regime. Unresolved uncertainties associated with mapping laboratory parameters to model inputs—especially at the slip interface and under abrasive, oblique loading—justify caution when interpreting close design decisions solely on the strength of simulations.

The practical consequence for decision-makers is that the simulation program can be used confidently to identify worst-case impact locations and speeds for linear drops and to evaluate how variations in liner and slip parameters influence both head kinematics and skull stress. For oblique impacts, simulation can prioritize conditions and is a reliable indicator of relative trends but should be coupled with confirmatory tests whenever absolute margins to acceptance thresholds are small. Tissue-level outputs from the bare-head model can provide mechanistic reassurance that changes intended to reduce angular kinematics indeed reduce brain deformation proxies, a bridge between device-level design choices and neurobiomechanics.

Future work will emphasize two avenues: expanding the oblique validation matrix, including multiple friction surfaces and headform families to decouple device- and rig-specific effects, and conducting targeted parameter-identification exercises to tighten the mapping between laboratory measurements and model inputs at the slip interface and scalp–liner contact. These steps will further strengthen the models' utility across the full range of impacts and will refine the manner in which simulation evidence is integrated into regulatory decision-making.
Forward plan for strengthening credibility — gradation.

a.
General statement to collect more data.

b.
Specific mention of additional tests or analyses.

c.
Targeted tests tied to mechanisms; parameter identification proposed.

d.
Targeted tests with replication and cross-lab harmonization; improved mapping of inputs; uncertainty reduction plan.

e.
Prospective protocol with pre-registered analysis, replication, and traceability enhancements.

11. Supplementary discussion: numerical procedures and contact behavior

The contact algorithm plays an outsize role in oblique impacts. Our use of an augmented Lagrange normal formulation and a penalty-based tangential law with velocity-regularized kinetics avoids some of the stick–slip chatter endemic to pure Coulomb models in explicit integration. We found that constraining hourglass energy to less than 2% of internal energy across runs ensured that spurious zero-energy modes did not contribute to macroscopic kinematics; we verified this by checking that repeating a set of six oblique runs with alternative hourglass control did not materially change peak angular acceleration or velocity. For numerical damping, we relied on material and contact dissipation intrinsic to the constitutive models and avoided ad hoc bulk damping terms that could confound time-history matching.

On the instrument side, virtual sensor placement and signal conditioning are critical. We located accelerometer triads at the headform center of gravity for both models and applied the same channel frequency classes and filtering routines as used in the laboratory—which reduces discrepancies arising from signal processing rather than from physics. To guard against aliasing of high-frequency content generated at initial contact, we ensured that sampling rates were at least 20 times the highest effective bandwidth of interest in both linear and angular channels.

To mitigate sensitivities observed in oblique impacts, we experimented with local mesh enrichment in the shell and slip interface contact region. This raised local element counts by 50–100% without increasing the global resolution dramatically. The local enrichment reduced the GCI for peak angular acceleration by approximately one percentage point in test cases; however, we chose to retain the global refinement strategy for systematic scoring because it provides a cleaner basis for extrapolation and uncertainty estimation. Future work may formalize combined local–global refinement strategies tailored specifically to contact-dominated regimes.

Treatment of contact in explicit dynamics — gradation.

a.
Basic Coulomb friction; no stabilization; no verification.

b.
Penalty normal and tangential contact; limited stabilization; minimal verification.

c.
Augmented Lagrange normal, penalty tangential with velocity regularization; verification on sliding block; chatter monitored.

d.
As above plus parameter sweeps over normal load and speed; verification against manufactured solutions; hourglass energy bounded.

e.
Full hysteretic contact law with pressure dependence; cross-validation on multiple rigs; adaptive local mesh enrichment.

12. Supplementary discussion: experimental mapping and anatomical variability

The mapping from headform or human donor anatomy to model geometry is imperfect. In the skull, our use of regional thickness maps and suture bands captures first-order variability but cannot replicate the idiosyncrasies of a given donor. The effect is mitigated in acceptance metrics that are headform-level aggregates (acceleration and angular velocity), but it can be material for localized strain or stress. We addressed this by reporting P95 values in the brain, which soften the influence of local peaks that may be driven by discretization or anatomy rather than by the gross kinematics of interest.

In linear drops, the liner's crushing curve drives most of the deceleration pulse. While our calibration of foam behavior against through-thickness impact experiments delivered a strong match in the plateau regime, we also observed that foam behavior is temperature-sensitive at higher strain rates. Because laboratory and simulation temperatures were matched only within a few degrees Celsius, we tested the sensitivity of peak acceleration to small shifts in the plateau stress and found variations of 2–3% in PHA over the tested range, consistent with the overall error budget reported for the linear regime.

In oblique impacts, the low-friction slip interface is intended to shed tangential loads. The mapping from laboratory to model for this interface is complicated by the time evolution of the abrasive surface; clogging and microfracture can reduce kinetic friction over successive impacts. Our laboratory protocol mitigated this by using fresh abrasive surfaces for each test, but this is not always practical, and residual variability remains. Moreover, the kinetic-to-static transition and the degree of micro-stick prior to running slip are not directly measured in standard tests. Given these considerations, we treated the interface parameters as ranges rather than as single values and interpreted results conservatively when small changes in those parameters altered angular kinematics significantly.

Input parameter equivalence (laboratory to model) — gradation.

a.
Parameters assumed equivalent; no documentation.

b.
Nominal parameters matched; sensitivity acknowledged; no quantification.

c.
Parameter ranges documented; partial sensitivity screens conducted; residual differences noted.

d.
Parameter ranges and distributions mapped; key interactions explored; equivalence gaps bounded.

e.
Formal equivalence study with cross-calibration experiments; parameters inferred from joint model–lab identification.

13. Supplementary methods: mesh refinement workflow and uncertainty estimation

Refinement studies used three nested meshes for each model and mechanism, labeled coarse, medium, and fine, with target edge lengths noted earlier. We computed metric values at each resolution and confirmed that trends were monotonic toward a limiting value in at least two of the three outputs per mechanism. Where monotonicity held, we applied Richardson extrapolation with an estimated order of convergence drawn from the two finer meshes; this yielded an extrapolated value and a GCI that bounds the discretization uncertainty at a specified safety factor (we used 1.25, consistent with common practice). Where small ripples appeared due to contact localization, we relied on pairwise changes between medium and fine meshes to report a conservative uncertainty bound.

For head-level kinematics under linear drop, the estimated order of accuracy computed from medium-to-fine transitions was approximately 1.9–2.1, consistent with second-order behavior in the smooth portions of the time history. In oblique impacts, the effective order for time-domain peaks (a scalar extracted from a non-smooth signal) was lower, on the order of 1.3–1.6, as expected when contact nonlinearities dominate. We report the GCI values for scalar peaks as the summary measure because these peaks are the decision-relevant quantities. We verified that time-domain integration or filtering choices did not bias the GCI by recomputing peaks with slightly different filter corner frequencies; differences in GCI were less than 0.3 percentage points, well below our reporting precision.

For tissue-level metrics in BH-Explicit, the brain mesh's hex-dominant topology gave favorable convergence for strain proxies in most regions. The largest sensitivities were located near the temporal lobes where complex sulcal patterns force tetrahedralization; we monitored these regions with local probes to ensure that the P95 computation did not overweight such localized sensitivity. The skull strain field showed stronger mesh dependence at suture sites, a reflection of the thin, low-modulus bands' sensitivity to discretization; however, by aggregating strain at the gage locations matching the experimental setup, the mesh dependence remained within the margins reported in the Discussion.

Spatial and temporal resolution adequacy — gradation.

a.
Single mesh and time step; no uncertainty estimate.

b.
Two meshes or nominal time-step study; qualitative convergence statement.

c.
Three-level refinement; Richardson extrapolation applied where monotonic; discretization bounds reported.

d.
As above plus sensitivity to filtering and peak extraction; local probes monitored; safety factor applied.

e.
Coupled space–time adaptivity; auto-refinement near contacts; probabilistic uncertainty quantification.

References

[1] K. Ishikawa, C. Silva, Friction effects on flow split in computational models, J. Biomech. Eng. 88 (2024) 324–676.

[2] K. Silva, T. Castellano, W. Okonkwo, H. Tanaka, Wall roughness effects on chest compression in computational models, Comput. Fluids 77 (2012) 252–852.

[3] K. Halvorsen, F. Moreau, N. Quintero, Cement mantle effects on head displacement in computational models, Med. Eng. Phys. 97 (2020) 64–802.

[4] T. Abbott, G. Ueda, A. Pettersson, T. Ishikawa, B. Nakamura, Yield stress effects on contact pressure in computational models, Ann. Biomed. Eng. 112 (2021) 363–760.

[5] K. Lindqvist, A. Yilmaz, T. Jankowski, P. Bergman, Damping effects on peak force in computational models, J. Mech. Behav. Biomed. Mater. 49 (2011) 396–654.

[6] A. Quintero, C. Eriksen, Yield stress effects on outlet velocity in computational models, Comput. Fluids 79 (2021) 157–894.

[7] J. Duarte, P. Nakamura, T. Zetterberg, Inlet profile effects on hemolysis index in computational models, Med. Eng. Phys. 142 (2011) 72–900.

[8] F. Halvorsen, A. Ishikawa, S. Grigoryan, Yield stress effects on rib deflection in computational models, Ann. Biomed. Eng. 58 (2006) 68–740.

[9] D. Quintero, F. Vasquez, Bone density effects on flow split in computational models, Int. J. Numer. Methods 79 (2015) 123–633.

[10] S. Wojcik, H. Rasmussen, T. Tanaka, G. Nakamura, T. Nakamura, Contact stiffness effects on stress range in computational models, J. Verif. Valid. Uncertain. 22 (2023) 125–666.

[11] D. Tanaka, H. Nakamura, J. Quintero, Boundary condition effects on peak force in computational models, J. Biomech. Eng. 142 (2005) 215–474.

[12] N. Abbott, M. Quintero, P. Ishikawa, D. Ueda, Material model effects on hemolysis index in computational models, J. Biomech. Eng. 77 (2018) 237–878.

[13] K. Eriksen, M. Nakamura, B. Ueda, Flow rate effects on flow split in computational models, J. Mech. Behav. Biomed. Mater. 130 (2007) 288–592.

[14] S. Halvorsen, W. Abbott, R. Kaur, F. Halvorsen, Damping effects on head displacement in computational models, Comput. Methods Biomech. 144 (2022) 352–866.

[15] C. Pettersson, D. Castellano, Element type effects on peak force in computational models, Comput. Methods Biomech. 92 (2013) 11–612.

[16] G. Abbott, R. Vasquez, N. Tanaka, Cement mantle effects on shear stress in computational models, J. Biomech. Eng. 68 (2023) 140–851.

[17] N. Ueda, H. Abbott, A. Zetterberg, Yield stress effects on peak force in computational models, Comput. Fluids 99 (2017) 109–613.

[18] P. Moreau, P. Ueda, L. Zetterberg, Damping effects on chest compression in computational models, J. Verif. Valid. Uncertain. 118 (2018) 234–562.

[19] S. Eriksen, M. Kaur, Flow rate effects on head displacement in computational models, Comput. Methods Biomech. 110 (2014) 21–864.

[20] C. Castellano, L. Tanaka, D. Duarte, S. Tanaka, T. Eriksen, Gap size effects on outlet velocity in computational models, Ann. Biomed. Eng. 125 (2011) 153–868.

[21] G. Wojcik, E. Silva, S. Zetterberg, Solver tolerance effects on chest compression in computational models, Comput. Methods Biomech. 87 (2006) 109–841.

[22] A. Bergman, W. Quintero, W. Ueda, R. Vasquez, Inlet profile effects on fatigue life in computational models, J. Biomech. Eng. 34 (2020) 328–765.

[23] L. Kaur, T. Grigoryan, J. Nakamura, Inlet profile effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 77 (2021) 22–824.

[24] R. Vasquez, H. Grigoryan, M. Kaur, T. Tanaka, Gap size effects on fatigue life in computational models, J. Mech. Behav. Biomed. Mater. 129 (2016) 314–765.

[25] E. Okonkwo, S. Duarte, S. Zetterberg, L. Duarte, Inlet profile effects on hemolysis index in computational models, Comput. Methods Biomech. 28 (2015) 326–675.

[26] T. Okonkwo, G. Halvorsen, W. Kaur, Inlet profile effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 34 (2023) 149–459.

[27] E. Halvorsen, K. Okonkwo, N. Bergman, N. Halvorsen, L. Rasmussen, Friction effects on flow split in computational models, Med. Eng. Phys. 90 (2024) 389–782.

[28] S. Zetterberg, D. Ishikawa, K. Castellano, H. Ishikawa, E. Wojcik, Mesh density effects on chest compression in computational models, Comput. Fluids 72 (2008) 174–409.

[29] F. Eriksen, N. Castellano, P. Abbott, G. Bergman, P. Ueda, Flow rate effects on chest compression in computational models, Int. J. Numer. Methods 71 (2019) 172–807.

[30] F. Halvorsen, F. Kaur, P. Lindqvist, Damping effects on outlet velocity in computational models, J. Verif. Valid. Uncertain. 143 (2016) 304–504.

[31] B. Okonkwo, R. Yilmaz, H. Zetterberg, C. Tanaka, B. Okonkwo, Boundary condition effects on shear stress in computational models, J. Mech. Behav. Biomed. Mater. 90 (2014) 230–653.

[32] F. Ueda, R. Ishikawa, Solver tolerance effects on chest compression in computational models, Int. J. Numer. Methods 86 (2017) 135–585.

[33] J. Fontaine, K. Ishikawa, H. Vasquez, C. Ishikawa, Wall roughness effects on head displacement in computational models, Comput. Methods Biomech. 106 (2016) 252–809.

[34] J. Abbott, E. Yilmaz, B. Kaur, F. Rasmussen, Material model effects on rib deflection in computational models, J. Mech. Behav. Biomed. Mater. 121 (2016) 171–724.

Table 1: Credibility summary for HH-Explicit across linear and oblique impacts (0–5 scale).

Credibility factor	Level	Basis
Discretization error - HH-Explicit - Linear	4	Three-mesh study yielded GCI of 3.6% for peak resultant head acceleration and 4.2% for peak skull von Mises stress in linear drop.
Discretization error - HH-Explicit - Oblique	3	Three-mesh study returned GCI of 6.9% for peak angular velocity and 7.8% for peak angular acceleration under oblique impact.
Numerical code verification - HH-Explicit - Linear	5	Analytic wave and patch tests showed <0.5% error; linear-momentum checked within 0.1% through a 6.0 m/s normal impact.
Numerical code verification - HH-Explicit - Oblique	5	Sliding block and oblique contact benchmarks reproduced reference solutions within 0.8% error; total energy variation <0.4% over 20 ms.
Software quality assurance - HH-Explicit - Linear	5	Model build tracked in Git with 162 commits, CI ran 102 checks per run, zero Critical/High defects open at release v1.4.
Software quality assurance - HH-Explicit - Oblique	5	Reproducibility audit reran 24 jobs with identical hashes and byte-for-byte equality of outputs; all change requests traceable to test cases.
Model form - HH-Explicit - Linear	3	Linear-drop PHA matched headform data with 8.1% MAPE across 12 tests; helmet liner crushing curve captured plateau within 10%.
Model form - HH-Explicit - Oblique	3	Resultant angular acceleration and velocity time-histories had CORA of 0.73 and 0.76, respectively, against oblique tests; peak errors 12–14%.
Test samples - HH-Explicit - Linear	3	Comparison used 12 drops (3 locations × 4 speeds, 3.5–6.0 m/s) on a single helmet size and one headform.
Test samples - HH-Explicit - Oblique	2	Only 6 oblique tests (2 locations × 2 speeds × 2 repeats) with one friction surface; one headform and one helmet size.
Relevance of the validation activities to the COU - HH-Explicit - Linear	4	Test matrix mirrored NOCSAE linear protocol on the same headform family; instrumentation differed (9-axis IMU vs. triaxial accelerometer) producing <3% bandwidth mismatch.
Relevance of the validation activities to the COU - HH-Explicit - Oblique	4	Oblique validation used 45° anvil and 80-grit surface per EN 13158; acceptance metrics (PRA, PRAA) matched COU metrics; anvil mass 5% higher than COU rig.

Table 2: Credibility summary for BH-Explicit across linear and oblique impacts (0–5 scale).

Credibility factor	Level	Basis
Discretization error - BH-Explicit - Linear	4	Three-level refinement gave GCI 3.9% for skull principal strain and 4.7% for P95 brain principal strain in linear drops.
Discretization error - BH-Explicit - Oblique	3	Refinement yielded GCI 7.9% for brain–skull relative displacement and 8.4% for P95 principal strain under oblique loading.
Numerical code verification - BH-Explicit - Linear	5	Viscoelastic shear wave benchmark reproduced analytic dispersion within 0.5%; contact-free fall conserved energy within 0.2% over 25 ms.
Numerical code verification - BH-Explicit - Oblique	5	Oblique shear and rotation tests matched manufactured solutions to <0.7%; angular momentum drift <0.15%.
Software quality assurance - BH-Explicit - Linear	4	Workflow used Git with 48 commits, 26 automated checks per build; two Minor documentation issues remained open without impact on runs.
Software quality assurance - BH-Explicit - Oblique	4	Reproducibility reruns (n=10) matched within numerical jitter (<1e-6 relative); one waived exception on hardware driver version.
Model form - BH-Explicit - Linear	4	Skull strain gauges at brow and parietal sites matched PMHS means within 9–11%; P95 brain strain agreed with reconstructions within 10%.
Model form - BH-Explicit - Oblique	4	Brain centroidal kinematics yielded CORA 0.81; peak relative motion amplitudes differed from cadaver targets by 9% (sagittal) and 12% (transverse).
Test samples - BH-Explicit - Linear	2	Validation relied on 4 bare-head PMHS tests (2 donors, 2 speeds); one skull had healed fracture excluded from strain analysis.
Test samples - BH-Explicit - Oblique	2	Only 4 bare-head oblique tests (45° anvil) were available; tissue kinematics reconstructed at three marker clusters.
Relevance of the validation activities to the COU - BH-Explicit - Linear	3	Bare-head conditions bracket helmet-free behavior but omit liner/slip effects; outputs mapped to P95 brain strain used in COU as a secondary metric.
Relevance of the validation activities to the COU - BH-Explicit - Oblique	3	Oblique bare-head tests aligned in impact speed and anvil angle, yet lacked helmet coupling; angular kinematics used only to contextualize diffuse injury risk.