Matthias Leuchter*, Laura Supp, Jan Oldenburg, Michael Stiehm, Klaus-Peter Schmitz and Finja Borowski

In Silico Clinical-Trials: Advancing Validation for Regulatory Approval

https://doi.org/10.1515cdbme-2025-0202

Abstract: The assesment of the safety and efficacy of medical devices can be time-consuming, expensive, and raise ethical concerns. Additionally, the high costs of innovation and lengthy development times may result in increased prices and unequal access for certain patient groups. *In silico* clinical trials (ISCTs) offer an innovative alternative by utilizing computational modeling and simulation to generate clinically relevant data. A recent draft by the FDA suggests the use of clinical data to support model validation. One example of this is a case study that investigates paravalvular leakage in the context of transcatheter aortic valve replacement.

Keywords: ISCT, Credibility, FDA, medical device regulation, regulatory affairs.

1 Introduction

The evaluation of safety and efficacy has traditionally been conducted through experimental methods, including *in vitro*, *ex vivo*, *in vivo* tests on animals, and clinical trials involving humans. There are several compelling reasons to incorporate computational models into the evidence-generation process. *In vitro* and *ex vivo* tests can be time-consuming, and using animals for experimentation and clinical trials raises ethical and translational concerns. Additionally, these tests can be costly and lengthy. The high costs associated with innovation and the lengthy time it takes to bring a product to market can result in increased prices and may lead to under-representation or discrimination based on factors such as gender, age, ethnicity, the rarity of the condition, or socioeconomic status [1]. Regulatory authorities, such as regulatory agencies and notified bodies, require evidence of the safety and efficacy of a new medical product to grant marketing authorization. This approval is essential for any company before their medical product can be sold in a specific country. The capacity of clinical studies acquire data on outlier patient cases or infrequently used device configurations is limited. *In silico* clinical trials (ISCTs) are a valuable source of clinically relevant data, which is obtained via computational modeling and simulation (CM&S) [2]. Technical aspects of an ISCT have been presented, however final regulatory decision on the ISCT application as a clinical trial alternative is still outstanding [3].

Currently, only the American regulatory authority FDA is dealing with the question of how clinical data can support model validation [4]. For CM&S applications, including ISCT, aimed at modeling clinical outcomes, both benchtop testing and clinical evidence are necessary to establish confidence in the modeling approach [5]. To illustrate the current framework for ISCT can guide and evaluate clinical validation activities, a case study is introduced around a high risk ISCT application. The application focuses on ParaValvular Leakage (PVL) in Transcatheter Aortic Valve Replacement (TAVR).

2 Model Credibility

2.1 Existing Regulatory Guidelines

The FDA guidance on Reporting of Computational Modeling Studies in Medical Device Submissions outlined the importance of providing a complete and accurate summary of computational modeling and simulation evidence in a dossier, referring to ASME VV-10:2019, VV-20:2009 (R2021), and VV-40:2018 [6-8]. The 2023 FDA guidance outlines an ASME VV-40:2018-based framework for assessing model credibility, proposing eight possible categories of credibility evidence (Tab.1). According to ASME VV-40:2018, the model's credibility activities are introduced with the formulation of the question of interest, which describes a specific question, decision, or concern addressed by the computational model. Building on this, the Context of Use (CoU) is defined, the objective of which is to explain the role and scope of the model and its relationship to other forms of

***Corresponding author: Matthias Leuchter:** Institute for ImplantTechnology and Biomaterials e.V., Rostock-Warnemünde, Germany, e-mail: matthias.leuchter@iib-ev.de;
Laura Supp, Jan Oldenburg, Michael Stiehm, Klaus-Peter Schmitz, Finja Borowski: Institute for ImplantTechnology and Biomaterials e.V., Rostock-Warnemünde, German

Table 1: Eight categories of credibility evidence. Reprinted from FDA guidance "Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions" [5]. An *in silico* clinical trial where a device safety/effectiveness question is addressed using a virtual cohort of patient models, generated by sampling parameter values across the patient population, should be supported by: code verification results (Category 1), bench test validation results (to validate the device model; Category 3), *in vivo* validation results (to validate the baseline patient model; Category 4), calculation verification results (Category 3, 4 or 8) and population-based validation results (Category 5)

	Category	Definition
1	Code verification results	Results showing that a computational model implemented in software is an accurate implementation of the underlying mathematical model
2	Model calibration evidence	Comparison of model results with the same data used to calibrate model parameters
3	Bench test validation results	Validation results using a bench test comparator. May be supported by calculation verification and/or UQ results using the validation conditions
4	In vivo validation results	Same as previous category except using in vivo data as the comparator
5	Population-based validation results	Comparison of population-level data between model predictions and a clinical data set. No individual-level comparisons are made
6	Emergent model behaviour	Evidence showing that the model reproduces phenomena that are known to occur in the system at the specified conditions but were not pre-specified or explicitly modeled by the governing equations
7	Model plausibility evidence	Rationale supporting the choice of governing equations, model assumptions, and/or input parameters only
8	Calculation verification/ UQ results using CoU evidence	Calculation verification and/or UQ results obtained using the CoU simulations, that is, the simulations performed to answer the question of interest

data, such as *in vitro* or *in vivo* data, to answer the question of interest. In general, the approval is specific for that CoU, and therefore, a new approval must be granted for each new CoU.

Bishopp et. al. proposed three independent factors of 'scope', 'coverage', and 'severity' for model risk assessment. 'Severity' is synonymous with severity of harm in the ISO medical device risk management process [2], representing the consequence to a patient if a device does not function as intended, and is specific to the patient outcome in the particular ISCT modeling application. The 'scope' factor reflects whether ISCT cohorts are outside the bounds of clinical trial data, whereas the proposed 'coverage' factor reflects the amount of data represented exclusively within the ISCT cohorts. While the scope factor captures qualitative differences between ISCT and clinical data, the coverage factor captures quantitative differences. This includes in relation to the investigation of the full intended clinical setting (including device indications and configurations, patient demographics, surgical approaches, etc.) for the device.
The scope reflects whether the ISCT application covers parts of the clinical setting that are not covered by traditional clinical data or whether the scope of the ISCT is entirely within the limits of other clinical data sources. The coverage factor reflects the quantitative impact of broader coverage of clinical variability etc. within the ISCT cohort (represented by anatomical variation of the aortic root area). Quantification of the coverage factor can be supported by various data sources, including literature and commercial data.

2.2 Credibility Targets

CM&S that reproduce the benchtop testing of products rely on benchtop validation activities [9]. The clinical relevance of the model must be ensured by extending the associated benchtop method. Therefore, complementing validation activities with clinical validation activities aimed at increasing confidence in the clinical relevance of the ISCT model is required.
Credibility factors are proposed to evaluate the credibility of the clinical referent data and of the model used to predict these data, to guide evaluation of the agreement between the comparator data and the model results, and to assess the applicability of the clinical validation activities to the intended ISCT application. A range of clinical sources for referent data can be used for establishing the clinical relevance for CM&S.

A variety of clinical sources can be utilized to establish the clinical relevance for CM&S (Clinical Measurement & Safety). These sources can provide data on individual patients, patient cohorts, or even from public resources like journal articles. They can highlight trends or support statistically significant conclusions regarding the differences in clinical performance between various patient groups or device configurations. These studies can be designed prospectively to maximize their utility for CM&S validation, or they can be used retrospectively. Ideally, the data and conclusions should reflect the clinical variability associated with the subject device, and the statistical power of the study's conclusions should be robust. The validity of ISCT applications depends on the representation of the expected range of clinical variability. Thus, a validation model that uses a single average model to represent the anatomy of interest is less credible than modeling a range of anatomies. The former approach can demonstrate that anatomical variability has no influence on the predicted outcome, whereas the latter assumes this outcome. This principle applies to all aspects of the model such as boundary conditions, device configurations or implantation configurations. Generating evidence of the impact of clinical variables on predicted clinical outcomes creates greater credibility.

As part of the evaluation of the validity of the clinical validation, both the congruence of the input conditions (validation inputs) and the clinical results (validation outputs) between the model and the clinical comparator are taken into account. The aim is to achieve agreement in the statistical distribution of the predictions from the *in silico* cohort and a corresponding clinical cohort. The assessment of credibility is also based on comparative data sets that show statistically significant differences between the clinical conditions. This allows accurate and quantitative prediction of outcomes, such as revision rates for specific failure modes. Lower credibility of assessments may mean that there is agreement on the relative performance of revision rates, but no accurate quantitative prediction is possible.

To enhance credibility, the clinical study involving the product of interest should be utilized to validate the ISCT approach for that same product. If there are differences in product characteristics between the product of interest and the product used in the comparison group for validation, it will directly affect the model.

3 ISCT Application

TAVR has emerged as a new standard for percutaneous treatment of severe aortic valve stenosis in patients at intermediate and high surgical risk, or who are inoperable. Nevertheless, its extensive utilisation is presently constrained by the occurrence of postoperative PVL, which portends elevated mortality during follow-up [10]. Aortic valve calcification has been found to be associated with PVL, but results from the available literature have been obtained in studies with limited sample sizes [11].

The CoU for this application is as follows: The ISCT aims to predict the occurrence of PVL by analyzing the variability in the severity and location of calcification within the target patient population. Clinical evidence has shown that the approved TAVR system performs well in anatomical configurations similar to those found in the ISCT patient population. The discussion regarding the risk assessment of the model focuses on its scope and coverage. Previous studies have already addressed the severity of clinical outcomes through simulations with a smaller sample size. A novel aspect of this study is the inclusion of variability in the clinical target environment, particularly with regard to the characteristics of calcification associated with the proposed implant system. The material characteristics of the TAVR reference product were adopted for the validation model using *in vitro* tests, by Finotello et al. [12]. Spanjaards et al. conducted a series of bench tests including a validation approach for the leakage flow simulations [13].

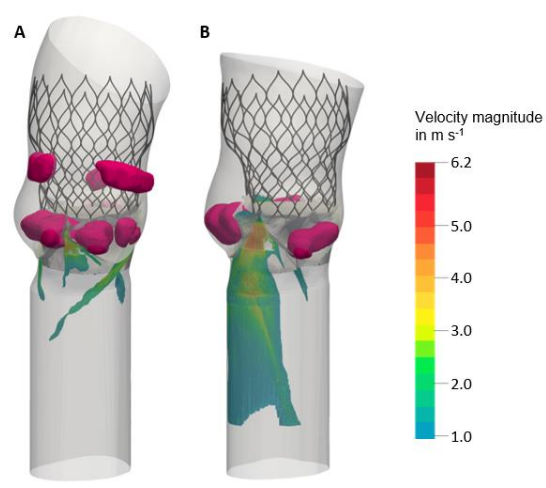

Figure 1: Exemplary flow results of patients with mild (A) and moderate (B) leakage.

The approach involves a non-parametric shape model designed to capture the geometric variations of the aortic root within a real patient cohort using CT imaging. For modeling the calcifications, a method based on the work of Oldenburg et al. was utilized [14]. As part of the clinical validation, a comprehensive collection of key input parameters was completed, covering the full range of expected clinical variability that are relevant to the model and the question of interest. The primary model inputs and comparison conditions are equivalent on a patient-specific basis [15]. The model quantitatively predicts the relative clinical performance of the patient cohort compared to published data with the identical device (24% vs. 16% moderate/severe PVL, p=0.250) [11]. Furthermore, the validation model was able to demonstrate the same statistical correlations as those reported in various published clinical studies using the identical reference product.

The lack of credibility of this model lies in the lack of follow-up of the patient cohort in order to be able to carry out a patient-specific comparison. This would require publicly available validation datasets for a cross-comparison of the predictive accuracy of the *in silico* technologies.

4 Discussion

Recent advancements in ISCT research have been impressive, resulting in several promising applications. However, the anticipated widespread adoption of ISCT is hindered by persistent challenges, including technological barriers, validation issues, regulatory deficiencies, scalability problems, communication difficulties, exploitation concerns, and training deficits [16]. When examining regulatory challenges, it becomes clear that the FDA's pathway for medical devices has proven effective and efficient, while Europe struggles to keep pace. Currently, no established processes exist for employing ISCT in the regulatory approval of medical devices. The establishment of a robust regulatory framework for ISCT applications in medical devices is imperative. The active involvement of regulatory authorities in this process is essential [17]. Moreover, establishing a solid theoretical and methodological foundation for generalizing verification, validation, and uncertainty qualification concepts in *in silico* models and ISCT is essential. Currently, the scarcity of public validation datasets limits accurate cross-comparisons of *in silico* technologies. Initiatives like the European Health Data Space (EHDS) to gather high-quality, diverse datasets from EU member states could greatly enhance the validation process for *in silico* models. Ultimately, selecting reliable and valid assessment methods remains a crucial challenge for researchers, policymakers, and clinicians alike [16].

Author Statement

Research funding: Financial support from the European Unions Horizon 2020 research and innovation program under grant agreement No 101017578. The EU is not responsible for any use that may be made of the results shown in this publication or any information contained thereinConflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study.

References

[1] Viceconti M, Emili L, Afshari P, et al. Possible Contexts of Use for In Silico Trials Methodologies: A Consensus-Based Review. *IEEE J Biomed Health Inform*. 2021;25(10):3977-3982

[2] Bischoff JE, Dharia MA, Favre P. A risk and credibility framework for in silico clinical trials of medical devices. *Comput Methods Programs Biomed*. 2023; 242:107813.

[3] J. Mathew, A. MacLeod, A. Toms, R. et al., Personalising high tibial osteotomy: 'ideal' preclinical stage evaluation to TOKA(R) patient specific 3D printed system, Knee 27 (S8) (202

[4] ASME standard "Assessing Credibility of Computational Modeling through Verification and Validation: Application to Medical Devices", 2018

[5] FDA guidance "Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions", Nov 2023.

[6] FDA guidance "Reporting of Computational Modeling Studies in Medical Device Submissions", Sep 2016.

[7] ASME standard "Standard for Verification and Validation in Computational Solid Mechanics V V 10 - 2019",

[8] ASME standard "Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer V V 20 - 2009 (R2021)",

[9] T.M. Morrison, P. Hariharan, C.M. Funkhouser, P. et al., Assessing computational model credibility using a risk-based framework: application to hemolysis in centrifugal blood pumps, ASAIO J. 65 (4) (2019) 349–360

[10] Laakso T, Laine M, Moriyama N, et al (2020) Impact of paravalvular regurgitation on the mid-term outcome after transcatheter and surgical aortic valve replacement. European Journal of Cardio-Thoracic Surgery 58(6):1145–1152.

[11] Pollari F, Dell'Aquila AM, S¨ohn C, et al (2019) Risk factors for paravalvular leak after transcatheter aortic valve replacement. The Journal of thoracic and cardiovascular 31 surgery 157(4):1406–1415.e3.

[12] Finotello A, Gorla R, Brambilla N, et al., Finite element analysis of transcatheter aortic valve implantation: Insights on the modelling of self-expandable devices. Journal of the Mechanical Behavior of Biomedical Materials 2021;123:104772

[13] Spanjaards M, Borowski F, Supp L, et al (2024) A fast in silico model for preoperative risk assessment of paravalvular leakage. Biomechanics and Modeling in Mechanobiology 23(3):959–985

[14] Oldenburg J, Borowski F, Supp L, et al (2024) Generation of shape models of calcified tavr populations for solid mechanics simulations by means of deep learning. Current Directions in Biomedical Engineering 10(4)

[15] Verstraeten S, Hoeijmakers M, Tonino P, et al (2024) Generation of synthetic aortic valve stenosis geometries for in silico trials. International journal for numerical methods in biomedical engineering 40(1):e3778.

[16] Karanasiou G, Edelman E, Boissel FH, et al. Advancing In Silico Clinical Trials for Regulatory Adoption and Innovation. *IEEE J Biomed Health Inform.* Published online November 8, 2024

[17] Pappalardo F, Wilkinson J, Busquet F, et al. Toward A Regulatory Pathway for the Use of in Silico Trials in the CE Marking of Medical Devices. *IEEE J Biomed Health Inform*. 2022;26(11):5282-5286