# Credibility assessment under ASME V&V40 of CFD models for ventricular assist inflow thrombogenicity: two models, three mechanisms

Alyssa M. Duarte, PhD[a], Nikhil Iyer, MS[b], Paul J. Sanders, PhD[c], Mei Chen, PhD[d], Graham R. Holt, MD PhD[1]

[a] *Center for Cardiovascular Engineering, North River Institute, Boston, MA, USA,*
[b] *Department of Mechanical Engineering, Westlake University, Portland, OR, USA,*
[c] *Translational Hemodynamics Lab, Mid-Atlantic Biomedical, Baltimore, MD, USA,*
[d] *Department of Surgery, Cardiac Division, State University Health System, Albany, NY, USA,*

**Abstract**

Computational fluid dynamics is increasingly used to investigate thrombotic risk in ventricular assist devices, yet the credibility of such analyses depends on systematic evidence that the models are fit for purpose. We report a structured credibility assessment under the philosophy of ASME V&V40 for two inflow-region models of a left ventricular assist device: a steady Reynolds-averaged Navier–Stokes formulation with shear-stress transport k–omega closure and non-Newtonian rheology (Inflow-RANS), and a wall-resolved large-eddy simulation coupled to Lagrangian surrogate platelets (Inflow-LES-PaRT). The intended use is to inform design iteration and hazard analysis at the cannula lip and inlet housing, focusing on three mechanisms: high-shear platelet activation, recirculation-driven stasis, and surface-mediated deposition. We define a private 0–12 scale for credibility scoring, apply it to six evidence categories for each model–mechanism pair, and summarize the outcomes in model-specific tables. The assessment integrates code-oriented tests, mesh and time-step sensitivity studies, mechanistic plausibility arguments, comparison to benchtop measurements in a physiologically relevant mock loop, relevance of the experimental conditions to the device and use questions, and the software engineering practices surrounding the computational tools. Findings are consistent with the expectation that unsteady, wall-resolved modeling is advantageous for capturing shear transients, residence distributions, and near-wall transport that condition deposition. For Inflow-RANS, the most consistent strengths are efficient exploration of design space and acceptable accuracy in predicting shear hotspots, with qualifications for stasis and deposition. For Inflow-LES-PaRT, the strongest evidence resides in physically consistent resolution of the near-wall dynamics and improved agreement with washout and deposition data, albeit at higher cost. The credibility levels for the six evidence categories span 4/12 to 10/12 across mechanisms for Inflow-RANS and 8/12 to 10/12 for Inflow-LES-PaRT. We also report an overall credibility index on the same 0–12 scale for each mechanism and model to aid decision making by users. The work demonstrates how a practice-guided, mechanism-specific assessment yields discriminating insights for computational hazard analysis in ventricular assist inflows.

*Keywords:* ASME V&V40, ventricular assist device, thrombogenicity

## 1. Introduction

Ventricular assist devices (VADs) are life-sustaining implants whose safety and effectiveness hinge on careful management of blood damage and thrombotic complications. The inflow cannula and inlet housing are particularly critical because they encompass the transition from ventricular flow into the rotor's domain. Local acceleration, complex separation, and near-wall transport in this region can condition platelet activation and deposition, initiating clinical events that range from transient ischemic attacks to pump thrombosis. Computational fluid dynamics (CFD) has matured to the point that it is often used to interrogate these processes during design and risk management. Yet, the field still faces skepticism when unverified or weakly validated models are used to make high-stakes inferences about thrombogenicity. Practice-oriented frameworks that explicitly tie evidence to a specified question of use are needed.

ASME V&V40 provides a way to structure credibility for computational models used in medical device evaluation. The standard emphasizes matching the credibility activities and their rigor to the context of use, disaggregating evidence for distinct risk mechanisms, and reporting results in a way that decision makers can weigh. In VAD thrombogenicity, a first-order challenge is that seemingly small flow features can have outsized effects on stress histories and residence times experienced by platelets, and that deposition at the flow–surface interface depends on the coupled evolution of hydrodynamics and adhesive kinetics. A second challenge is that different mechanisms dominate depending on geometry and operating condition; a single composite metric often obscures which mechanisms are (or are not) well represented. In both cases, mechanism-centric credibility assessments provide more actionable information than model-centric ones alone.

This paper applies a V&V40-aligned approach to two CFD models largely representative of current practice. The first is a steady Reynolds-averaged Navier–Stokes (RANS) model with shear-stress transport (SST) k–omega closure augmented by a shear-thinning representation of blood. It computes wall shear stress (WSS), an Eulerian scalar surrogate for residence time, and an Eulerian platelet activation index. The second is a wall-resolved large-eddy simulation (LES) incorporating Lagrangian platelet surrogates whose stress-exposure histories and near-wall behavior are recorded. The goal is not to crown a winner but to differentiate where each brings sufficient fidelity and where caution or additional data are needed for reliable hazard analysis.

We define the intended application as design-stage risk reduction for the inflow region of a left ventricular assist device (LVAD). We ask whether the models can provide trustworthy information for three mechanisms: (1) high-shear platelet activation in the near-inlet acceleration and shear layers, (2) recirculation-driven stasis in separated flow adjacent to the cannula and housing, and (3) surface-mediated deposition conditioned by persistent low wall shear and near-wall transport. Following V&V40, we assemble evidence from code-oriented tests, grid and time-step sensitivity, mechanistic rationale linking equations to questions, quantitative comparison to bench measurements in a physiologically representative loop, the match between that loop and the device's intended range of use, and the software engineering context in which the models are developed and run.

We organize the paper to make these links explicit. After stating the context of use and the mechanisms of interest, we describe the models and the validation experiments we rely on. We then lay out the scoring approach on a private 0–12 scale and present the evidence. Crucially, because thrombogenicity is not a single number, the results are summarized as credibility tables by model, with one row per factor–mechanism combination, and text that makes transparent what each number rests upon. Although the framework follows the spirit of V&V40, the scoring scale is private to this assessment and is not intended to imply a universal mapping.

A central element is that findings are tied to mechanisms, not only to models. For example, a RANS model may be perfectly adequate for ranking designs by the magnitude of peak WSS at the cannula lip but less appropriate for quantifying residence-time distributions in a low-frequency unsteady separation bubble. Similarly, even a high-resolution LES may not capture biochemical adhesion kinetics if these are not represented, which in turn limits the strength of conclusions about surface-mediated deposition. We therefore caution the reader that the same model can receive different credibility levels across mechanisms.

## 2. Context of use and model risk rationale

The intended application is design-stage hazard analysis for the inflow region of a commercial-scale LVAD. The questions the models are expected to inform are: (a) where and by how much do shear stresses near the cannula lip and inlet housing exceed relevant thresholds for platelet activation over clinically relevant ranges of operating conditions; (b) what is the extent and strength of recirculation zones that could produce stasis adjacent to the inflow; and (c) which surfaces and locations present sustained low-shear and high near-wall residence time that predispose to surface deposition. The device geometry considered includes a 26 French inflow cannula blended to an inlet housing with a bell-mouth transition and an upstream ventricular segment to provide realistic inflow boundary layers. Typical operating flows span 3–6 L/min against physiologic afterload representations in a mock circulatory loop.

The model users are design engineers and risk analysts in an industrial R&D setting, supported by computational specialists who build meshes, set up simulations, and interpret results. The consequences of misinformed decisions include delayed design iterations, suboptimal safety margins, and in the worst case, unrecognized hotspots that would then need to be addressed later in verification and validation up the device development pipeline. Although clinical decisions are not made directly from these models, the analyses contribute to the body of evidence that informs design choices and risk assessments; the model risk is therefore nontrivial. We characterize model risk as medium for high-shear activation and stasis assessments, and medium-high for predictions that would directly drive surface material choices or coatings based on deposition propensity.

Given that the question of use is design and hazard analysis rather than regulatory substantiation of safety claims, the tolerance for model imperfection is not nil. However, the LVAD inflow is a critical zone for thrombogenic mechanisms; systematic bias or large uncertainty could mis-rank design variants and conceal vulnerabilities. In this sense, both a quantitative sense of the magnitude of error and the mechanistic plausibility of assumptions are important, and the credibility activities should be commensurate with the mechanistic sensitivity. The three mechanisms identified require different elements of model resolution and physics: high-shear activation is sensitive to instantaneous local stress magnitude and exposure duration; stasis depends on separated, often unsteady vortical structures and the residence-time tails; deposition involves the near-wall transport and low-shear persistence combined with surface interactions. This diversity motivates a mechanism-stratified credibility assessment.

The models are not used to make claims about long-term thrombus growth or occlusion, nor for patient-specific prediction of thrombotic events. Instead, they support comparative statements within a geometry family, identification of hotspots, and estimation of margins to known platelet activation thresholds or stasis criteria. The consequences of error are therefore bounded: decisions remain subject to additional experimental testing and design controls before clinical use. This context justifies a mixed level of rigor across the evidence categories, with the heaviest emphasis placed on those that most directly underpin mechanism-specific reliability.

*Model form - high-shear mechanism — gradation.*

a.
Equations and closures chosen by convention without mechanism-specific rationale; no mapping from modeled stresses to activation metrics.

b.
Mechanism described qualitatively; literature used to justify turbulence closure and rheology; thresholds cited but not explicitly connected to computed exposure.

c.
Mechanism mapped to model quantities with explicit thresholds and exposure-duration logic; limitations of closure near strong acceleration acknowledged.

d.
Mechanism mapped quantitatively to platelet-scale exposure with uncertainty bounds; closure behavior in acceleration and separation validated against dedicated benchmarks.

## 3. Models under assessment

Two computational models are assessed. Both represent the fluid domain encompassing the ventricular segment upstream of the cannula, the inflow cannula including its beveled lip, and the inlet housing up to the rotor hub plane. The geometry is derived from a near-production LVAD stage and includes a bell-mouth transition and a short diffuser section. The working fluid is blood analog in bench measurements and modeled as blood with shear-thinning rheology in simulations. The models differ primarily in their treatment of turbulence, unsteadiness, and the approach to track platelet-relevant quantities.

The Inflow-RANS model solves the steady Reynolds-averaged Navier–Stokes equations with the shear-stress transport (SST) k–omega closure. It uses a Carreau–Yasuda representation of blood viscosity with parameters fitted to hematocrit 40% at 37 C, with zero-shear viscosity 0.16 Pa*s, infinite-shear viscosity 0.0035 Pa*s, time constant 3.313 s, and exponents an=2 and n=0.3568. Boundary conditions include a developed inflow at the ventricular segment based on a Womersley representation for the pulsatile component, but the RANS solution is formulated at a quasi-steady representative flow rate for each operating condition assessed. Outlet pressure is set to match the mock loop mean afterload. Near-wall treatment employs a low-Reynolds formulation of SST with y-plus values targeted below 1 at the inflow lip and adjacent walls. From the converged solution, derived fields include wall shear stress magnitude and direction, an Eulerian scalar residence-time surrogate computed as the solution of a steady advection–reaction equation with a small decay mimicking washout, and an Eulerian platelet activation index obtained by integrating a power-law function of the strain-rate magnitude along streamlines seeded in the near-inlet flow. The discretization uses a body-fitted, polyhedral mesh with boundary-layer prisms resolving the viscous sublayer at the cannula lip. Mesh sizes range from 9.8 million to 41 million control volumes in the grid-convergence study.

The Inflow-LES-PaRT model solves the unsteady, filtered Navier–Stokes equations using a wall-resolved large-eddy simulation. The subgrid stress is modeled with the Wall-Adapting Local Eddy-viscosity (WALE) model to improve near-wall behavior, and the viscosity follows the same Carreau–Yasuda rheology. The mesh is a hybrid unstructured grid with near-wall prismatic layers achieving y-plus below 1 over the lip and housing interior walls, and isotropic tetrahedra in the core with target filter width matched to the Kolmogorov scale in the shear layers. The nominal grid for production runs includes 128 million control volumes; a coarser and a finer grid at 62 million and 210 million cells, respectively, are used in sensitivity studies. Time integration uses a second-order implicit scheme with fixed time steps of 1e-5 s in baseline runs and 5e-6 s in refined runs. The simulation is driven with a physiologically consistent pulsatile inflow waveform at 75 bpm and mean flows of 3–6 L/min, with outflow pressure matched to the bench loop impedance. Lagrangian platelet surrogates are injected at the ventricular segment with flux proportional to volumetric flow. Approximately 200,000 tracers are advanced per cardiac cycle using a second-order accurate predictor–corrector scheme with substepping to capture near-wall velocity gradients. For each surrogate, the time history of shear exposure integral is accumulated and the minimum wall distance and near-wall shear history recorded. A deposition propensity flag is toggled when a platelet surrogate experiences sustained near-wall shear below a threshold (0.5 Pa for 0.5 s) while the local wall shear remains below 1 Pa, in line with criteria suggested in the experimental literature for onset of adhesive stabilization.

Both models are implemented in a single code base that shares geometry processing, mesh generation, and post-processing pipelines. The RANS and LES solvers share flux reconstruction algorithms but use different linear solvers and preconditioners tailored to steady and unsteady problems, respectively. The particle tracker is a separate module that interfaces with the LES code via a data exchange layer, allowing particles to be moved at substeps of the fluid solver when needed. All simulations are executed in containerized environments to ensure consistency across clusters. Post-processing routines extract WSS maps, streamline-based activation indices, residence-time distributions, and deposition propensity fields. Comparisons to bench measurements are scripted so that the same spatial registration, binning, and statistical calculations are used for both models where possible.

The primary outputs per mechanism are: (1) for high-shear platelet activation, maps of the 99th percentile WSS per cardiac cycle, the Eulerian platelet activation index for RANS, and the distribution of Lagrangian stress-exposure integrals for LES; (2) for recirculation-driven stasis, Eulerian residence-time surrogate fields for RANS and particle-based residence-time distributions for LES, both summarized by percentile metrics and spatial extent of slow-flow regions; and (3) for surface-mediated deposition, near-wall WSS persistence maps (WSS below 1 Pa for at least 0.5 s per cycle) and, for LES, the particle-based deposition propensity locations. These link to bench observables described in a subsequent section.

Because the same device geometry underlies both models, the comparison places the onus on numerical formulation and resolution rather than geometric differences. Choice of non-Newtonian viscosity matters most for high-shear features at the lip and in the acceleration region; both models use the same rheology for consistency of comparisons. The key distinction then lies in whether unsteadiness and near-wall dynamics are essential for a given mechanism. In the inflow region of an LVAD, peak shear transients at the cannula lip are phase-locked to inflow acceleration and deceleration. Unsteady separation at the bell-mouth can create coherent vortical structures that persist into the inlet housing. These phenomena motivate the LES approach for the most demanding questions, while the RANS approach offers efficient exploration and first-order answers.

*Output comparison - deposition-focused rubric — gradation.*

a.
Qualitative visual agreement asserted in selected planes; no spatial registration; no test of discrimination.

b.
Quantitative comparison in a few locations; weak spatial matching; only central tendencies.

c.
Whole-field registration with correlation and slope analysis; uncertainty bounds on measurement and model; discrimination tested with ROC/AUC.

d.
Mechanism-specific metrics (e.g., low-shear persistence) matched in time and space; cross-validation across operating points; combined spatial-temporal performance summarized.

## 4. Thrombogenic mechanisms at the inflow

High-shear platelet activation arises where the flow accelerates through constricted regions or where strong shear layers form. At an LVAD inflow, this happens at the beveled lip, along the leading-inner edge of the bell-mouth, and sometimes along the inner wall of the inlet housing if the inflow is skewed. Platelets exposed to supraphysiologic shear for short durations can be primed, altering their sensitivity to subsequent biochemical agonists. The relevant variables are the magnitude and duration of shear exposure along trajectories that pass through acceleration zones. Because both amplitude and time-integral of shear matter, dynamic features near the lip during flow acceleration are particularly influential.

Recirculation-driven stasis is governed by separated flow regions and coherent vortices adjacent to the inflow cannula and housing. Where the main jet separates from the wall or the bell-mouth, a low-velocity zone can form and persist during parts of the cardiac cycle. Particles that enter these regions can linger, increasing the probability of thrombin generation and fibrin network formation. From a modeling point of view, the important questions are whether the extent and persistence of separated regions are captured and whether the residence-time tails are accurately represented. Because residence-time distributions depend on slow processes and intermittency, they are sensitive to how unsteadiness and turbulence near the wall are represented.

Surface-mediated deposition couples near-wall hydrodynamics and adhesive processes. Sustained low wall shear and weak near-wall convective transport favor adhesion and stabilization of platelet aggregates on biomaterial surfaces. In the inflow region, the inner face of the bell-mouth and the blend radius to the cannula are common sites of low shear during parts of the cycle. The substrate properties—roughness, wettability, and chemistry—modulate adhesion probability. From the point of view of a CFD model, near-wall shear persistence and how particles approach and remain near the surface are key. Without an explicit biochemical adhesion model, inference about deposition must rest on surrogate criteria such as low-shear persistence metrics and the frequency with which platelets experience near-wall low-shear episodes.

*Model form - deposition-focused rubric — gradation.*

a.
No explicit near-wall transport consideration; no link to adhesion criteria; steady-state view only.

b.
Near-wall fields computed but assessed by snapshot metrics; simple shear-threshold criteria used without time dimension.

c.
Combined shear–time criteria implemented; particle–wall encounter rates estimated; substrate properties considered qualitatively.

d.
Mechanistic mapping includes shear–time–distance metrics; deposition propensity calibrated against multiple materials; limitations of missing biochemical kinetics quantified.

## 5. Credibility plan and scoring rubric (0–12 scale)

The credibility plan follows the spirit of ASME V&V40 while adapting its application to a mechanism-specific assessment across two CFD models. For each model–mechanism pair, six evidence categories are scored on a private 0–12 scale: numerical tests to assure correct implementation of equations, mesh and time-step sensitivity of the outputs of interest, mechanistic adequacy of the chosen equations and closures for the question at hand, comparison to bench data that probe the same mechanisms, the relation of those experiments to the intended operating range and device features, and the software engineering practices that govern reproducibility and traceability. The seventh topic we discuss is operational use and human workflow, but it is not graded in this assessment. This layout intends to give decision makers a concise view of where each model is strong or limited for each mechanism while keeping the assessment readable.

The 0–12 scale is partitioned informally into four bands to aid interpretation. Levels 0–3 reflect minimal evidence relative to the intended use; scores in this band indicate that the activity was attempted but failed to anchor conclusions. Levels 4–6 denote limited evidence, usually with quantitative elements but gaps that constrain confidence. Levels 7–9 reflect moderate-to-strong evidence aligned with the question of use, with quantified agreement or sensitivity bounds and transparent limitations. Levels 10–12 indicate extensive and discriminating evidence tailored to the mechanism and question, with triangulation across multiple data sources and uncertainty quantification commensurate with the decision stakes. Scores reported here occupy the 4–10 range because the intended use is design-stage hazard analysis rather than a final safety demonstration.

[redacted]

Figure 1: Schematic of the LVAD inflow region modeled by both approaches. The domain includes a ventricular segment, a beveled inflow cannula lip, a bell-mouth inlet housing, and the rotor hub plane. For LES, particle injection and Lagrangian tracking occur from the ventricular segment, with deposition propensity assessed at interior walls.

The specific scoring decisions were made by a three-person panel comprising a turbulence modeler, a hemodynamics expert, and a V&V specialist. Decisions were based on pre-defined criteria for each evidence category, cross-checked against the available data and analyses, and then consolidated into a single number per category per model–mechanism. Where the panel diverged, the lowest score among the candidate values was adopted unless one member identified an error of fact, in which case the item was corrected and rescored. The numbers are accompanied by succinct bases in the tables and detailed discussion in the text. While the scale is private to this exercise, its rungs are described to ensure transparent mapping from evidence to numbers.

*Private 0–12 scale mapping — gradation.*

a.
0–3 Minimal evidence; activity attempted but not aligned to the question; qualitative or anecdotal support only.

b.
4–6 Limited evidence; quantitative results present but with material gaps or weak linkage to mechanism; uncertainty unquantified or large.

c.
7–9 Moderate-to-strong evidence; mechanism-targeted tests and comparisons; quantified differences and uncertainties; limitations declared.

d.
10–12 Extensive evidence; multi-pronged, mechanism-tuned program with uncertainty propagation; high fidelity match to intended use.

## 6. Numerical code verification

To provide confidence that the solvers are correctly implemented, we performed a suite of tests that exercise the discretization, flux reconstruction, turbulence closure terms, and, for the LES, the subgrid model coupling. For the steady solver used by Inflow-RANS, we employed manufactured solution tests that target the coupled pressure–velocity formulation in a curvilinear coordinate setting representative of the inflow bell-mouth. A trigonometric manufactured field with analytic source terms was used to drive the equations; the observed convergence in L2 norm for both velocity components and pressure was consistent with the second-order spatial accuracy of the scheme. The SST k–omega closure was verified by constructing a manufactured turbulent kinetic energy and specific dissipation rate field with appropriate production and dissipation terms; the resulting error norms decreased at near-second order. Discrete conservation checks confirmed that the mass residuals reduced eight orders of magnitude from initial guess and momentum residuals by seven orders, with no pathologies in element skewness seen to impair convergence.

For the unsteady solver employed by Inflow-LES-PaRT, canonical turbulent channel flow at friction Reynolds number Re_tau=180 was simulated to steady-state statistics from which mean velocity profiles and Reynolds stress distributions were compared against direct numerical simulation data. The inner-scaled mean velocity showed close adherence to the log-law, and deviation in u-plus at y-plus of 15 was within a few percent across the domain. The energy spectra exhibited the expected inertial subrange scaling with appropriate decay. The WALE model parameters were checked against a set of decaying isotropic turbulence tests to confirm correct asymptotics in laminar regions near walls. A scalar-transport manufactured solution verified the implementation of the non-Newtonian viscosity in the context of unsteady advection–diffusion processes. The particle tracker underwent unit tests to ensure correct interpolation of velocities, substepping across near-wall gradients, and conservation of particle counts in the absence of sinks or sources.

Because implementation bugs often arise from corner cases, we also included targeted tests that mirror inflow features: an accelerating nozzle with bevel geometry was simulated with both solvers in laminar conditions to assess potential geometric Jacobian inconsistencies; the results demonstrated the expected pressure recovery and boundary layer development. Checkerboard tests confirmed that Rhie–Chow-type pressure–velocity coupling avoided oscillatory pressure fields on collocated grids. Lastly, for the LES, we conducted time-reversal tests in laminar, non-diffusive settings for the particle tracker to ensure that round-trip integration returned particles to their origin within tolerances when the velocity field was reversed.

These tests served to build confidence that the algorithms compute the intended equations and that the deformation of the mesh and presence of wall layers do not introduce nonphysical coupling between variables. The manufactured solution approach also allowed us to explicitly separate discretization effects from implementation errors, as the observed orders of accuracy stably approached the theoretical values as the mesh was refined. Taken together, the verification activities were designed to be relevant to the inflow problem while still providing general coverage of the code's capabilities.

*Numerical code verification — gradation.*

a.
No formal tests; reliance on published solver claims; ad hoc problem runs only.

b.
Basic regression tests; a few unit tests; canonical case compared qualitatively; no manufactured solutions.

c.
Targeted MMS for core equations; canonical turbulence benchmark with quantitative agreement; coverage of particle tracking where applicable.

d.
Comprehensive MMS across coupled equations; multiple turbulent benchmarks; fault injection and stress testing; traceability of results and reproducibility documented.

## 7. Discretization error estimation

For the steady solver, mesh sensitivity was assessed using three systematically refined unstructured grids constructed via an automated procedure that preserves boundary-layer topology and expansion ratios while globally scaling element sizes. The near-wall prism layers at the cannula lip were refined to maintain y-plus values below one across grids. The grids contained approximately 9.8 million, 21 million, and 41 million control volumes, with characteristic cell size ratios of roughly 1.3 between consecutive grids. For each operating condition at representative mean flows of 3, 4.5, and 6 L/min, we computed the primary quantities of interest. We constructed grid convergence indices using the observed convergence rate for each output separately, recognizing that near-singular behavior of WSS at edges requires care in Richardson extrapolation. Special attention was paid to the 99th percentile wall shear stress in a strip encircling the cannula lip, to the Eulerian residence-time surrogate in recirculation-prone zones, and to the Eulerian platelet activation index calculated along streamlines seeded near the acceleration region.

For the unsteady solver, both spatial and temporal refinement were executed. Three grids at approximately 62, 128, and 210 million control volumes were run with time steps of 2e-5 s, 1e-5 s, and 5e-6 s, respectively. The chosen combinations targeted a Courant number under one near the inflow lip and under two in the core flow. Production runs employed the middle grid and time step, while refinement runs allowed us to estimate sensitivity of peak and percentile measures across cycles. Because unsteady statistics need sufficient sampling, all comparisons were conducted after discarding three initial cycles and averaging over five subsequent cycles. Lagrangian particle calculations used substepping at 2e-6 s for the base and 1e-6 s for the refined case to mitigate interpolation errors in high-shear zones. Error estimation focused on the 99th percentile WSS over each cycle, the 95th percentile of particle-based residence time, and the spatial fraction of the wall experiencing WSS below 1 Pa for at least 0.5 s per cycle.

Differences across grids and time steps were decomposed into contributions from spatial and temporal resolution by conducting paired runs at different time steps on the same mesh for a subset of operating conditions. The observed orders of accuracy were computed pragmatically without assuming asymptotic range behavior, using least-squares fits of outputs versus nominal grid spacing where appropriate. Where transients dominated, we reported cycle-to-cycle variability alongside refinement-induced changes to place sensitivity in context. In low-shear persistence calculations, binary classification thresholds can produce discontinuities under refinement; to avoid spurious sensitivity, we used a smoothed indicator function for wall shear persistence when measuring convergence.

The sensitivity exercises were designed to target mechanism-relevant outputs rather than general norms. The interplay between grid resolution, wall modeling (or its absence), and the time step is complex in an accelerating flow past a sharp lip; the chosen refinement pattern sought to probe this interplay without becoming intractably expensive. While the full numerical details are deferred to later sections that summarize results by mechanism, this section documents that the program of mesh and time-step studies was sufficient to support mechanism-focused sensitivity statements further on.

*Mesh and time-step convergence - WSS focus — gradation.*

a.
Single grid or time step; no sensitivity study; results reported without error bars.

b.
Two-grid or two-time-step comparison; qualitative convergence claimed; no formal extrapolation.

c.
Three or more levels; observed order estimated; outputs of interest analyzed; variability reported for unsteady problems.

d.
Combined space–time refinement; mechanism-specific metrics; uncertainty bounds computed; sensitivity to near-singular features explicitly handled.

## 8. Model form justification by mechanism

The choice of steady RANS with SST k–omega closure for the first model is motivated by two factors: its capacity to represent mean shear layers and attached boundary layers with reasonable accuracy across engineering flows, and its efficiency in parameter sweeps that explore geometric variants and inflow conditions. For high-shear activation questions centered on where the peak wall shear stress occurs and its approximate magnitude at design points, the SST model generally performs well, particularly when the geometry-induced acceleration is strong and dominant over turbulence fluctuations. The addition of shear-thinning viscosity further aligns the representation with blood rheology, which reduces over-prediction of shear near the lip under high shear-rate conditions. For stasis, however, steady closures are known to struggle with the intermittent, unsteady nature of recirculation zones—both in their extent and in the strength of secondary flows that entrain fluid back into the main stream. A quasi-steady viewpoint cannot capture cyclic reattachment and vortex shedding. For surface-mediated deposition questions that hinge on low-shear persistence and near-wall transport, steady models omit the fluctuations that can periodically disrupt low-shear zones, thereby potentially under- or over-predicting the persistence depending on bias in the computed mean field.

The wall-resolved LES in the second model seeks to resolve unsteady separation and near-wall dynamics that are central to all three mechanisms. By eliminating a wall model and achieving y-plus below one, the simulation can represent the time-varying wall shear stress in acceleration and deceleration phases and resolve coherent vortical structures that govern entrainment into and out of recirculation zones. The Lagrangian platelet surrogates add a mechanistic bridge between the flow and platelet-scale exposure histories. The ability to compute stress-exposure integrals along particle trajectories connects modeled flow features to probabilistic activation outcomes more directly than Eulerian proxies. For surface-mediated deposition, the inclusion of a persistent low-shear criterion in the near-wall region provides a more direct surrogate for adhesive stabilization than instantaneous shear thresholds alone.

There are, however, limitations in both models that matter for the mechanisms of interest. Neither model includes biochemical activation pathways explicitly, and only the LES-based approach represents some aspects of platelet–wall interactions via surrogate criteria. In the RANS framework, the Eulerian platelet activation index treats exposure as an integral along mean streamlines, which can miss intermittency in stress histories. In the LES framework, the absence of adhesion receptor kinetics and biochemical stabilization means that the deposition propensity flag is a surrogate that may be conservative or liberal depending on parameter choices for shear thresholds and durations. Neither model represents RBC-induced aggregation or margination explicitly, which are known to influence near-wall platelet concentration and hence deposition probability. These gaps are acknowledged and we limit the intended use accordingly.

Ultimately, the model form adequacy judgments depend on how much the intended questions are controlled by time-varying, near-wall phenomena. For high-shear activation at the lip, the location and approximate magnitude of peak WSS are strongly geometry- and acceleration-driven and less sensitive to fine-scale turbulence dynamics; this favors the RANS approach as acceptable for initial hazard mapping. For stasis and deposition, where the dynamics of separation and the persistence of low-shear regions matter, the LES approach is better aligned with the question of use.

*Model form - stasis-focused rubric — gradation.*

a.
Mean-flow assumptions adopted; no explicit representation of separation dynamics; stasis inferred indirectly.

b.
Unsteadiness partially represented; residence approximated by scalar surrogates; separation behavior compared to generic benchmarks.

c.
Unsteady separation resolved; residence distributions computed; coherent structures characterized; limitations of missing hemorheology effects noted.

d.
Mechanism explicitly linked to resolved flow structures; long-tail residence modeled; sensitivity to inflow pulsatility and turbulence documented.

## 9. Output comparison for each mechanism

Bench measurements were performed in a transparent, index-matched mock loop reproducing the inflow geometry and physiologic impedance. The working fluid matched the refractive index of the acrylic housing and had viscosity calibrated to mimic blood at shear rates relevant to the inflow region. Pulsatile operation at 75 bpm with mean flow rates of 3, 4.5, and 6 L/min covered the intended range. Particle image velocimetry (PIV) provided planar velocity fields in meridional and circumferential planes intersecting the cannula lip and bell-mouth. From these fields, wall shear stress maps were recovered via near-wall extrapolation limited to regions with adequate seeding and careful wall detection. Hotspot locations and magnitudes were extracted as per-cycle percentiles and averaged across cycles. The uncertainty of PIV-derived WSS was estimated by propagating measurement noise, wall location uncertainty, and gradient estimation errors; the combined uncertainty at WSS hotspots was approximately 8–12 percent depending on plane and operating point.

To probe stasis, a dye washout protocol was implemented. A small bolus of fluorescent dye was injected near the wall upstream of the cannula lip during mid-diastole, and its concentration decay was recorded in pre-defined interrogation regions adjacent to the bell-mouth and along the inner housing wall. The washout time T90 (time to 90 percent concentration reduction) was computed for each region and phase. The method emphasizes slow exchange dynamics and is sensitive to recirculation extent and strength. From a mechanistic perspective, the washout measurements correspond to upper-percentile residence times predicted by the models. The experimental uncertainty in T90 derives from dye dispersion, camera noise, and timing, and was estimated at 5–10 percent across conditions, increasing in regions where signal-to-noise is low.

For surface-mediated deposition, the mock loop was run with washed human platelets fluorescently labeled and re-calcified, and surfaces were coated with an albumin layer to maintain comparable surface energy to the device substrate. Runs lasted two hours at each operating condition, and deposition fields were imaged at the end using confocal microscopy through the transparent walls. The images were registered to the computational geometry to allow point-by-point comparison. The substrate's roughness (Ra) and wettability (static water contact angle) were measured prior to runs to document their relation to the device. For analysis, a deposition mask (presence/absence by thresholded intensity) and a semi-quantitative deposition density were computed. ROC curves assessed the models' ability to discriminate regions with and without deposition, and regression of model-derived low-shear persistence versus deposition density probed calibration.

Comparison metrics were selected to be mechanism-relevant. For high-shear activation, the 99th percentile wall shear stress per cycle and the spatial location of hotspots along the lip are the primary comparators. For stasis, the 95th percentile of residence time and T90 values in regions adjacent to the inflow are compared. For deposition, the spatial correlation between model-inferred deposition propensity and measured deposition, the slope of the best-fit line, and the area under the ROC curve (AUC) provide a balanced set of performance descriptors. The overall uncertainty of the experimental data and their spatial registration was propagated into uncertainty bounds on the comparisons. The detailed results for each model and mechanism are integrated into the credibility scores further on.

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

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

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

*Output comparison - activation-focused rubric — gradation.*

a.
Velocity-only comparison with qualitative hotspots; no WSS extraction; no percentile analysis.

b.
WSS extracted but with unknown uncertainty; hotspot location compared; magnitude assessed at a few points.

c.
Per-cycle percentile metrics; hotspot location error measured; uncertainty bounds propagated; bias and RMSE reported.

d.
Multi-plane WSS fields; cycle-resolved analysis; uncertainty-deconvolved comparisons; cross-condition consistency demonstrated.

## 10. Relevance of the validation activities to the context of use

To ensure that the bench comparisons bear directly on the intended use, the mock loop and protocols were designed to mirror the inflow geometry and operating regimes of the LVAD under study. The cannula lip bevel angle and radius, the bell-mouth curvature, and the inlet housing diameter match the device within machining tolerances of 0.1 mm measured with coordinate metrology. The working fluid's viscosity–shear rate relation was calibrated to approximate blood at hematocrit 35–45 percent and 37 C. The impedance of the afterload module was tuned to reproduce physiologic pressure–flow relations over the range of mean flows considered, including comparable pulse pressures and duty factors. Inflow turbulence intensity at the ventricular segment was controlled via upstream screens to match levels measured in a separate in vivo swine model at the device implant site. Surfaces were prepared to match the device's substrate wettability and roughness within practical limits.

The operating conditions spanned the central 60–120 percent of the device's intended Reynolds number range for the inflow and covered commonly encountered heart rates. Hemodynamic variability was incorporated by testing at different flows and pulse pressures. Not all patient-relevant conditions can be reproduced in bench settings; particular omissions include pathological ventricular flow patterns and extreme hematocrit values. Nevertheless, the ensemble of conditions probed here covers the nominal design envelope and therefore addresses the most likely operating conditions for the LVAD inflow. In addition, the mock loop allows repeatable measurements that reduce confounding factors.

Spatial registration between the bench images and computational geometry was performed using fiducial markers embedded in the transparent housing and by reconstructing the geometry from micro-CT scans of the printed model. Registration errors were below 0.3 mm in regions used for WSS and deposition comparisons. The velocity field planes were chosen to intersect clinically relevant features and to avoid optical distortions. While there are differences between the mock loop materials and the clinical device, notably the absence of long-term protein adsorption effects and cellular responses, the short-term deposition protocol and its thresholds are calibrated against prior studies that used the same surface chemistry as the device, increasing transferability of the findings.

Given the mechanism-specific needs, relevance was judged most stringent for deposition comparisons, where surface chemistry and micro-scale topography modulate outcomes. Here, we ensured that surface energy, wettability, and roughness were within defined bounds of the device in regions of comparison. For high-shear activation and stasis, the primary dependencies are on geometry and flow waveform, which were closely matched. The choice of measurement modalities was also guided by mechanism relevance, with PIV for WSS, dye washout for residence time, and fluorescent platelet imaging for deposition.
*Relevance to use — gradation.*

a.
Generic experiments not matching geometry or operating range; weak traceability.

b.
Partial geometric or operating point overlap; qualitative mapping to use; limited material matching.

c.
Close match in geometry and operating range; mechanism-specific observables; partial material matching.

d.
High-fidelity geometric and operating match; mechanism-

tuned observables; material properties aligned; registration quantified.

## 11. Software quality assurance

The computational workflows operate under a software engineering framework that emphasizes reproducibility, traceability, and robustness. All solver and pre-processor code is maintained in a version-controlled repository with branch protections and code review required prior to merging. Continuous integration pipelines run on each commit, executing unit tests for core numerical kernels, solvers, and the particle tracker. Daily builds execute a larger regression suite that includes canonical flow cases, the manufactured solution tests, and selected end-to-end runs on simplified inflow geometries to detect integration issues. Test coverage is measured and reported to developers, with a target above 90 percent for numerical kernels and above 85 percent for the rest of the code base.

Simulation environments are containerized with pinned compiler versions, math libraries, and MPI stacks to prevent environmental drift across clusters. Meshing and post-processing scripts are likewise versioned and containerized. Data provenance is tracked via manifest files that capture code versions, mesh hashes, boundary condition files, and runtime parameters. A project-specific registry maps each simulation to its inputs and derived outputs, ensuring that results included in comparisons can be reproduced by an independent user. For high-performance runs such as the LES, resilience to hardware faults is provided by periodic checkpoints with validation of checkpoint integrity.

The team operates standard operating procedures (SOPs) for meshing, boundary condition specification, and validation comparisons, with checklists to reduce human error. Peer reviews of major simulation set-ups are documented, and paired analysts cross-check each other's scripts for data extraction. Issues raised by failed tests or anomalies in results are tracked in an issue management system with documented resolution. The adoption of containers also extends to the plot-generation scripts that build figures and tables for this paper, providing an audit trail from raw data to reported metrics.

Although the software engineering practices are mature, limitations do exist, particularly around portability to novel architectures and stresses induced by very large problem sizes. For example, while the LES code can run on GPU-accelerated clusters, the most thoroughly tested path remains the CPU-based builds. Nonetheless, the existing infrastructure is adequate to support the analyses and comparisons reported here with high confidence in reproducibility and to minimize inadvertent errors due to environment or workflow inconsistencies.

*Software quality assurance — gradation.*

a.
Ad hoc scripts; no version control; results not reproducible; no tests.

b.
Basic version control; a few unit tests; limited documentation; manual environments.

c.
Comprehensive version control; CI with unit and regression tests; containerized environments; SOPs for key workflows.

d.
High-coverage tests; automated provenance; fault-tolerant runs; independent reproduction of results demonstrated.

## 12. Operational considerations

The practical use of the models encompasses how analysts construct boundary conditions, choose operating points, and interpret outputs in the context of design and hazard analysis. For the RANS approach, its speed enables exploration of geometric variants and operating points that helps screen for potential hotspots before engaging more expensive simulations or bench cycles. The primary operational challenge is the discipline to avoid over-interpreting steady results in situations where unsteadiness is known to be influential. For the LES approach, the computational cost is substantial: a baseline run over eight cardiac cycles on the 128 million cell grid with 200,000 particles per cycle requires approximately 240,000 core-hours on a 3 GHz CPU cluster. This cost necessitates careful selection of operating points and targeted use when the questions demand it. Turnaround time is thus a factor in planning design iterations.

Post-processing scripts and visualization standards are critical to avoid misinterpretation. Hotspot maps, percentile metrics, and residence-time distributions are reported with clearly labeled units and uncertainties. Spatial registration between model and bench reference frames is repeated for each comparison batch to ensure consistency. Analysts are trained to report both central tendencies and tails of distributions, since tails are often decisive for thrombogenic mechanisms. The team maintains checklists for set-up and review to ensure that workflows are executed consistently across analysts and projects.

User workflow can influence the robustness of conclusions through choices in meshing, filtering of unsteady data, or how surrogate metrics are calculated. Training materials and SOPs are designed to minimize variability, but human factors—such as inconsistent masking of wall regions for WSS extraction or misplacement of dye washout interrogation boxes—remain potential sources of discrepancy across teams and time. These risks are mitigated by automation where feasible, paired analyses, and scripted pipelines that enforce consistent choices.

*Mesh and time-step convergence - residence-time focus — gradation.*

a.
Single-cycle analysis; no tails reported; residence proxies uncalibrated.

b.
Multi-cycle averages without sensitivity; percentiles reported; limited tail stability checks.

c.
Multi-level space–time refinement; tail percentiles stable

across cycles; smoothing used to reduce thresholding artifacts.

d.

Ensemble-of-ensembles analysis; tail stability documented across operating points; sensitivity of thresholds quantified.

## 13. Credibility scores by model and mechanism

This section consolidates the evidence into numeric levels on the private 0–12 scale for each model–mechanism pair and each evidence category. For each category, we provide a brief rationale and the key numbers that anchor the level. The model-specific summary tables present the same information in compact form. The paragraphs here contain the sole statements of findings for each category, placed away from the sections where the activities were described, in keeping with the reporting discipline of this assessment.

Numerical tests for Inflow-RANS were judged robust. In manufactured-solution tests for coupled pressure–velocity and SST closure fields, the observed L2 error convergence rates were 1.96 for velocity, 1.91 for pressure, 1.88 for k, and 1.90 for omega across three grids. The residual norms decreased by at least seven orders of magnitude, and discrete mass conservation errors remained below 1e-12 of mass flux across all grids. The largest remaining bias occurs in corner cells at the bell-mouth, where curvature-induced skewness reduces local order to around 1.6, but integrated outputs showed no sensitivity to this effect within the 0.5 percent level. We assign a level of 8/12 for this category across all three mechanisms for Inflow-RANS on the basis of these numbers.

For Inflow-LES-PaRT, the code-oriented tests showed very strong performance. In turbulent channel flow at Re_tau=180, the inner-scaled mean velocity deviated by 2.3 percent at y-plus of 15 and by less than 1.5 percent across most of the log-layer; the Reynolds shear stress peaked within 3 percent of the DNS reference. The energy spectra displayed the expected -5/3 inertial range. A manufactured scalar-transport test in the presence of non-Newtonian viscosity yielded L2 errors below 0.1 percent at the finest grid. The particle tracker conserved particle count to within 0.02 percent over 100,000 time steps in a closed domain, and round-trip time-reversal tests returned particles to within 0.1 percent of their origin. We therefore assign a level of 9/12 for this category across mechanisms for Inflow-LES-PaRT.

Sensitivity to spatial and temporal resolution for Inflow-RANS was moderate, with outputs stabilizing across three grids in most metrics. For high-shear activation, the three-grid study at 4.5 L/min yielded an observed order of 1.92 for the 99th percentile wall shear stress around the lip and a Richardson-extrapolated uncertainty of 2.8 percent relative to the finest grid value. The location of the WSS hotspot moved by less than 0.4 mm between grids, within the 0.3 mm spatial registration accuracy of bench measurements. For stasis, the 95th percentile of the Eulerian residence-time surrogate in the primary recirculation zone changed by 4.6 percent between the finest two grids, with a grid convergence index (GCI) of 5.1 percent. For deposition-relevant near-wall scalar gradients, sensitivity was larger; the near-wall gradient magnitude in the bell-mouth interior near the blend radius exhibited a 7.4 percent uncertainty under refinement, with y-plus values of 1.6 in that region. We assign levels of 8/12 for high-shear activation, 7/12 for stasis, and 6/12 for surface-mediated deposition for the Inflow-RANS model on the basis of these sensitivities.

Inflow-LES-PaRT showed strong but not negligible sensitivity, as expected in wall-resolved unsteady simulations. For high-shear activation, using 62, 128, and 210 million control volumes with time steps of 2e-5, 1e-5, and 5e-6 s respectively, the 99th percentile wall shear stress per cycle changed by 4.1 percent from coarse to medium and 2.4 percent from medium to fine; the observed order was approximately 1.7 in space–time combined. The y-plus maxima remained below 0.9 over the inflow lip in all grids. For stasis, the 95th percentile of Lagrangian particle residence time varied by 7.5 percent coarse-to-medium and 4.8 percent medium-to-fine. For deposition propensity, the fraction of wall area classified as persistently low shear (below 1 Pa for at least 0.5 s per cycle) changed by 8.6 percent coarse-to-medium and 5.9 percent medium-to-fine; particle-based deposition flags within a 1 mm band of the wall shifted by 6.1 percent between the medium and fine runs. From these results, we assign levels of 9/12 for high-shear activation, 8/12 for stasis, and 8/12 for surface-mediated deposition for Inflow-LES-PaRT.

The mapping from equations to mechanistic needs supported moderate levels for the RANS approach and higher levels for the LES-based approach. For Inflow-RANS, the SST closure with shear-thinning viscosity aligned reasonably with the need to localize shear hotspots: the peak WSS along the lip and the leading inner edge of the bell-mouth are geometry-driven, and the mean stress field positions these reliably. However, representation of unsteadiness is absent and intermittent separation effects are not captured; as a result, pathlines through low-speed regions miss variability important for the stasis question. For surface precipitation, the steady field does not provide a time-resolved picture of low-shear persistence. We therefore assign levels of 7/12 for high-shear activation, 5/12 for stasis, and 4/12 for surface-mediated deposition for the RANS model on adequacy of the formulation.

For Inflow-LES-PaRT, the wall-resolved nature of the simulation and the Lagrangian particle framework render a close match to mechanistic needs. In the high-shear regime, the temporal dynamics of acceleration and deceleration phases are resolved, and the particle stress histories span the relevant range from 2 to 200 Pa*s per cycle near the lip in the 4.5 L/min condition. For stasis, coherent vortical structures adjacent to the bell-mouth are resolved, and their entrainment of particles into and out of the near-wall low-speed zones is captured. For deposition, the use of a combined shear–time criterion for near-wall persistence provides an interpretable surrogate for adhesive stabilization, though the absence of explicit receptor kinetics remains a limitation. On this basis, we assign levels of 9/12 for high-shear activation, 8/12 for stasis, and 8/12 for surface-mediated deposition to Inflow-LES-PaRT for model adequacy.

Quantitative comparison to bench data yields differentiated performance. For Inflow-RANS in high-shear activation, the PIV-derived WSS hotspots at 4.5 L/min showed a root-mean-square error of 18 percent and a mean bias of -6 percent compared to simulation for the 99th percentile per cycle, with hotspot location errors below 0.6 mm. Across 3 and 6 L/min, RMSE ranged 16–21 percent with similar bias. For stasis, dye washout T90 values in interrogation regions adjacent to the bell-mouth differed by 23 percent on average, and the Kolmogorov–Smirnov (KS) distance between the model's Eulerian residence-time distribution and the experimental decay curves was 0.22 at 4.5 L/min, with similar values at the flanking flow rates. For surface-mediated deposition, the spatial correlation (Pearson r) between model low-shear persistence maps and deposition density was 0.42 at 4.5 L/min, with a best-fit slope of 0.55 and an AUC of 0.68 for discriminating deposit versus no-deposit. These comparisons motivate levels of 7/12 for activation, 6/12 for stasis, and 4/12 for deposition for Inflow-RANS.

For Inflow-LES-PaRT, the high-shear activation comparisons show stronger agreement: PIV-derived WSS hotspots yielded an RMSE of 9.7 percent and bias of -2.1 percent for the 99th percentile per cycle at 4.5 L/min, with hotspot location errors under 0.6 mm. At 3 and 6 L/min, RMSEs were 11.2 and 8.9 percent, respectively. For stasis, dye washout T90 values differed by 12 percent on average across interrogation regions at 4.5 L/min, and the KS distance between particle-based residence-time distributions and experimental decay curves was 0.18; similar agreement was seen at the other flow rates. For deposition, the spatial correlation between model-inferred deposition propensity (from the combined shear–time criterion) and measured deposition density was 0.71 with a slope of 0.94 at 4.5 L/min, and the AUC for the deposit mask was 0.86. These metrics support levels of 9/12 for activation, 8/12 for stasis, and 8/12 for deposition for Inflow-LES-PaRT in output comparison.

The match between the bench tests and the intended use is an important aspect of credibility. The mock loop geometry matches the device within 0.1 mm on key inflow features, the operating Reynolds numbers span 60–120 percent of the nominal design range, and hematocrit equivalence is achieved through viscosity matching in the working fluid to the 35–45 percent range. For Inflow-RANS, we emphasize that these conditions are directly relevant to what the model is used to answer, especially for hotspot identification and shear quantiles; consequently, we assign 9/12 for high-shear activation, 8/12 for stasis, and 7/12 for deposition on relevance. The deposition score is slightly lower to reflect the challenges of precisely matching surface chemistry over long durations, even though short-term wettability and roughness were aligned within 3 degrees and 10 percent Ra, respectively.

For Inflow-LES-PaRT, the same mock loop conditions provide arguably an even tighter match to the questions asked of the model, because unsteadiness and near-wall dynamics are central in both model and experiment. The particle-based deposition propensity can be registered to the deposition images, and the residence-time distributions are probed by dye washout in the same regions where particles linger. Given this alignment, we assign 9/12 for all three mechanisms on relevance for Inflow-LES-PaRT.

The software engineering environment in which these models are developed and executed is common to both and confers strong reproducibility. The code base has 164 unit and regression tests run on continuous integration with 94 percent coverage reported for numerical kernels and 88 percent overall. Containers pin compilers and libraries, and manifests ensure data provenance. We therefore assign 10/12 for both models across all mechanisms for this category.

Combining the category levels into an overall index for each model–mechanism pair provides a compact sense of readiness. For Inflow-RANS, averaging the six category scores (with equal weights, as the stakes are similar across categories for design-stage hazard analysis) yields overall indices of 8.2/12 for high-shear activation, 7.3/12 for stasis, and 6.0/12 for surface-mediated deposition. For Inflow-LES-PaRT, the corresponding indices are 9.3/12, 8.3/12, and 8.5/12. These numbers are not substitutes for the detailed tables and narrative but offer a summary view consistent with the mechanism-specific evidence.

Finally, a brief word on human workflows is warranted. Interpretation of hotspot maps, configuration of dye washout interrogation regions, and consistent application of deposition thresholds in image analysis are susceptible to variability across analysts. While SOPs and automation reduce this risk, vigilance is required to maintain consistency as staff and tools evolve.
*Output comparison - stasis-focused rubric — gradation.*

a.
No residence comparison; qualitative flow recirculation noted; no time metrics.

b.
Single-point washout or tracer test; limited regions; no distributional comparison.

c.
Regional T90 and percentile metrics; distributional tests (e.g., KS) reported; spatial registration quantified.

d.
Multi-region, multi-condition washout with uncertainty; tail behavior assessed; mapping to particle-based distributions demonstrated.

## 14. Credibility assessment summary for Inflow-RANS

The following table consolidates the credibility levels for Inflow-RANS across the three thrombogenic mechanisms. Each row provides the evidence category, the numeric level on the private 0–12 scale, and a succinct basis that is also documented in the narrative. The table is intended for quick reference; the surrounding sections provide the detailed context and interpretation. As emphasized above, the steady RANS approach demonstrates consistent strengths in hotspot localization and magnitude estimation for high-shear activation, with qualifications for stasis and deposition where unsteadiness and near-wall dynamics are central.

## 15. Credibility assessment summary for Inflow-LES-PaRT

The following table consolidates the credibility levels for Inflow-LES-PaRT across the three thrombogenic mechanisms. The wall-resolved unsteady formulation and Lagrangian particle framework align closely with the mechanistic needs for high-shear activation, stasis, and deposition analyses. The table reports the levels and the succinct bases, with detailed numerical evidence and interpretation provided in the text. Computational expense and complexity are acknowledged elsewhere; the table focuses on mechanism-specific credibility elements.

## 16. Discussion

The credibility assessment presented here offers an example of how to operationalize mechanism-specific evaluations under the V&V40 philosophy. In VAD inflow analyses, the inability of any single model to fully capture all mechanisms across all operating regimes necessitates a discriminating approach. By focusing on three mechanistic questions—high-shear activation, stasis, and surface-mediated deposition—we were able to isolate what each model did well and where it struggled, and to do so in a way that maps to decision-relevant metrics.

For high-shear activation, both models provide useful information. The RANS approach localizes hotspots at the cannula lip and bell-mouth leading edge with high consistency, provides reasonable magnitude estimates of peak wall shear stress, and achieves quantitative agreement with PIV-based WSS percentiles within approximately 20 percent RMSE across operating points. The LES-based approach improves agreement to under 12 percent RMSE and aligns hotspot locations within sub-millimeter error. The time-resolved nature of the LES is arguably more true to the physics of platelet exposure during acceleration and deceleration phases. However, at the design-screening stage where the objective is to identify and rank hotspots across variants, the RANS approach may offer a better speed–fidelity balance.

For stasis, the strengths of the LES approach are more pronounced. Steady models lack the ability to capture intermittent reattachment, vortex shedding, and the entrainment dynamics that control how fluid lingers near the inflow surfaces. The washout comparisons corroborate this: dye-based T90 values match the LES results within roughly 12 percent on average, whereas RANS-based surrogates differ by about 23 percent. The KS distance between model-predicted distributions and measured decay curves similarly favors LES. Although the RANS model remains a useful tool for flagging obvious recirculation zones and for relative comparisons across designs, its use for quantitative statements about residence-time tails should be qualified.

For surface-mediated deposition, both the physics and the measurement challenges are substantial. The LES-based approach, augmented with a combined shear–time criterion and Lagrangian particle tracking, offers a mechanistic bridge from near-wall dynamics to deposition propensity. The resulting maps correlate with measured deposition densities with a coefficient near 0.7 and a slope close to unity in our experiments, and the model discriminates deposit from no-deposit regions with an AUC of 0.86. The RANS approach, limited to Eulerian proxies for low-shear persistence, correlates more weakly and exhibits a slope under unity, indicating systematic underestimation of deposition density variations. Neither approach includes biochemical adhesion kinetics, which likely limits absolute predictive power; yet, the LES-derived proxies appear sufficiently discriminating to assist in design-stage prioritization of surfaces for material or geometric modification.

The numerical tests and sensitivity studies anchor the trust in computed quantities before any experimental comparison is drawn. The code-oriented tests show near-second-order convergence in manufactured solutions in the steady solver and satisfactory reproduction of turbulent channel benchmarks in the unsteady solver. Sensitivity of key outputs to mesh and time-step refinement is mostly in the low single digits for activation metrics and larger but controlled for stasis and deposition metrics. These numbers give context for interpreting differences between model and experiment and support the assigned credibility levels.

Relevance of the experimental activities to the intended use emerges as a vital contributor to trust. The deliberate design of the mock loop to match geometry, flow conditions, and surface properties within defined tolerances ensures that disagreement between model and experiment is more likely due to model limitations than to apples-to-oranges comparisons. When paired with software engineering practices that safeguard reproducibility and traceability, the assessment provides a coherent picture of what the models can and cannot be trusted to do.

The scoring itself deserves comment. A private 0–12 scale allows tailoring granularity to the decision context, but it also introduces the burden of explicating what each level means for each category. Our rubrics attempted to make the mapping transparent. The overall indices offer a digestible summary, but their value lies primarily in communicating that the LES-based approach reaches high single-digit scores across mechanisms, whereas the RANS approach lags particularly for deposition and, to a lesser degree, stasis. Importantly, the context of use is design-stage exploration and hazard analysis; a different context—such as a regulatory submission intending to demonstrate safety margins quantitatively—would likely demand higher levels in some categories or additional activities such as uncertainty propagation through to a risk metric.

There are limitations that temper the conclusions. Bench measurements themselves carry uncertainties and practical deviations from the device, particularly in long-term surface chemistry effects. The Lagrangian approach uses surrogate criteria for deposition, which cannot replicate the complex biology of platelet adhesion and stabilization. Blood is modeled with a shear-thinning viscosity but without explicit particulate phases; this choice ignores RBC-induced margination, which could alter near-wall platelet concentrations and deposition rates. Some of these limitations can be mitigated in future work by adding richer hemorheology, incorporating adhesion kinetics calibrated against targeted assays, and expanding the experimental program to include different surface chemistries and longer durations.

Finally, the assessment shows the importance of sustained attention to operational details, even when high-quality models and data are available. Standardized post-processing, rigorous spatial registration, and disciplined interpretation of percentile metrics prevent over- or under-stating performance. User workflows remain a potential source of variability, underlining the value of SOPs and automation. If this assessment is to serve as a template, it should encourage teams to think with similar specificity about which mechanisms their questions hinge on and how to assemble targeted, balanced evidence to address them.

*Aggregation and reporting discipline — gradation.*

a.
Ad hoc scoring; no rubrics; findings mixed with methods; repeated or conflicting statements.

b.
Basic rubrics; some mixing of methods and findings; limited transparency in mapping to scores.

c.
Mechanism-specific rubrics; single-location findings; tables aligned to narrative; private scale explained.

d.
Cross-checked rubrics; independent review; uncertainty integration; traceable mapping from evidence to scores.

## 17. Conclusions

We presented a mechanism-specific credibility assessment under the V&V40 philosophy for two CFD approaches to analyzing thrombogenic risk in the inflow region of an LVAD. The steady RANS model with SST closure and non-Newtonian rheology is credible for localizing and ranking high-shear hotspots that condition platelet activation, with quantitative agreement to bench WSS metrics within approximately 20 percent RMSE and moderate sensitivity to mesh refinement. Its use for quantitative statements about residence-time tails and deposition propensity should be qualified, as it lacks unsteady dynamics central to these mechanisms. The wall-resolved LES coupled to Lagrangian platelet surrogates attains higher credibility levels across all three mechanisms, with stronger agreement to washout and deposition data and space–time sensitivity in the low single digits for key metrics. This performance comes with increased computational cost, demanding careful targeting to the most consequential questions.

The private 0–12 scale used here differentiates performance across mechanisms and models in a way intended to inform design-stage hazard analysis. Across the six evidence categories, Inflow-RANS levels for high-shear activation cluster around 7–10, for stasis around 6–10 with modestly lower values in model adequacy and comparison, and for deposition around 4–10 with the lowest values in model adequacy and comparison. Inflow-LES-PaRT levels are 8–10 across mechanisms, with the lower end driven by the absence of explicit adhesion kinetics and residual sensitivity in residence-time tails. Overall indices summarize these distinctions at 8.2/12, 7.3/12, and 6.0/12 for RANS across activation, stasis, and deposition, and 9.3/12, 8.3/12, and 8.5/12 for LES, respectively.

For practitioners, the main practical guidance is: use the steady RANS approach to efficiently survey designs for high-shear hotspots and to estimate margins to activation thresholds, and deploy the wall-resolved LES with particles for focused questions about residence-time distributions and deposition propensity where design decisions could hinge on resolving unsteady near-wall features. For sponsors and reviewers considering evidence packages, this assessment illustrates how to weight mechanistic adequacy and targeted comparisons for inflow thrombogenicity questions, and underscores the importance of disciplined software practices and relevance of validation activities. The methods and rubrics reported here can be adapted to other device regions or mechanisms, provided that the context of use and the mechanistic sensitivities are carefully reconsidered.

Future work should expand the mechanistic link between flow and biology by including RBC-induced margination models, explicit adhesion receptor kinetics, and coupled thrombin generation where feasible. Further, adding operating conditions that probe the edges of the design envelope and surfaces with different chemistries will enhance relevance and stress-test the models. Finally, formal uncertainty propagation from numerical and model-form sources through to mechanism-specific risk metrics will make the linkage to design decisions even tighter and more defensible.

*Model form - activation-focused rubric — gradation.*

a.
Shear magnitude only; no exposure time; uses generic thresholds without justification.

b.
Exposure-time proxy included; thresholds from literature; limited sensitivity analysis.

c.
Trajectory-based exposure computed; thresholds tied to in vitro activation curves; phase-specific dynamics considered.

d.
Full stress-history distributions computed; mapping to activation probability calibrated; uncertainty reported.

## References

[1] D. Halvorsen, D. Zetterberg, L. Ueda, Gap size effects on hemolysis index in computational models, J. Verif. Valid. Uncertain. 62 (2020) 146–758.

[2] B. Okonkwo, A. Rasmussen, W. Duarte, C. Tanaka, S. Silva, Material model effects on peak force in computational models, Int. J. Numer. Methods 110 (2021) 313–634.

[3] C. Duarte, B. Bergman, L. Wojcik, C. Moreau, C. Rasmussen, Wall roughness effects on rib deflection in computational models, Med. Eng. Phys. 106 (2013) 83–402.

[4] J. Zetterberg, G. Wojcik, Flow rate effects on chest compression in computational models, J. Mech. Behav. Biomed. Mater. 66 (2015) 304–466.

[5] A. Pettersson, R. Ishikawa, H. Okonkwo, Element type effects on chest compression in computational models, J. Mech. Behav. Biomed. Mater. 20 (2006) 248–634.

[6] G. Quintero, K. Moreau, Element type effects on rib deflection in computational models, Med. Eng. Phys. 20 (2019) 76–847.

[7] E. Grigoryan, W. Yilmaz, P. Rasmussen, S. Ishikawa, N. Rasmussen, Boundary condition effects on rib deflection in computational models, J. Biomech. Eng. 41 (2016) 379–471.

[8] J. Ueda, P. Fontaine, E. Jankowski, J. Yilmaz, E. Rasmussen, Material model effects on stress range in computational models, Int. J. Numer. Methods 126 (2013) 238–623.

[9] F. Castellano, L. Okonkwo, C. Quintero, Damping effects on hemolysis index in computational models, Int. J. Numer. Methods 51 (2008) 314–783.

[10] T. Grigoryan, N. Nakamura, E. Nakamura, N. Kaur, W. Silva, Time step effects on shear stress in computational models, Int. J. Numer. Methods 32 (2024) 304–557.

[11] D. Lindqvist, B. Castellano, A. Tanaka, B. Halvorsen, Friction effects on hemolysis index in computational models, J. Verif. Valid. Uncertain. 31 (2012) 241–589.

[12] F. Ueda, L. Fontaine, A. Ishikawa, P. Yilmaz, Flow rate effects on flow split in computational models, Med. Eng. Phys. 108 (2016) 299–637.

[13] S. Wojcik, R. Ishikawa, H. Wojcik, D. Wojcik, W. Quintero, Contact stiffness effects on fatigue life in computational models, Int. J. Numer. Methods 117 (2023) 175–812.

[14] B. Ueda, S. Castellano, Gap size effects on rib deflection in computational models, Ann. Biomed. Eng. 107 (2019) 162–408.

[15] D. Lindqvist, H. Zetterberg, S. Castellano, Time step effects on peak force in computational models, Comput. Methods Biomech. 133 (2018) 94–821.

[16] K. Okonkwo, C. Silva, Yield stress effects on head displacement in computational models, Comput. Fluids 144 (2007) 275–555.

[17] T. Zetterberg, S. Silva, K. Fontaine, Bone density effects on flow split in computational models, Med. Eng. Phys. 127 (2005) 334–676.

[18] S. Silva, P. Ishikawa, C. Halvorsen, D. Fontaine, R. Grigoryan, Material model effects on stress range in computational models, Comput. Methods Biomech. 29 (2017) 358–778.

[19] R. Duarte, K. Okonkwo, F. Yilmaz, E. Duarte, Material model effects on stress range in computational models, J. Biomech. Eng. 22 (2012) 206–798.

[20] P. Zetterberg, N. Silva, Cement mantle effects on contact pressure in computational models, Comput. Methods Biomech. 122 (2014) 137–686.

[21] K. Castellano, B. Ishikawa, L. Vasquez, B. Kaur, Boundary condition effects on flow split in computational models, J. Mech. Behav. Biomed. Mater. 65 (2010) 268–591.

[22] N. Quintero, E. Ishikawa, Wall roughness effects on outlet velocity in computational models, J. Mech. Behav. Biomed. Mater. 139 (2017) 183–633.

[23] A. Lindqvist, W. Eriksen, Material model effects on peak force in computational models, Comput. Fluids 145 (2010) 199–699.

[24] R. Zetterberg, B. Grigoryan, Boundary condition effects on hemolysis index in computational models, Med. Eng. Phys. 137 (2019) 173–629.

[25] L. Rasmussen, B. Bergman, A. Eriksen, A. Abbott, B. Eriksen, Cement mantle effects on shear stress in computational models, Int. J. Numer. Methods 43 (2018) 243–699.

[26] E. Pettersson, K. Halvorsen, Cement mantle effects on contact pressure in computational models, J. Verif. Valid. Uncertain. 59 (2009) 216–532.

[27] G. Eriksen, G. Bergman, W. Okonkwo, Flow rate effects on shear stress in computational models, J. Biomech. Eng. 81 (2014) 205–723.

[28] T. Moreau, C. Quintero, Cement mantle effects on peak force in computational models, Comput. Fluids 119 (2013) 389–616.

[29] G. Grigoryan, C. Pettersson, B. Tanaka, N. Vasquez, Inlet profile effects on outlet velocity in computational models, Ann. Biomed. Eng. 133 (2024) 213–709.

[30] K. Silva, D. Moreau, K. Okonkwo, R. Rasmussen, K. Quintero, Inlet profile effects on hemolysis index in computational models, J. Verif. Valid. Uncertain. 89 (2023) 343–586.

[31] N. Pettersson, E. Silva, H. Quintero, Contact stiffness effects on shear stress in computational models, Med. Eng. Phys. 102 (2019) 311–429.

[32] B. Quintero, M. Okonkwo, P. Yilmaz, H. Okonkwo, Cement mantle effects on outlet velocity in computational models, Comput. Fluids 121 (2005) 254–839.

[33] M. Silva, N. Yilmaz, B. Lindqvist, T. Rasmussen, P. Ueda, Inlet profile effects on peak force in computational models, Comput. Fluids 32 (2010) 363–829.

[34] H. Lindqvist, T. Ishikawa, T. Tanaka, D. Castellano, T. Abbott, Flow rate effects on outlet velocity in computational models, J. Biomech. Eng. 102 (2017) 288–653.

Table 1: Credibility factors for Inflow-RANS by thrombogenic mechanism on a private 0–12 scale.

| Credibility factor | Level | Basis |
|---|---|---|
| Numerical code verification - Inflow-RANS - High-shear platelet activation | 8/12 | MMS observed orders ~1.9 for velocity/pressure/k/omega; residuals reduced 7–8 orders; mass conservation error <1e-12. |
| Numerical code verification - Inflow-RANS - Recirculation-driven stasis | 8/12 | Same solver tests apply; corner-cell skewness lowers local order to ~1.6 but integrated outputs stable within 0.5%. |
| Numerical code verification - Inflow-RANS - Surface-mediated deposition | 8/12 | Manufactured turbulence variables converged near second order; no defects in flux reconstruction in bell-mouth geometry. |
| Discretization error - Inflow-RANS - High-shear platelet activation | 8/12 | Three-grid study: observed order 1.92; Richardson-extrapolated WSS(99th) uncertainty 2.8%; hotspot shift <0.4 mm. |
| Discretization error - Inflow-RANS - Recirculation-driven stasis | 7/12 | 95th percentile Eulerian RT changed 4.6% (fine vs medium); GCI 5.1% in primary recirculation zone. |
| Discretization error - Inflow-RANS - Surface-mediated deposition | 6/12 | Near-wall scalar gradient uncertainty 7.4%; wall y+ ~1.6 at blend radius; larger sensitivity in low-shear persistence proxies. |
| Model form - Inflow-RANS - High-shear platelet activation | 7/12 | SST with shear-thinning adequately localizes shear hotspots; steady view acceptable for geometry-driven peak WSS. |
| Model form - Inflow-RANS - Recirculation-driven stasis | 5/12 | Quasi-steady approach cannot capture intermittent separation and reattachment; residence tails under-resolved. |
| Model form - Inflow-RANS - Surface-mediated deposition | 4/12 | Time-resolved low-shear persistence not represented; no particle–wall encounter modeling; surrogate metrics only. |
| Output comparison - Inflow-RANS - High-shear platelet activation | 7/12 | PIV WSS(99th) RMSE 18%, bias -6% at 4.5 L/min; hotspot location error <0.6 mm; similar at 3 and 6 L/min. |
| Output comparison - Inflow-RANS - Recirculation-driven stasis | 6/12 | Dye washout T90 mean difference 23%; KS distance 0.22 between model RT and experimental decay curves. |
| Output comparison - Inflow-RANS - Surface-mediated deposition | 4/12 | Deposition correlation r=0.42; slope 0.55; AUC 0.68 at 4.5 L/min with registered fields. |
| Relevance of the validation activities to the COU - Inflow-RANS - High-shear platelet activation | 9/12 | Geometry matched within 0.1 mm; Reynolds numbers 60–120% of intended; hematocrit-equivalent viscosity 35–45%. |
| Relevance of the validation activities to the COU - Inflow-RANS - | 8/12 | Loop pressure and pulsatility matched within 8% and 0.06 duty factor; inflow turbulence intensity controlled. |

Table 2: Credibility factors for Inflow-LES-PaRT by thrombogenic mechanism on a private 0–12 scale.

| Credibility factor | Level | Basis |
|---|---|---|
| Numerical code verification - Inflow-LES-PaRT - High-shear platelet activation | 9/12 | Channel flow Re_tau=180 u+ within 2.3%; Reynolds shear stress within 3%; MMS scalar L2 error <0.1%; particle count conserved to 0.02%. |
| Numerical code verification - Inflow-LES-PaRT - Recirculation-driven stasis | 9/12 | WALE model verified; energy spectra show -5/3 slope; time-reversal tests return particles within 0.1% of origin. |
| Numerical code verification - Inflow-LES-PaRT - Surface-mediated deposition | 9/12 | Near-wall resolution y+<1; scalar transport and viscosity coupling verified; no defects detected in wall shear computation. |
| Discretization error - Inflow-LES-PaRT - High-shear platelet activation | 9/12 | 128M grid y+ max 0.9; WSS(99th) changed 2.4% medium-to-fine; observed combined space–time order ~1.7. |
| Discretization error - Inflow-LES-PaRT - Recirculation-driven stasis | 8/12 | 95th percentile particle residence time changed 4.8% medium-to-fine; ensemble-averaged over 5 cycles. |
| Discretization error - Inflow-LES-PaRT - Surface-mediated deposition | 8/12 | Low-shear wall area fraction changed 5.9% medium-to-fine; deposition flags within 1 mm band shifted 6.1%. |
| Model form - Inflow-LES-PaRT - High-shear platelet activation | 9/12 | Wall-resolved unsteadiness captures acceleration-phase peaks; particle stress integrals span 2–200 Pa*s near lip. |
| Model form - Inflow-LES-PaRT - Recirculation-driven stasis | 8/12 | Coherent vortices resolved; entrainment into/out of low-speed zones captured; RBC aggregation not modeled. |
| Model form - Inflow-LES-PaRT - Surface-mediated deposition | 8/12 | Combined shear–time near-wall criterion implemented; no receptor kinetics; deposition propensity is a calibrated surrogate. |
| Output comparison - Inflow-LES-PaRT - High-shear platelet activation | 9/12 | PIV WSS(99th) RMSE 9.7%, bias -2.1% at 4.5 L/min; hotspot location error <0.6 mm; similar at 3 and 6 L/min. |
| Output comparison - Inflow-LES-PaRT - Recirculation-driven stasis | 8/12 | Dye washout T90 mean difference 12%; KS distance 0.18 between particle RT and experimental decay curves. |
| Output comparison - Inflow-LES-PaRT - Surface-mediated deposition | 8/12 | Deposition correlation r=0.71; slope 0.94; AUC 0.86 at 4.5 L/min with registered fields. |
| Relevance of the validation activities to the COU - Inflow-LES-PaRT - High-shear platelet activation | 9/12 | Geometry and operating range matched; unsteady waveform at 75 bpm; viscosity-equivalent hematocrit 35–45%. |
| Relevance of the validation activities to the COU - Inflow-LES-PaRT - Recirculation-driven stasis | 9/12 | Washout regions coincide with modeled recirculation; pressure and pulsatility within 8% and 0.06 duty factor. |