Credibility assessment of CFD models for an intravascular catheter: trackability and flow resistance under ASME V&V 40

Elena M. Duarte, PhD[a], Harish K. Rao, MS[b], Marta M. Brunner, PhD[c], Jonas P. Li, MS[1]

[a]*Center for Device Modeling and Simulation, MedTech Research Institute, Minneapolis, MN, USA,*
[b]*Department of Biomedical Engineering, Riverbend University, Riverbend, IL, USA,*
[c]*Computational Mechanics Group, Vascular Dynamics Ltd., Cambridge, UK,*

Abstract

Computational fluid dynamics (CFD) is increasingly applied to assess the hydraulic resistance and trackability of intravascular catheters. Under ASME V&V 40, credibility must be established in relation to the specific clinical decisions that the models inform and the risks carried by those decisions. We present a credibility assessment for two models of a steerable intravascular catheter intended for coronary and peripheral arterial navigation. Model A resolves steady, incompressible flow inside the catheter lumen with and without a guidewire to quantify pressure drop, wall shear stress, and flow partitioning. Model B solves transient, pulsatile external flow around the catheter advancing through a curved vessel to extract surface shear/pressure distributions, net lateral tip force, and unsteady drag. Four hydrodynamic mechanisms were targeted: M1, lumen pressure drop at clinical infusion rates; M2, outer-surface shear and drag during advancement; M3, lateral tip force from asymmetric outlet geometry; and M4, flow-induced distal shaft oscillations. The context of use spans design selection and labeling support for maximum deliverable flow rate, and design trade-offs for tracking forces in moderately tortuous paths. Credibility scores were assigned on a 0–7 index separately for each (model × mechanism) pair against six evidence areas: input specification, numerical solution quality, experimental agreement, pertinence of outputs, pertinence of the tests to the intended use, and physiological realism of boundary settings. Evidence included dimensional metrology of 20 catheters across three production lots; viscometry of contrast and blood analogs; mesh and time-step refinement leading to observed orders near 2.0 and grid-convergence indices below 1.2% for lumen pressure drop and below 3% for external forces; and benchtop measurements using a flow loop with pulsatile coronary waveforms, index-matched particle image velocimetry, and a six-axis microbalance to record forces. Model A attained high credibility for M1 (scores 5–7 across factors) with mean absolute error of 4.3% versus measured pressure drop across 1–15 mL/s and strong mesh convergence. It scored low for M2–M4, as internal CFD is not informative for external drag, lateral biasing, or oscillations. Model B achieved mid-to-high credibility for M2 (scores 4–6), with root-mean-square error of 7% for unsteady drag across three vessel diameters and three bend severities, and moderate credibility for M3 (errors of 12% MAE in lateral tip force) and M4 (frequency prediction within 6% but amplitude underpredicted by 27%). Physiological coverage included coronary diameters 2.5–4.0 mm, bend radii down to 20 mm, and Womersley numbers 2.8–5.1; extremes of tortuosity and highly calcified conditions were not emulated. Two credibility matrices—one per model—enumerate all 48 scored rows across mechanisms and factors. The integrated judgment supports use of Model A for claims and design choices tied to lumen hydraulics within the studied ranges, and use of Model B for comparative design screening of external drag and lateral biasing in moderate tortuosity, with additional evidence required for oscillation amplitudes and for severe anatomical scenarios.

Keywords: catheter CFD, ASME V&V 40, trackability

1. Introduction

Vascular access devices must balance hydraulic performance with navigational controllability in living vessels whose geometry and hemodynamics vary widely across patients. Catheter designers have traditionally relied on bench experiments to establish pressure–flow characteristics, trackability through benchtop tortuosity models, and compatibility with guidewires and ancillary tools. Computational fluid dynamics (CFD) now offers complementary insight, allowing designers to probe internal and external flows, explore local stresses on both device and vessel wall, and anticipate how geometry changes translate to performance shifts. However, to meaningfully support product decisions or labeling, these models require a structured assessment of their trustworthiness in relation to the intended decisions and the hazards of getting those decisions wrong.

We evaluate the credibility of two CFD models developed for a steerable intravascular catheter intended for coronary and peripheral arterial navigation. The first resolves the lumen hydraulics and predicts pressure drop under steady infusion and contrast delivery, with and without an in-lumen guidewire. The second resolves unsteady, pulsatile external flow during prescribed advancement of the catheter through a curved vessel segment and outputs time-varying drag, surface shear, and net lateral force at the distal tip associated with asymmetric outlet geometries such as bevels and side holes. Four fluid–structure mechanisms frame the inquiry: lumen resistance (M1), outer drag during motion (M2), lateral biasing from outlet asymmetry (M3), and flow-driven tip oscillations (M4).

The assessment follows the ASME V&V 40 framework, aligning the degree of evidence to the risk that model-informed decisions could lead to adverse outcomes. For lumen resistance, model-informed labeling of maximum deliverable flow is directly safety-relevant: overestimation could delay therapy or promote vascular injury from excessive syringe pressure. For navigational forces, the primary use is design selection and optimization, where errors are mitigated by subsequent integrated bench testing. We therefore expect to require more stringent support for M1 than for M2–M4.

To reduce ambiguity in interpreting the presented evidence, we score each (model × mechanism) pair across six evidence areas on a 0–7 Credibility Index. The index is private to this analysis but monotonically ordered: increasing scores indicate stronger support along the corresponding dimension. Rather than aggregating these scores, which can mask strengths and weaknesses, we present them in two model-specific matrices and interpret them mechanism by mechanism.

The parameterization of catheter, fluid, and vessel conditions is central to the practical utility of these models. The devices' dimensional features, fluid properties, and guidewire fits were pinned down through a targeted measurement campaign on 20 production catheters across three lots and two clinical guidewires, augmented by independent viscometry of contrast and blood analogs. The residual variability in the geometric and material descriptors entering the models was under 5% in all cases, with luminal diameter tolerance contributing the dominant share for M1 and outlet bevel and side-hole edging dominating M3. This anchoring of the parameter space stabilized predicted outputs across mechanisms and supported a consistent linkage between model runs and bench conditions.

Model inputs — gradation.

a.
Inputs set to catalog or nominal values without measurement or sourcing traceability.

b.
Key dimensions measured on a single lot or prototype; limited property measurements with basic instruments.

c.
Multiple lots measured with calibrated tools; fluid properties characterized across operating temperatures; inputs documented and version-controlled.

d.
Lot-to-lot sampling with traceable calibration and environmental control; independent metrology and viscometry; quantified uncertainties propagated into sensitivity analyses.

2. Device description and decisions supported

The subject device is a torqueable, steerable catheter with an overall working length of 110 cm, outer diameter 2.6–3.3 mm depending on size, and a nominal inner diameter (ID) of 1.40–1.50 mm. The distal 20–40 mm comprises a soft atraumatic tip with side holes in certain models and a beveled leading edge in others; the main shaft is a polymeric composite with embedded braiding, and the lumen may accommodate a 0.014 in (0.36 mm) or 0.018 in (0.46 mm) guidewire depending on size. The clinical use cases include selective engagement of coronary ostia, navigation through moderate tortuosity in proximal segments, and infusion of vasodilators or contrast agents into coronary or peripheral vessels. Flow rates in practice range from 1 to 15 mL/s for contrast boluses and 0.1 to 1 mL/s for medication infusion. Peak-to-mean ratios exceeding 1.8 occur in coronary waveforms, and diameters of interest range from 2.5 to 4.0 mm in coronaries and 4 to 8 mm in peripheral vessels.

Two categories of decisions rely on the CFD outputs. First, lumen hydraulics directly inform labeling for maximum deliverable flow rate under specified syringe pressures and guidewire presence. This sits near the clinical interface and demands tight error control and traceable agreement with benchtop measurements. Second, tracking performance is shaped in part by hydrodynamic loading as the shaft advances and turns, with particular interest in the forces required to traverse a bend, the tendency to drift laterally when outlet geometry is asymmetric, and any tendency to oscillate under pulsatile flow. These outputs are combined with frictional and structural considerations in design selection and do not, by themselves, directly support labeling about navigation forces.

The quantities computed by the models map onto the product-level metrics that the engineering organization uses to compare designs and, in the case of pressure–flow behavior, to substantiate claims. For example, steady-state pressure drop per unit length at specified volumetric flow is the same quantity read out during syringe-driven flow bench testing. The net lateral force calculated at the tip in a curved vessel relates to a bias score used by the design team to rank alternative outlet patterns; unsteady drag over one pulsatile cycle corresponds to the load cell trace recorded as the catheter is fed through the curvature in a flow loop; and the dominant frequency of tip displacement or force fluctuation ties directly to a repeatability index used in trackability ratings. By aligning the computed outputs to these decision metrics, we ensured that the simulation produces the signals that matter for the intended decisions, rather than surrogate quantities detached from engineering practice.

Relevance of the quantities of interest — gradation.

a.
Model outputs are surrogate values that do not appear in decision processes.

b.
Model outputs are qualitatively related to decisions but require interpretation or scaling to be used.

c.
Model outputs are the same signals used in engineering decisions and bench tests, with clear mapping.

d.
Model outputs are identical to decision metrics; thresholds and acceptance criteria are pre-defined and consistently applied across both simulation and experiment.

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

3. Context of use (ASME V&V 40) and rationale

The context of use (COU) for the internal-flow model is to support design selection and labeling for maximum deliverable flow rate through the catheter lumen when driven by a standard syringe or pump, with and without a guidewire, for contrast bolus delivery and medication infusion. The model addresses steady-state operating points in Newtonian fluids with viscosity matched to contrast and blood analogs measured at room and physiological temperatures, within the flow range 0.1–15 mL/s. The clinical decisions this supports center on infusion protocols and safety margins to avoid overpressurization or under-delivery of agents.

For the external-flow model, the COU is to support design selection and optimization for navigation in moderately tortuous environments, covering curved vessel segments with bend angles of 30°, 60°, and 90° and radii 20–40 mm, for coronary-sized vessels (2.5–4.0 mm diameter). The catheter advancement is prescribed to a constant feed rate typical of cautious navigation (2–5 mm/s) during pulsatile flow. The model produces cycle-resolved drag, surface shear distributions, and net lateral forces as a function of side-hole and bevel patterns. These outputs inform internal design ranking but do not directly drive labeling.

Within these COUs, the bench-relevant physiological operating space was explicitly selected and reproduced in both simulation models and experiments. Coronary-like waveforms with heart rates of 60–90 bpm, mean volumetric flows of 1–3 mL/s and peak-to-mean ratios near 1.8 were used; Womersley numbers for 2.5–4.0 mm vessels ranged from 2.8 to 5.1 when computed with an effective kinematic viscosity consistent with our fluid analogs. Vessel diameters of 2.5, 3.0, and 4.0 mm were included, and curvature radii of 20, 30, and 40 mm were tested. We did not emulate extremes such as severe ostial stenosis, highly eccentric calcified lesions, or anomalous branching geometries, and our conclusions should not be extrapolated to those conditions without additional evidence.

The balance between model detail, computational cost, and relevance was struck deliberately. For lumen flow, steady-state simulations were chosen because the pressure–flow relationship is not materially influenced by coronary pulsatility at the scales of interest and because the bench tests producing comparative data are steady. For external flow, unsteady simulations were necessary to capture load fluctuations correlated with the cardiac cycle and motion of the catheter. For both, geometric idealizations were introduced but minimized in dimensions known to strongly affect the outputs under study.

The operating scenarios simulated and tested encompass the middle 90% of physiologic combinations for the intended vessels by design, providing coverage that is neither excessively narrow nor so broad as to dilute relevance. This selection was grounded in a review of published coronary flow ranges and in-house intraprocedural recordings and was confirmed by a clinical advisory board. While not exhaustive, the range of diameters, curvatures, flow rates, and pulsatility indices used is sufficient for decisions about design selection and flow-performance labeling in typical coronary applications.

Test conditions — gradation.

a.
Boundary and operating settings are chosen for simulation convenience with little bearing on in vivo ranges.

b.
Some physiologic parameters are represented, but key ranges or combinations are absent.

c.
Operating points and ranges reflect typical in vivo conditions; extremes may be missing.

d.
A structured coverage of physiologic variability across key parameters with justification; boundary settings are harmonized between simulation and experiment.

4. Model risk level and rationale

Under ASME V&V 40, the level of evidence needed from a model depends on how the model influences decisions and on the potential clinical impact of those decisions. We classified the internal-flow model, when used to support labeling for maximum deliverable flow at specified pressures, as a moderate-to-high risk use. Errors here could produce clinical misjudgments about achievable infusion rates or induce excessive operator effort to attain a desired bolus, with downstream effects on efficacy and safety. That classification led us to demand tight uncertainty control, direct comparison to traceable measurements over the relevant flow range, and rigorous numerical verification for the internal-flow model.

For the external-flow model, the use case is design selection and optimization of catheter geometries, including side-hole patterns and bevel profiles, with outputs combined with frictional and structural tracking metrics in subsequent integrated tests. This was classified as a low-to-moderate risk use: while the model informs design choices, the consequences of incorrect rankings are mitigated by further bench evaluation before any clinical exposure. Consequently, we accepted somewhat looser acceptance thresholds for agreement and broader modeling assumptions (e.g., idealized vessels) provided the model captured the dominant physics for comparative purposes.

We invested in validation studies that mirror the decision environments described above. The external-flow experiments replicated coronary-scale curvature, flow pulses, and advancement rates to a degree consistent with practical benchtop constraints, though we did not force anatomical extremes. The internal-flow experiments duplicated the bench syringe-driven tests, matching geometric, fluid, and thermal conditions. By bringing the benches close to the decision context—mirroring diameters, curvatures, and waveforms—we ensured that the tests are not just generic but illuminate the very scenarios the engineering teams will rely on when applying the models in design review and labeling substantiation.

Relevance of the validation activities to the COU — gradation.

a.
Tests are conducted in generic conditions with no clear alignment to the decisions supported by the model.

b.
Some aspects of the decision environment are mirrored in the tests; others are missing or simplified.

c.
Key aspects of the decision environment (geometry, loads, operating ranges) are mirrored, with documented rationale for simplifications.

d.
Validation activities faithfully emulate the decision context across the principal variables; residual gaps are identified and bounded.

5. Mechanisms of interest and quantities of interest

Four mechanisms guide our analysis. M1, lumen pressure drop at clinical infusion rates, addresses the relationship between volumetric flow and pressure loss across the catheter, with and without a guidewire. This mechanism is driven by internal viscous losses and secondary flows near side holes; the quantities of interest are $\Delta p(Q)$, wall shear stresses in the lumen, and flow partitioning among distal outlets. M2, outer-surface shear and drag during advancement, concerns the hydrodynamic loads experienced by the catheter as it moves through a curved vessel under pulsatile conditions. Here the primary quantities are the axial drag over the cardiac cycle and the distribution of shear on the outer surface and adjacent vessel wall as the catheter feeds. M3, lateral tip force from asymmetric outlet geometry, reflects cross-stream loading induced by bevels and side holes that can bias the distal segment to push toward or away from the inner curvature during turns. The net lateral force integrated over the distal tip and adjacent shaft is the key quantity. M4, flow-induced distal shaft oscillations, captures unsteady coupling between pulsatile flow and the compliant distal segment, which can lead to oscillatory forces and motion that degrade controllability and elevate local stresses; the quantities include the dominant frequency of tip force oscillations and a measure of amplitude over a cycle.

For M1, assumptions of steady, incompressible, Newtonian flow are justified because Reynolds numbers in the lumen remain below 1500 for the studied rates and the test fluids show nearly constant viscosity at the shear rates encountered. Wall shear stress distributions inform material compatibility but serve mainly as internal checks for consistency with expected profiles (e.g., peak near constrictions). For M2–M4, pulsatility cannot be neglected: Womersley effects introduce phase shifts between pressure and velocity that govern instantaneous drag and lateral loading, and the prescribed advancement interacts with the local curvature to generate unsteady boundary layers and secondary Dean vortices. Despite this complexity, we isolate outputs that are both tractable to measure and stable with respect to minor model details.

Relevance of the quantities of interest — gradation.

a.
Selected outputs have an indirect or weak connection to the physical mechanisms of interest.

b.
Outputs track the mechanisms qualitatively but omit a key measurable signal.

c.
Outputs cover the primary measurable signals that best characterize the mechanisms.

d.
Outputs are matched to the most decision-critical and measurable signals, with definitions harmonized between models and experiments.

6. Computational model descriptions

Model A (Lumen Flow CFD) solves the steady incompressible Navier–Stokes equations in the catheter lumen geometry, including the distal outlets: end hole and, where applicable, side holes near the tip. The domain extends from the proximal hub to a plane 10 mm distal to the tip, with outlet boundary conditions set to zero gauge pressure. The guidewire, if present, is represented as a rigid, coaxial cylinder of the measured diameter and an offset eccentricity distribution consistent with observed tolerances; in the baseline case it is centered. No-slip boundary conditions are applied on all solid surfaces. The fluid is Newtonian, with viscosity set to either 1.0 mPa·s (saline contrast analog at 22 °C), 3.5 mPa·s (blood analog at 37 °C), or measured contrast-grade viscosity (2.1 ± 0.1 mPa·s at 22 °C) to match test conditions. The discretization uses a finite-volume scheme with second-order accuracy in space. The mesh employs boundary layer refinement with 10–14 prism layers to achieve $y+ < 1$ on internal walls, and a core unstructured tetrahedral fill. Mesh densities range from 0.8 million to 8.1 million cells across refinement levels. Convergence is established when residuals decrease by 6 orders of magnitude and Δp between successive iterations varies by less than 0.1%.

Model B (External Flow CFD with Imposed Advancement) solves the transient incompressible Navier–Stokes equations in an idealized curved vessel domain into which the catheter advances at a prescribed constant rate against a pulsatile background flow. The vessel is a toroidal segment of diameter Dv = 2.5, 3.0, or 4.0 mm with bend angle θ = 30°, 60°, or 90° and radius of curvature Rc = 20, 30, or 40 mm, smoothly connected to straight sections upstream (10 Dv) and downstream (15 Dv). The catheter is modeled as a rigid body with measured outer diameter and distal geometry (bevel and side holes), and its position is updated in time via a user-defined motion law at 2–5 mm/s. Pulsatile inflow is prescribed as a Womersley solution matching a measured coronary waveform (mean 1.8 mL/s, peak-to-mean 1.8, heart rate 60–90 bpm) with velocity profiles consistent with straight-tube solutions; outlet static pressure is set to achieve the desired mean flow. The fluid is Newtonian with viscosity 3.5 mPa·s at 37 °C. Turbulence is not expected for the Reynolds numbers encountered (Re < 1000), and we employ a direct unsteady laminar solution with second-order temporal accuracy (Crank–Nicolson-like scheme) and second-order spatial accuracy. Time steps of 0.1–0.5 ms were tested; baseline runs use Δt = 0.2 ms, with at least 10 cardiac cycles simulated to reach periodic steady state before gathering statistics. Meshes are hybrid unstructured with prismatic boundary layers ensuring y+ < 1 at the vessel and catheter walls, and total cell counts from 2.5 million to 22 million depending on Dv and Rc.

Both models were implemented in a single commercial CFD code, with custom scripts managing geometry import from measured CAD (distal outlet architecture from micro-CT scans), mesh generation parameters, and post-processing to extract the quantities of interest. For Model B, the vestibule entry and exit were modified to minimize reflection of pulsatile waves and spurious recirculation induced by the boundaries. The catheter motion is imposed via an arbitrary Lagrangian–Eulerian (ALE) mesh motion approach with periodic remeshing in a buffer region to maintain mesh quality; mesh skewness remained below 0.4 and cell aspect ratios below 8 in all runs.

Further refinement of the meshes and time steps in both models produced changes in primary outputs that were small relative to the acceptance thresholds defined for validation. In Model A, raising the mesh from 2.0 million to 8.1 million cells changed Δp at 10 mL/s by 0.6% and the mean wall shear stress by 1.1%. In Model B, halving Δt from 0.2 ms to 0.1 ms and increasing the mesh from 7.9 million to 15.8 million cells reduced the root-mean-square unsteady drag over a cycle by 2.1% and the peak lateral force by 2.6%. The power spectral density of the unsteady drag changed negligibly with further refinements: the integrated area over 0–20 Hz varied by 2.7% when halving Δt from 0.2 ms to 0.1 ms. These changes suggest that numerical resolution does not materially limit the precision of the predictions in the studied regimes.

Numerical solver error — gradation.

a.
No quantitative assessment of discretization error; single mesh and time step used.

b.
Mesh and time step varied coarsely; qualitative statement of stability and convergence.

c.
Systematic mesh and time-step refinement with observed order close to formal; changes in QoIs quantified and small.

d.
Full verification including grid-convergence index or equivalent; error estimates reported and propagated to decision metrics.

7. Credibility plan and scoring method (0–7 Credibility Index)

We defined a Credibility Index with levels 0 through 7, where a higher number indicates stronger support along a given evidence dimension for a specific (model × mechanism) pair. The six dimensions scored were: input specification, numerical solution quality, experimental agreement, pertinence of outputs, pertinence of tests to the decision context, and physiological realism of boundary settings. We did not aggregate scores across dimensions or mechanisms; instead, we present two model-specific matrices with 24 rows each (six factors times four mechanisms).

Our scoring references both internal process maturity and external evidence. For example, a high score in input specification requires not just dimensional measurements, but traceable calibration, coverage of multiple lots, and quantified uncertainties fed into sensitivity analyses that demonstrate output robustness. A high score in numerical solution quality requires observed orders close to formal, small changes in outputs upon mesh/time-step refinement, and reported uncertainty estimates (e.g., grid-convergence index). For experimental agreement, we required direct comparisons of the defined quantities at matched conditions with acceptance criteria pre-specified based on decision needs; the acceptance criteria are tighter for M1 than for M2–M4 consistent with risk classifications.

The acceptance criteria for M1 set a mean absolute error (MAE) threshold of 5% across 1–15 mL/s and a maximum deviation threshold of 10% for any given operating point in pressure–flow curves. For M2, the root-mean-square (RMS) difference between simulated and measured unsteady drag over a cardiac cycle must be ≤10% and the phase of the dominant harmonic within ±10°. For M3, mean lateral tip force must agree within 15% across tested outlet geometries and bend severities. For M4, the dominant frequency of tip force oscillations must match within ±10%, and amplitude within ±25% is considered acceptable for design comparison purposes though not for absolute prediction. These thresholds reflect the degree to which inaccuracies would change design rankings or labelable quantities.

Scores were assigned by two independent reviewers who examined the detailed modeling reports, raw metrology, and experimental data packages. Differences in preliminary scoring were reconciled by discussion, with tie-breakers resolved by a third reviewer. Where evidence was missing or weak, low scores were assigned rather than omitting the row. The resulting matrices therefore present a complete picture of evidence coverage and strength.

Output comparison — gradation.

a.
No experimental comparison; qualitative plausibility only.

b.
Limited comparison at a few conditions; acceptance criteria set post hoc.

c.
Direct comparison at multiple matched conditions with pre-defined acceptance criteria; errors quantified.

d.
Comprehensive comparison across the operating space with uncertainty bounds and traceability; results meet pre-set criteria with margin.

8. Model inputs: definition, sources, and sensitivity by mechanism

Dimensional metrology was conducted on 20 catheters sampled from three production lots using a combination of optical microscopy for distal features (side holes, bevel angle) and pin gages/micro-CT for lumen diameter and wall thickness. The inner diameter (ID) along the shaft was measured at 30 evenly spaced locations; the mean ID was 1.42 mm with a standard deviation of 0.03 mm across devices, and within-device variation had a standard deviation of 0.01 mm. The outer diameter (OD) at the distal 40 mm averaged 2.80 mm with a 0.04 mm standard deviation. For models with side holes, diameters were 0.25 ± 0.02 mm and arranged in a staggered pattern over the last 6 mm; the bevel angle at the tip was 25° ± 2° measured relative to the shaft axis. Guidewires of 0.014 in and 0.018 in nominal diameters were measured at 10 locations using a micrometer with NIST-traceable calibration; actual diameters were 0.356 ± 0.005 mm and 0.458 ± 0.005 mm respectively.

Fluid properties were measured on two fluids: a saline/contrast analog at 22 °C (1.00 ± 0.02 mPa·s across shear rates 10–1000 s1) and a blood analog at 37 °C formulated to 3.5 ± 0.1 mPa·s Newtonian viscosity over the same range, verified with a cone-and-plate rheometer. For contrast-grade fluids used in some pressure–flow testing, the viscosity was 2.1 ± 0.1 mPa·s at 22 °C. Density was measured by picnometry as 997 ± 1 kg/m3 at 22 °C and 993 ± 1 kg/m3 at 37 °C.

Sensitivity studies for M1 indicated that Δp at 10 mL/s varied approximately with μ/ID^4 as expected; varying viscosity by ±3% shifted Δp by ±3.0 ± 0.2%, whereas varying ID by ±0.03 mm shifted Δp by 8.1 ± 0.7%. Side-hole presence altered end distribution of flow by up to 12% of total for the last 6 mm but had negligible effect on total Δp for the lengths studied. The presence of a centered 0.014 in guidewire increased Δp by 21% at 10 mL/s, consistent with the reduced cross-section. For M3, lateral tip force changed by up to 18% when the bevel angle was varied within the measured tolerance (±2°), and by up to 9% when side-hole diameters were varied by ±0.02 mm. For M2, external drag was sensitive to OD (±0.04 mm produced ±5% changes in peak drag) but relatively insensitive to minor variations in distal tip length within ±2 mm.

Model inputs — gradation.

a.
Dimensions and properties assumed; sensitivity unexplored.

b.
Some measurements with basic tools; limited sensitivity exploration.

c.
Multiple calibrated measurements; targeted sensitivity analyses on key parameters.

d.
Comprehensive metrology across lots with uncertainty budgets; structured sensitivity analyses informing design tolerances.

9. Numerical solver error: spatial/temporal discretization studies

Spatial convergence for Model A was assessed by constructing three systematically refined meshes (0.8, 2.0, and 8.1 million cells) using the same meshing algorithm with a refinement factor r 1.6 in characteristic cell size. The formal order of the spatial scheme is two. The observed order of accuracy for Δp at 10 mL/s was 1.98, with a grid-convergence index (GCI) of 1.2% between the fine and medium meshes using a 95% confidence factor. For mean wall shear stress over the distal 20 mm, the observed order was 1.87 and GCI 1.6%. Similar behavior was observed at other flow rates (1–15 mL/s).

Temporal and spatial convergence for Model B were evaluated by refining both the time step and the mesh. The baseline mesh of 7.9 million cells and Δt = 0.2 ms yielded stable solutions with three cycles to periodic steady state (checked by overlaying the last two cycle traces of drag and lateral force). Halving the time step to 0.1 ms changed the cycle-averaged drag by 1.8%, the RMS drag over a cycle by 2.1%, and the peak lateral force by 2.6%. Doubling the mesh to 15.8 million cells (with Δt = 0.2 ms) changed those quantities by 2.4%, 2.7%, and 2.3% respectively. The power spectral density (PSD) of unsteady drag exhibited a primary peak at 1.1–1.5 Hz depending on heart rate, and that peak's frequency shifted by less than 0.05 Hz on refinement; the integrated PSD over 0–20 Hz shifted by 2.7% with halved Δt. Extrapolating with Richardson's method for selected points produced changes under 3% in both drag and lateral force metrics. Cells adjacent to walls maintained y+ below 1 in all meshes, with prism-layer coverage of at least 10 cells across the boundary layer thickness.

Numerical solver error — gradation.

a.
Verification not attempted; discretization unspecified.

b.
Coarse refinement attempted; changes in outputs reported without error estimation.

c.
Observed orders near formal; GCI or equivalent reported for primary outputs.

d.
Verification with demonstrated asymptotic range and error bars propagated to decision metrics.

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

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

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

10. Output comparison: experimental benchmarks and acceptance criteria

Experimental benchmarks were developed to mirror the mechanisms and quantities targeted by the models. For M1, a steady flow loop with deionized water, blood analog at 37 °C, and contrast analog at 22 °C drove fluid through the catheter lumen at controlled volumetric rates from 0.1 to 15 mL/s using a calibrated syringe pump. Differential pressure was recorded between proximal and distal taps 20 cm apart using a differential pressure transducer with 0.25% full-scale accuracy, yielding a resolution better than 0.2% of the measured Δp at typical flows. Tests were conducted for bare lumen and with 0.014 in and 0.018 in guidewires in place, centered. For each configuration, five repeats were performed. The acceptance criterion was a mean absolute error ≤ 5% across the flow range with no point exceeding 10% error.

For M2–M4, an index-matched flow loop with a curved glass vessel section (D = 2.5, 3.0, 4.0 mm; Rc = 20, 30, 40 mm; bend angles 30°, 60°, 90°) was used. A programmable pump generated coronary-like pulsatile flow (mean 1–3 mL/s, 60–90 bpm), and the catheter was advanced at 2, 3.5, and 5 mm/s using a linear stage synchronized with the flow controller. A six-axis microbalance attached to the catheter hub recorded axial and lateral forces at 1 kHz. Particle image velocimetry (PIV) captured in-plane velocity fields at selected cross-sections in a midbend plane to assess shear and flow structures; seeding particles were 10 μm hollow glass spheres, and images were acquired at 500 Hz with spatial resolution ~25 μm. The acceptance criterion for M2 was RMS unsteady drag error ≤ 10% and dominant harmonic phase error ≤ 10°. For M3, mean lateral tip force error ≤ 15% across outlet geometries and bend severities. For M4, dominant frequency error ≤ 10% and amplitude error ≤ 25%.

Model A predictions of Δp(Q) matched the bench data within a mean absolute error of 4.3% across the 1–15 mL/s range for bare lumen, 4.9% with a 0.014 in guidewire, and 6.2% with a 0.018 in guidewire. Maximum pointwise deviation was 8.6% at 12 mL/s with the 0.018 in guidewire in contrast analog at 22 °C. Model-to-measurement agreement improved as viscosity increased, consistent with Reynolds-number scaling and reduced sensitivity to minor geometric roughness. Wall shear stress distributions at the distal outlets had no direct experimental comparator but showed patterns consistent with known internal-flow behavior near sudden expansions and side holes.

For Model B, the RMS unsteady drag error across all vessel diameters and bend severities was 7.0% (N = 27 condition combinations), with dominant harmonic phase errors under 6°. Mean lateral tip force across three outlet geometries (end hole only, end hole plus two side holes, and beveled tip with two side holes) and three bend severities (30°, 60°, 90°) had a mean absolute error of 12%, with the largest errors (up to 18%) occurring in the most severe bend (90°, Rc = 20 mm) at 5 mm/s advancement. The power spectral density of lateral force displayed a dominant frequency near the heart rate and its first harmonic; the dominant frequency was predicted within 6% across conditions, while amplitude at that frequency was underpredicted by 27% on average and up to 34% at the highest advancement rate. PIV velocity fields agreed qualitatively in secondary flow structures (Dean vortices) and quantitatively within 10–15% of centerline velocity profiles, though the models lacked the fine-scale waviness induced by slight glass surface imperfections observable in experiment.

Output comparison — gradation.

a.
Single-condition comparison with qualitative agreement only.

b.
Multiple conditions with quantitative comparison; acceptance thresholds not pre-specified.

c.
Pre-specified acceptance thresholds met across the bulk of the operating space; uncertainties reported for instruments.

d.
Pre-specified thresholds met with margin across operating space; measurement and model uncertainties combined to support decisions.

11. Relevance of the quantities of interest to trackability and resistance

The outputs for M1—pressure drop as a function of flow rate with and without a guidewire—feed directly into flow-performance charts and labeling. For M2, the time trace of axial drag during prescribed advancement in pulsatile conditions is the signal that determines the hydrodynamic portion of the force budget engineers use to evaluate trackability; it is combined with frictional and structural contributions measured separately. For M3, the net lateral force at the tip is used to compute a bias index that penalizes designs exhibiting a sustained sideways push in bends, with thresholds set to ensure controllability. For M4, the dominant frequency and amplitude of tip force oscillations are used to compute a repeatability index

that is factored into the trackability scorecard. These outputs are defined identically in model post-processing and in bench data analysis, making one-to-one comparisons and transfer to decision-making straightforward.
Relevance of the quantities of interest — gradation.

a.
Outputs are defined differently across model and bench, complicating mapping to decisions.

b.
Outputs share definitions but require conversion to align with decisions.

c.
Outputs are identically defined and directly comparable to the decision metrics.

d.
Outputs are identically defined with shared analysis pipelines across simulation and experiment.

12. Relevance of the validation activities to the COU

The lumen-flow validation repeated the same syringe-driven conditions used in product verification and labeling substantiation: identical fluid types, temperatures, and volumetric-rate sweeps. The external-flow validation mirrored the coronary-scale diameters, pulsatile waveforms, and advancement speeds used in internal design comparisons and in the downstream integrated trackability benches. The curvature severities selected align with the device's intended use: moderate tortuosity, not extreme or pathological cases. While vessel walls were rigid in the external-flow experiments and in the model, this simplification aligns with the intent to isolate hydrodynamic loads as a design variable rather than to predict wall stress in elastically deforming vessels, which is outside the COU. We therefore consider the tests to be strongly aligned with the decisions these models inform, with deliberate simplifications documented and confined to aspects outside the direct decision metrics.
Relevance of the validation activities to the COU — gradation.

a.
Tests remote from the intended decision environment; transfer requires major assumptions.

b.
Tests share some features with the decision environment; key conditions are missing.

c.
Tests match the main features of the decision environment; remaining gaps are justified.

d.
Tests replicate the decision environment across the key variables that drive the decision metrics.

13. Test conditions: boundary conditions and physiological ranges

For the internal-flow model and its companion experiments, boundary conditions were volumetric-flow inlets set by syringe pumps and outlets held at atmospheric pressure. Fluid temperatures were controlled to 22 ± 1 °C for saline/contrast analog tests and 37 ± 0.5 °C for blood analog tests. Flow rates covered 0.1–15 mL/s. The geometry spanned the entire catheter length for pressure-drop measurement but focused numerically on the hydraulic resistance segment between taps to align with the experimental Δp. Guidewire presence was modeled and tested at centered position for baseline comparison; sensitivity to eccentricity was explored numerically but not verified experimentally due to positioning limitations. The minor losses associated with connectors and luer fittings were measured and subtracted from bench data to isolate catheter-only Δp.

For the external-flow model and experiments, inflow boundary conditions were specified using coronary-like pulsatile waveforms scaled to three mean flow levels (1.0, 1.8, and 3.0 mL/s) and three heart rates (60, 75, and 90 bpm), using Womersley profiles for straight segments. Outflow pressure was adjusted to achieve the desired mean flow. Vessel walls were rigid with no-slip conditions; catheter outer surface was no-slip. The catheter moved at fixed rates of 2.0, 3.5, or 5.0 mm/s through the curved segment. Bursts of microbubbles were avoided to preserve PIV quality. All settings were mirrored between simulation and bench to the extent practical, including synchronized acquisition of force and flow data for phase alignment.
Test conditions — gradation.

a.
Boundary settings are arbitrary or undocumented; physiologic ranges not considered.

b.
Some physiologic settings used; limited range; boundary mismatches between model and bench.

c.
Physiologic ranges covered with harmonized boundary settings between model and bench.

d.
Structured coverage of physiologic combinations; rigorous boundary matching and documentation.

14. Use error: operator and setup considerations

Trackability experiments and simulations can be influenced by how operators handle devices and how setups are assembled. In bench tests, catheter feed rates, alignment at entry, and preconditioning protocols can alter frictional baselines or introduce geometric bias. In simulation, the way distal geometry is reconstructed from scans, or how motion profiles are imposed, can likewise nudge results. We noted in passing that clinical operators vary in the cadence of advancement and in choice of guidewire support, and these behaviors may not be fully captured by constant-rate advancement assumptions. While these topics frame the context in which models and experiments are applied, they sit outside the present scoring and are considered qualitatively in design reviews.
Model inputs — gradation.

a.
Operator- and setup-dependent parameters are not identified.

b.
Parameters are listed but not quantified.

c.
Parameters are listed with typical ranges from expert input; limited quantification.

d.
Parameters are quantified or bounded by measurement, and their influence is explored in sensitivity analyses.

15. Factor-by-factor score matrices for Model A across M1–M4

The following matrix presents the 24 scored entries for Model A (Lumen Flow CFD) across the four mechanisms. Scores are on the 0–7 Credibility Index, where higher values denote stronger evidence. Each row includes a concise basis statement aligned with evidence detailed in the body, with specific numerical values matching those reported in the sections on inputs, numerical studies, and comparisons.

16. Factor-by-factor score matrices for Model B across M1–M4

The following matrix presents 24 scored entries for Model B (External Flow CFD with Imposed Advancement) across the four mechanisms. Scores and basis statements mirror the evidence discussed in preceding sections. Model B was developed specifically for M2–M4 and was not used to evaluate lumen pressure drop (M1).

17. Integrated credibility judgment by mechanism and model

The mechanism-by-mechanism synthesis clarifies where each model carries decision-grade weight. For lumen resistance (M1), Model A demonstrated close agreement with benchtop pressure–flow measurements across 1–15 mL/s in all tested fluids and guidewire configurations, with mean absolute errors below 5% and a maximum deviation of 8.6%. Numerical verification produced observed orders near the formal second order and grid-convergence indices below 1.6% for both Δp and wall shear. Together with a tightly controlled input set—traceable metrology on 20 catheters and fluid viscometry—this yields a coherent picture: the internal-flow model can support labeling statements and design trade-offs in the studied range.

For hydrodynamic loading during advancement (M2), Model B produced RMS unsteady drag within 7% of measured values across a broad set of vessel diameters (2.5–4.0 mm), bend angles (30–90°), curvature radii (20–40 mm), flow rates (1–3 mL/s mean), heart rates (60–90 bpm), and advancement speeds (2–5 mm/s). Phase agreement of the dominant harmonic was better than 6°. Numerical convergence checks indicated that changes in RMS drag and peak lateral force were ≤ 2.7% under halved time steps and doubled meshes, suggesting numerical errors are below acceptance thresholds. The tested conditions mirror the intended decision context—moderate tortuosity under coronary-like pulses. In consequence, Model B earns mid-to-high scores for M2, adequate to support comparative design screening.

For lateral biasing due to outlet geometry (M3), Model B captured mean lateral tip forces with a mean absolute error of 12% across three outlet configurations and three bend severities, meeting the pre-set 15% threshold but with degradation at the most severe bend and fastest advancement, where errors reached 18%. Input control was strong for the geometric features that most affect M3 (bevel angle and side-hole diameter), and numerical sensitivities were small. The conditions exercised address the design environment for the intended use but stop short of extreme anatomies or vessel compliance effects. We consider the external-flow model serviceable for ranking designs on lateral bias in moderate tortuosity, with caution for the severe combinations where errors grew.

For flow-induced oscillations (M4), Model B reproduced dominant frequencies within 6% across conditions but under-predicted amplitude by 27% on average. The absence of structural compliance in the model likely limits amplitude fidelity even though hydrodynamic time resolution and mesh quality were sufficient, as indicated by PSD stability with refinement. Experiments used rigid glass vessels by design to isolate hydrodynamic contributions, mirroring the model. Given these results, the model is appropriate for identifying frequency regimes of potential concern and for relative ranking based on frequency content, but not for making absolute amplitude judgments without additional modeling or evidence.

Across mechanisms, the patterns underscore the principle of mechanism-specific credibility: evidence that suffices for lumen hydraulics (M1) does not transfer to external drag (M2–M4), and vice versa. Our matrices reflect this separation explicitly. Notably, the computed quantities aligned cleanly with the decision metrics used in design reviews and labeling. Against bench measures, differences remained within the pre-stated thresholds for M1–M3, while M4 amplitude did not meet the 25% amplitude acceptance criterion. The numerical resolution had a small influence on results: further refinements altered M1 Δp by less than 0.8% and changed M2/M3 loads by less than 2.6%, and halving the time step changed the force spectra area by 2.7%. These observations support the central conclusion: the internal-flow model supports label claims on flow resistance in the studied regimes, and the external-flow model supports design ranking for drag and lateral biasing in moderate tortuosity; for oscillation amplitudes and for severe tortuosity or highly non-ideal geometries, additional validation and model enhancement would be needed.

Output comparison — gradation.

a.
Agreement discussed qualitatively without thresholds.

b.
Thresholds applied after observing data; partial coverage.

c.
Pre-set thresholds met for main mechanisms; shortfalls

identified and bounded.

d.
Pre-set thresholds met with margin for all mechanisms in scope; implications for decisions quantified.

18. Discussion

The present assessment demonstrates how to apply ASME V&V 40 to a pair of CFD models that target different but complementary aspects of catheter performance. By structuring evidence around mechanisms and by resisting the temptation to lump across them, the analysis preserves transparency: Model A is fit for purpose on lumen resistance; Model B is fit for purpose on hydrodynamic components of trackability in moderate coronary-like geometries. The explicit mapping of quantities of interest to decision metrics—Δp(Q) for labeling, time-resolved drag and lateral force for design ranking, dominant oscillation frequency for early warning—keeps the focus on signals that matter to engineering choices.

Several modeling choices warrant comment. First, both models assume Newtonian fluids. For the flow rates and vessel sizes under consideration, non-Newtonian effects in blood analogs are weak; our rheometry confirmed near-constant viscosity across relevant shear rates, and the difference between Newtonian and simple shear-thinning models would be small relative to measurement error in our benches for these ranges. Second, Model B adopts rigid walls and prescribes catheter advancement at constant speed. These choices isolate the hydrodynamic loads from structural and frictional couplings, consistent with the intent to assess one component of trackability in a controlled way. The down side is that amplitude of oscillations (M4) is not reliably predicted without a compliant-structure model; the benefit is an interpretable mapping from geometry changes to hydrodynamic loading, which is useful in design iteration.

The numerical verification results show that both models operate in the asymptotic regime of their discretizations for the quantities of interest. Observed orders near two and small changes upon refinement indicate low discretization error burdens. We found that time-step and mesh changes in Model B had limited effect on RMS drag, peak lateral force, and spectral content, with shifts under 3% for most metrics. In Model A, Δp was nearly invariant with further mesh refinement (<1%), reflecting adequate resolution of the boundary layers and mild geometric complexity in the lumen.

On experimental comparison, Model A's close match to pressure–flow curves likely benefits from both the problem's relative simplicity and the rigor of input measurement. Small discrepancies at high flows with the larger guidewire (0.018 in) are plausibly explained by minor eccentricities not captured in the centered-wire baseline and by sensitivity to ID tolerances compounded by the high Reynolds number end of the tested range. An eccentricity sweep in simulation suggests that off-center wire positioning by 0.1 mm would increase Δp by an additional ~3% at 12 mL/s, which brackets the observed discrepancies. For Model B, the better agreement in M2 than in M3 and M4 follows intuition: axial drag is primarily a function of projected area and wake dynamics that are comparatively insensitive to minor geometric details, whereas lateral force and oscillation amplitude depend on the interplay of asymmetric outlet jets, curvature-induced secondary flows, and structural dynamics not included in the fluid-only model.

Limitations include the use of idealized vessel geometries for external-flow modeling and testing; while diameter and curvature ranges were representative, real vessels feature skews, branches, and wall compliance. Extending the validation to vessel phantoms that incorporate elastic walls or to ex vivo vessels could bridge this gap, though introducing additional uncertainties in material properties and repeatability. For oscillation amplitudes, coupling the CFD model to a reduced-order structural model of the distal shaft (e.g., Euler–Bernoulli beam with measured stiffness and damping) could improve fidelity for M4. For M3, further attention to the precise reconstruction of bevel edges and side-hole chamfers may tighten agreement, as sensitivity studies indicated notable influence of ±2° bevel changes on lateral force.

We also note process learnings. The alignment of analysis pipelines between bench and model—consistent definitions of RMS drag over a cycle, dominant frequency extracted from PSD, and averaging windows—reduced interpretive loss and made acceptance checks mechanical. The pre-definition of acceptance criteria before running comparisons prevented subconscious drift of thresholds and clarified when targeted mechanisms were adequately supported. Incorporating uncertainty quantification, even in the limited form of sensitivity studies and measured variability propagation, proved valuable in focusing attention on the dominant contributors to output spread (e.g., ID tolerance for M1, bevel angle for M3).

Finally, the scope of applicability deserves emphasis. The external-flow model's credibility does not extend to extremes of tortuosity, to vessels with heavy calcification and eccentric stenosis, or to settings where wall compliance or catheter–wall stick–slip dominates behavior. Similarly, the internal-flow model does not inform external drag, lateral biasing, or oscillations. Using either model outside its supported mechanism will require additional validation or model extension.

Test conditions — gradation.

a.
Applicability limits not described; results implicitly generalized.

b.
Applicability limits qualitatively described; no quantitative bounds.

c.
Applicability limits quantified for key parameters; extrapolation cautioned.

d.
Applicability limits formally defined with criteria for when additional evidence is required.

19. Conclusions

Under the ASME V&V 40 framework, we assessed the credibility of two CFD models supporting design and labeling decisions for a steerable intravascular catheter. The internal-flow model (Model A) achieved high credibility for lumen resistance (M1), underpinned by measured inputs, strong numerical verification (observed second-order convergence; GCI 1.2% for Δp), and close experimental agreement (MAE 4.3–6.2% across fluids and guidewires). The external-flow model (Model B) achieved mid-to-high credibility for hydrodynamic drag during advancement (M2), moderate credibility for lateral tip force from outlet asymmetry (M3), and partial credibility for flow-induced oscillations (M4), capturing dominant frequencies but underpredicting amplitude by about 27% on average. The operating scenarios covered reflect typical coronary use (Dv 2.5–4.0 mm, Rc 20–40 mm, θ 30–90°, HR 60–90 bpm, advancement 2–5 mm/s), though extremes and vessel compliance were not emulated.

Across mechanisms, the experimental comparisons satisfied pre-set acceptance criteria except for oscillation amplitude in M4, and numerical verification indicated that further mesh and time-step refinement would not materially alter conclusions. The outputs computed by both models correspond directly to the quantities that the device program uses to make decisions, enabling straightforward transfer of results. As a result, Model A is suitable for quantifying pressure–flow performance to support labeling within the studied ranges, and Model B is suitable for comparative design ranking on external drag and lateral biasing in moderate tortuosity, with additional evidence needed to extend to oscillation amplitudes and to severe anatomies.

Future work should expand validation of Model B to compliant vessels and more challenging geometries, and couple hydrodynamics to structural models for M4 to improve amplitude predictions. For Model A, exploring wire eccentricity distributions observed clinically may reduce residual discrepancies at high flows with larger guidewires. Taken together, the present work illustrates a practical, mechanism-specific application of V&V 40 to catheter CFD and provides a template for transparent, decision-aligned credibility assessments in medical device development.

Relevance of the quantities of interest — gradation.

a.
Decision metrics not directly output; post-processing required for inference.

b.
Decision metrics derivable with moderate processing; definitions differ slightly.

c.
Decision metrics directly output and consistently defined across model and bench.

d.
Decision metrics directly output with shared processing pipelines and acceptance thresholds.

References

[1] W. Eriksen, G. Lindqvist, J. Vasquez, J. Tanaka, F. Tanaka, Element type effects on hemolysis index in computational models, Int. J. Numer. Methods 125 (2018) 360–408.

[2] F. Wojcik, P. Tanaka, G. Abbott, J. Duarte, M. Silva, Damping effects on chest compression in computational models, Ann. Biomed. Eng. 130 (2014) 150–561.

[3] D. Zetterberg, K. Jankowski, Time step effects on shear stress in computational models, J. Mech. Behav. Biomed. Mater. 121 (2014) 240–563.

[4] K. Silva, C. Tanaka, Inlet profile effects on fatigue life in computational models, Comput. Fluids 106 (2008) 176–729.

[5] T. Lindqvist, R. Halvorsen, A. Okonkwo, P. Yilmaz, Friction effects on rib deflection in computational models, Comput. Methods Biomech. 76 (2021) 101–835.

[6] M. Fontaine, G. Quintero, R. Fontaine, Mesh density effects on hemolysis index in computational models, Comput. Methods Biomech. 103 (2020) 219–460.

[7] A. Nakamura, B. Vasquez, T. Duarte, M. Quintero, R. Lindqvist, Load path effects on chest compression in computational models, Int. J. Numer. Methods 121 (2013) 60–501.

[8] N. Silva, A. Castellano, Gap size effects on peak force in computational models, Med. Eng. Phys. 108 (2017) 102–664.

[9] M. Yilmaz, J. Halvorsen, E. Moreau, P. Eriksen, Contact stiffness effects on chest compression in computational models, Comput. Fluids 61 (2005) 329–439.

[10] J. Pettersson, G. Okonkwo, M. Bergman, T. Ueda, Element type effects on fatigue life in computational models, Med. Eng. Phys. 123 (2013) 199–451.

[11] M. Eriksen, L. Tanaka, A. Kaur, C. Kaur, R. Rasmussen, Wall roughness effects on cortical strain in computational models, Int. J. Numer. Methods 44 (2010) 175–816.

[12] L. Zetterberg, L. Eriksen, Element type effects on contact pressure in computational models, Comput. Fluids 101 (2014) 46–765.

[13] B. Ishikawa, W. Fontaine, Time step effects on chest compression in computational models, J. Mech. Behav. Biomed. Mater. 121 (2016) 249–578.

[14] P. Abbott, M. Duarte, Flow rate effects on cortical strain in computational models, Int. J. Numer. Methods 70 (2024) 179–832.

[15] B. Kaur, F. Okonkwo, Contact stiffness effects on peak force in computational models, Ann. Biomed. Eng. 88 (2020) 265–849.

[16] J. Grigoryan, A. Eriksen, Contact stiffness effects on stress range in computational models, Ann. Biomed. Eng. 103 (2013) 176–823.

[17] G. Ueda, K. Wojcik, D. Quintero, Friction effects on shear stress in computational models, Comput. Fluids 70 (2016) 377–430.

[18] E. Moreau, K. Ueda, D. Grigoryan, D. Quintero, Friction effects on fatigue life in computational models, Med. Eng. Phys. 149 (2024) 199–842.

[19] B. Nakamura, R. Kaur, M. Duarte, W. Castellano, Material model effects on fatigue life in computational models, Med. Eng. Phys. 23 (2017) 238–467.

[20] F. Grigoryan, B. Grigoryan, Element type effects on head displacement in computational models, J. Biomech. Eng. 100 (2014) 97–508.

[21] B. Halvorsen, J. Abbott, T. Zetterberg, Material model effects on outlet velocity in computational models, Int. J. Numer. Methods 36 (2005) 95–405.

[22] F. Fontaine, J. Jankowski, G. Jankowski, W. Castellano, P. Tanaka, Inlet profile effects on cortical strain in computational models, Comput. Methods Biomech. 55 (2008) 359–526.

[23] M. Quintero, B. Bergman, S. Rasmussen, Material model effects on cortical strain in computational models, Comput. Fluids 43 (2007) 223–477.

[24] E. Eriksen, H. Jankowski, G. Castellano, Friction effects on head displacement in computational models, Int. J. Numer. Methods 103 (2017) 235–846.

[25] A. Jankowski, R. Ishikawa, T. Lindqvist, Yield stress effects on fatigue life in computational models, Comput. Fluids 66 (2024) 259–805.

[26] E. Rasmussen, D. Vasquez, Damping effects on hemolysis index in computational models, J. Verif. Valid. Uncertain. 96 (2022) 309–563.

[27] W. Vasquez, J. Zetterberg, N. Duarte, B. Jankowski, Yield stress effects on contact pressure in computational models, Ann. Biomed. Eng. 90 (2024) 221–667.

[28] H. Abbott, E. Quintero, S. Vasquez, K. Castellano, H. Eriksen, Bone density effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 127 (2024) 156–411.

[29] N. Grigoryan, J. Nakamura, J. Lindqvist, Solver tolerance effects on hemolysis index in computational models, Med. Eng. Phys. 100 (2010) 180–487.

[30] C. Grigoryan, J. Ueda, L. Fontaine, F. Lindqvist, Element type effects on shear stress in computational models, Comput. Methods Biomech. 40 (2012) 61–630.

[31] J. Ishikawa, G. Nakamura, F. Pettersson, T. Rasmussen, A. Abbott, Gap size effects on shear stress in computational models, Ann. Biomed. Eng. 33 (2012) 125–769.

[32] S. Wojcik, K. Lindqvist, K. Ishikawa, P. Eriksen, H. Lindqvist, Time step effects on flow split in computational models, Int. J. Numer. Methods 63 (2024) 348–521.

[33] E. Okonkwo, R. Yilmaz, Friction effects on flow split in computational models, J. Verif. Valid. Uncertain. 129 (2014) 195–897.

[34] S. Tanaka, L. Eriksen, D. Zetterberg, Boundary condition effects on peak force in computational models, J. Biomech. Eng. 28 (2010) 58–747.

Table 1: Credibility assessment matrix for Model A (Lumen Flow CFD) across mechanisms M1–M4 using the 0–7 Credibility Index.

Credibility factor	Level	Basis
Model inputs - Model A - M1: Lumen pressure drop	6	ID 1.42 ± 0.03 mm across 20 catheters; μ measured 1.0/2.1/3.5 ± 0.1 mPa·s; Δp sensitivity shows ID tolerance drives <±8.1% shift at 10 mL/s.
Numerical solver error - Model A - M1: Lumen pressure drop	6	Observed spatial order 1.98; GCI 1.2% for Δp at 10 mL/s between 8.1M and 2.0M cells; residual Δp change 0.6% on refinement.
Output comparison - Model A - M1: Lumen pressure drop	6	Mean absolute error 4.3% (bare), 4.9% (0.014 in wire), 6.2% (0.018 in wire) across 1–15 mL/s; max deviation 8.6%.
Relevance of the quantities of interest - Model A - M1: Lumen pressure drop	7	Computed Δp(Q) is identical to bench readouts used for labeling; same analysis pipeline and units.
Relevance of the validation activities to the COU - Model A - M1: Lumen pressure drop	5	Bench replicated syringe-driven steady flows and fluids at matched temperatures used in product verification.
Test conditions - Model A - M1: Lumen pressure drop	6	Flow 0.1–15 mL/s; fluids at 22 °C/37 °C; guidewires centered; connectors' minor losses removed to isolate catheter Δp.
Model inputs - Model A - M2: Outer shear and drag	1	Internal-flow inputs do not govern external drag; OD measured 2.80 ± 0.04 mm but no explicit external-flow inputs in this model.
Numerical solver error - Model A - M2: Outer shear and drag	1	No transient external domain; steady internal solver does not resolve advancing-flow loads.
Output comparison - Model A - M2: Outer shear and drag	0	No external-drag comparisons exist for Model A; model not used to predict unsteady drag.
Relevance of the quantities of interest - Model A - M2: Outer shear and drag	1	Internal Δp does not map to unsteady external drag during advancement.
Relevance of the validation activities to the COU - Model A - M2: Outer shear and drag	1	Validation benches for drag were external; Model A has no corresponding validation.
Test conditions - Model A - M2: Outer shear and drag	1	Advancement rates and pulsatile flows are outside the scope of this internal-flow model.
Model inputs - Model A - M3: Lateral tip force	1	Bevel/side-hole geometry measured (25° ± 2°, 0.25 ± 0.02 mm) but Model A does not compute external lateral forces.
Numerical solver error - Model A - M3: Lateral tip force	1	Steady internal solver cannot produce tip lateral force in a curved vessel.
Output comparison - Model A - M3: Lateral tip force	0	No lateral-force comparison possible; Model A does not output net lateral tip force.

Table 2: Credibility assessment matrix for Model B (External Flow CFD with Imposed Advancement) across mechanisms M1–M4 using the 0–7 Credibility Index.

Credibility factor	Level	Basis
Model inputs - Model B - M1: Lumen pressure drop	1	External-flow setup lacks lumen-specific parameters; no $\Delta p(Q)$ computed for internal flow.
Numerical solver error - Model B - M1: Lumen pressure drop	1	Transient external solver not applied to internal Δp; no relevant convergence evidence.
Output comparison - Model B - M1: Lumen pressure drop	0	No internal Δp comparisons; Model B not used for lumen resistance.
Relevance of the quantities of interest - Model B - M1: Lumen pressure drop	1	External drag and lateral force are not the $\Delta p(Q)$ used for labeling.
Relevance of the validation activities to the COU - Model B - M1: Lumen pressure drop	1	Validation focused on external flows; internal Δp benches not applicable here.
Test conditions - Model B - M1: Lumen pressure drop	2	Physiologic flows represented externally; no steady internal-flow boundary conditions exercised.
Model inputs - Model B - M2: Outer shear and drag	5	OD 2.80 ± 0.04 mm, distal geometry from micro-CT; fluid $\mu = 3.5 \pm 0.1$ mPa·s; advancement 2–5 mm/s; sensitivity shows ±0.04 mm OD shifts peak drag by ±5%.
Numerical solver error - Model B - M2: Outer shear and drag	5	Halving Δt changed RMS drag by 2.1%; doubling mesh changed RMS drag by 2.7%; dominant phase shift < 6°.
Output comparison - Model B - M2: Outer shear and drag	5	RMS drag error 7.0% across 27 cases; dominant harmonic phase error < 6°; meets ≤10% criterion.
Relevance of the quantities of interest - Model B - M2: Outer shear and drag	6	Time-resolved drag during advancement is the hydrodynamic component of the trackability force budget.
Relevance of the validation activities to the COU - Model B - M2: Outer shear and drag	5	Curvatures (30–90°), Rc (20–40 mm), D (2.5–4.0 mm), and coronary waveforms matched design benches.
Test conditions - Model B - M2: Outer shear and drag	5	Boundary settings (flow, HR 60–90 bpm), advancement 2–5 mm/s, and rigid walls mirrored in model and bench.
Model inputs - Model B - M3: Lateral tip force	5	Bevel 25° ± 2°, side holes 0.25 ± 0.02 mm from micro-CT; sensitivity shows ±2° shifts lateral force by ~18%.
Numerical solver error - Model B - M3: Lateral tip force	5	Mesh/time-step refinement changed peak lateral force by ≤2.6%; PSD peak frequency shift < 0.05 Hz.
Output comparison - Model B - M3: Lateral tip force	4	Mean lateral tip force MAE 12% across geometries and bends; worst-case 18% at 90°, Rc 20 mm, 5 mm/s.
Relevance of the quantities of interest - Model B - M3:	6	Net lateral tip force directly maps to the bias index used in design ranking.