EU Horizon 2020 Research & Innovation Program
Digital transformation in Health and Care
SC1-DTH-06-2020
Grant Agreement No. 101016496

SimCardioTest - Simulation of Cardiac Devices & Drugs for in-silico Testing and Certification

SIM CARDIO TEST

Technical Report

D 6.4: Integration of standardisation working groups within standardisation organisations

Work Package 6 (WP 6)

Verification, validation, uncertainty quantification & certification

Task Lead: MPC, France

WP Lead: MPC, France

PUBLIC

DELIVERABLE INFORMATION

Deliverable number	D 6.4
Deliverable title	Integration of standardisation working groups within standardisation organisations
Description	Summary of actions (including exchanges with standardisation groups) for proposing standardised guidelines on verification, validation, and uncertainty quantification (VVUQ) of computational models applied to medical devices and drugs
Lead authors	Romano SETZU (MPC)
Contributors	For UC1 (WP2): Yves COUDIERE (UBx) For UC2 (WP3): Oscar CAMARA (UPF) For UC3 (WP4): Beatriz TRENOR (UPV), Maria Teresa MORA (UPV) For standardisation policy stakeholders: Janaki RANGARAJAN
Due date	M54
Submission date	30/06/2025
Comments	N.A.

Document history			
Date	**Version**	**Author(s)**	**Comments**
05/06/2025	V1	R. SETZU	First Draft
26/06/2025	V2	Y. COUDIERE, B. TRENOR, V. CARBONE, J. RANGARAJAN	First consolidated draft including feedback from WP6 stakeholders
30/06/2025	V3	Y. COUDIERE	Quality Review
30/06/2025	V4	R. SETZU	Final Version
30/06/2025	F	M. BARBIER	Final editing

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report and its attachments comprise Deliverable D6.4 of the SimCardioTest (SCT) project's Work Package 6 (WP6), scheduled for completion by June 2025. It describes the exchanges undertaken with selected standardisation bodies during years 2023 and 2024 addressing standardised guidelines on verification, validation, and uncertainty quantification (VVUQ) of computational models applied to medical devices and drugs. Additionally, the deliverable briefs on the ongoing initiatives on the integration of the VVUQ learnings into working groups within standardisation organisations.

During the first half of 2023, together with VPHi (SCT partner), WP6 undertook a joint revision of the ISO/DTS 9491-1 draft technical specification which was under ballot. Comments were sent on April 27th, 2023, to the ISO secretariat through DIN (Deutsche Institut für Normung). Through our comments we aimed to provide ISO/DTS 9491-1 with relevant feedback, offering our perspective on the VVUQ assessment of numerical models in medicine and highlighting what we considered sensible points and potential axes of improvements. The consolidated list of comments is included in the attachments of this deliverable.

During the second half of 2024, SCT WP6 reached out to ASME VV40 technical committee in order to organise a joint meeting which was held on the side of the VPH conference in Stuttgart (Germany) on September 5th, 2024. The aim of this exchange was twofold. On one hand, for the SimCardioTest consortium to get constructive feedback on the pertinence of their interpretation and implementation of the VV40 standard guidelines on the WP6 VVUQ pipeline. On the other hand, to provide feedback to ASME VV40 committee on the VV40 standard usability based on SCT experience. The materials from the scheduled meeting with the ASME VV40 committee as well as minutes from this event are included in the attachments of this deliverable.

The above standardisation initiatives within SCT (among others in the wider computer modelling and simulation projects) helped the ecosystem organisations like VPH Institute (SCT partner) and Avicenna Alliance to gain support of ISO-IEC standardisation. This has led to the formation of an ad–hoc working group (ahG) predicating the creation of an international ISO-IEC standard. The progress directly relates to the T6.4 of SCT, which sets its ambition to initiate the development of a European standard. This deliverable briefly states the current state of this process and an indicative outlook, as creation of standards involves multiple stages and iterations which go beyond the scope of the SCT project.

Acronyms

Table 1: List of Acronyms.

Acronym	Meaning
AF	Atrial Fibrillation
ahG	Ad–Hoc Group
ASME	The American Society of Mechanical Engineers
Avg.	Average (abbreviation)
CEPS	Cardiac ElectroPhysiology Solver (cf. Use Case 1)
CFD	Computational Fluid Dynamic
CI	Continuous Integration
CiPA	Comprehensive in-vitro Proarrhythmia Assay (cf. Use Case 3)
COU	Context of Use
DE	Discretization Error (in Verification)
DRT	Device-Related Thrombosis
ECG	Electrocardiogram
EP-0D	0D Electrophysiology Model (cf. Use Case 3)
EP-3D	3D Electrophysiology Model (cf. Use Case 3)
EXC	ExactCure
FDA	US Food and Drug Administration
GSP	Good Simulation Practice
IEC	International Electrotechnical Commission
IST	INSILICOTRIALS TECHNOLOGIES SRL Also referring to the Cloud service hosting the models
LA	Left Atrium
LAAO	Left Atrial Appendage Occluder
MOTS	Modified Off-the-Shelf Software
MPC	MICROPORT CRM - SORIN CRM SAS
MV	Mitral Valve
N.A. / n.a.	Not Applicable

Acronym	Meaning
NCV	Numerical Code Verification
NSE	Numerical Solver Error (in Verification)
ODE	Ordinary Differential Equations
OTS	Off-the-Shelf Software
PK	Pharmacokinetics Model (cf. Use Case 3)
PR	Pulmonary Ridge
PV	Pulmonary Vein
QI	**Question** of Interest
QoI	**Quantity** of Interest
SCT	SimCardioTest
SQA	Software Quality Assurance (in Verification)
SRL	SIMULA RESEARCH LABORATORY AS
TAWSS	Time-Averaged Wall Shear Stress
TC	Test Condition (in Validation)
TdP	Torsade de Pointes
TS	Test Sample (in Validation)
UB / U.B.	Uncertainty Budget
UBx	Université de Bordeaux
UC	Use Case
UD	User Developed (Software)
UE	Use Error (in Verification)
UPF	UNIVERSIDAD POMPEU FABRA
UPV	UNIVERSITAT POLITECNICA DE VALENCIA
V&V, VV	Verification & Validation
VVUQ	Verification, Validation, and Uncertainty Quantification
WP	Work Package
WSS	Wall Shear Stress

1. Introduction and Objectives

In order to capitalize on all the work done to support the credibility assessment of the computational models developed within SimCardioTest (VVUQ activities conducted in WP6), it is paramount to share our experience with stakeholders implicated in the development of common standards to be adopted by the wider community [1].

This report and its attachments comprise Deliverable D6.4 of the SimCardioTest (SCT) project's Work Package 6 (WP6), scheduled for completion by June 2025. It describes the exchanges undertaken with selected standardisation bodies during years 2023 and 2024 addressing standardised guidelines on verification, validation, and uncertainty quantification (VVUQ) of computational models applied to medical devices and drugs.

Two international documents were identified:

- ISO/DTS 9491-1: [2]
- ASME VV40: Assessing Credibility of Computational Modelling Through Verification and Validation: Application to Medical Devices [3]

ISO/DTS 9491-1 draft technical specification was identified as undergoing international ballot and call for comments during the time of the SimCardioTest project. Although such document's primary focus is research in personalized medicine, we felt compelled to express our feedback for sake of harmonization with other international documents (such as ASME VV40 and FDA's guidelines [4]). During the first half of 2023, together with VPHi (SCT partner), SCT WP6 team undertook a joint feedback on the revision of ISO/DTS 9491-1. Through our comments we aimed to provide ISO/DTS 9491-1 with relevant suggestions, offering our perspective on the VVUQ assessment of numerical models in medicine and highlighting what we considered sensible points and potential axes of improvements.

ASME VV40 international standard was identified as main document upon which the current VVUQ strategy implemented within WP6 is based. During the second half of 2024, SCT WP6 approached the ASME VV40 committee to organise a joint meeting. The aim of this exchange was twofold. On one hand, for the SimCardioTest consortium to get constructive feedback on the pertinence of their interpretation and implementation of the VV40 standard guidelines on the WP6 VVUQ pipeline. On the other hand, to provide feedback to ASME VV40 committee on the VV40 standard usability based on SCT experience.

In addition to the two references cited above, another relevant document for computational model credibility assessment is the recent joint work hosted by the In Silico World community of practice, and supported by the VPH Institute and the Avicenna Alliance: "Toward Good Simulation Practice (GSP) - Best Practices for the Use of Computational Modelling and Simulation in the Regulatory Process of Biomedical Products" [5].
This initiative involved hundreds of in silico trial experts worldwide working in academia, healthcare, industry and regulatory bodies. Moreover, a team of 13 FDA M&S experts, covering all three medical product centres (CDRH, CDER and CBER) provided feedback on the whole draft document.

The document, published in February 2024, is available in open access and has registered so far more than 44k accesses. It explores multiple aspects related to using computational modelling and simulations in medical applications, including the model credibility assessment. It should be noted that the GSP is intended as a position report and is not an official standard. Although no direct liaison has been established between the SCT consortium and the GSP taskforce, two members of the consortium (VPHi and IST) have been actively involved in the development and editorial process of this document, including the chapter concerning the Model Credibility. This ensured that, although not actively involved, the SCT project was aware of this upcoming publication. From the Model Credibility chapter, we can infer that GSP largely aligns with the credibility approach and definitions proposed by ASME VV40 document. For this reason, and since the GSP was largely completed by the time the SCT consortium had gained enough maturity on the VVUQ framework, we decided not to seek further involvement on the GSP taskforce within the timeframe of the SCT project.

2. Methodologies

2.1 ISO/DTS 9491-1

For what concerns document ISO/DTS 9491-1, comments were jointly sent by VPH institute and the SimCardioTest consortium in April 27th, 2023 to the ISO/DTS 9491-1 document under ballot through DIN (Deutsche Institut für Normung). The consolidated list of comments is included in Appendix 1.
NOTE: the outline of the comments has been reformatted to fit the SimCardioTest project document template.

2.2 ASME VV40

Thanks to the intermediation of the VPH institute, a joint meeting between SCT WP6 members and ASME VV40 committee members was planned to take place in Stuttgart (Germany) on September 5th, 2024, on the side of the VPH 2024 conference.
A list of questions and discussion points was compiled by the SCT WP6 working group beforehand and shared with the ASME VV40 committee members taking part in the meeting.
Some of the questions were addressed directly during the meeting, while others (due to time constraints) were addressed offline, through e-mail exchanges.
All questions that were addressed are compiled in section 3, together with the feedback collected from the VV40 committee members. The materials from the scheduled meeting with the ASME VV40 committee is also included within Appendix 2, while the minutes from this event are included in Appendix 3.

3. Results on the Interaction with Standardisation Bodies

3.1 ISO/DTS 9491-1

All VPHi-SCT editorial comments have been accounted for by the ISO/DTS 9491-1, however, the comments regarding content (both general and technical) could not be implemented at the current stage of the standard's revision process. Therefore, the document has been voted, approved and published as such, while Martin Golebiewski from HITS (Heidelberg Institute for Theoretical Studies) co-leading the ISO working group communicated to VPHi that they appreciated the feedback and could likely take into account the detailed comments when enhancing the document in ISO/TC 276/WG 5 from ISO/TS to the level of an international standard, which is already planned. VPHi (SimCardioTest partner) expressed its intention to follow the future development of this document, as part of the Institute's work on standardisation.

3.2 ASME VV40

This section compiles all questions addressed during our exchange with ASME VV40 committee, and include the feedback collected from the VV40 committee members. The materials from the scheduled meeting with the ASME VV40 committee is included in Appendix 2, while the minutes from this event are included in Appendix 3.

All opinions expressed in this section by ASME VV40 committee members are personal opinions of the people involved in the meeting, and do not represent the official perspective of the ASME organisation nor of their respective companies.

3.2.1 Q1 - General: when and how-often should we run VV40 on a model?

From our experience, running VV40 on models under development is particularly challenging. The quantity of the work is significant, and we are tempted to consider that it should be run again each time the model undergoes some significant changes (endless work). It seems to us like a non-optimal way to proceed. What do you preconize in this circumstance? A guidance statement in this sense in ASME VV40 may be useful.

[Addressed remotely, feedback from Mr. Bischoff]

It is helpful to have a credibility plan at the start of developing the model. As you go through model iterations / light validation activities / etc, this plan can help to guide information you may want to hold on to. However, typically, all the rigorous V&V is done once model development is done.
In my experience, verification is more important to be done sequentially.

3.2.2 Q2 – General: V&V on modular models

VV40 does not address the case of a model consisting of several independent sub-models working in pipeline (e.g. SCT UC3).
Should we assume that we just apply VV40 to each sub-model independently, isolating inputs and outputs, or do we need additional work?

A guidance statement in this sense may be useful in ASME VV40.

[Addressed remotely, feedback from Mr. Bischoff]

Correct, V&V40 is silent on this topic. I think the hierarchical approach you mention is the right way. Distributing risk across sub-models is not clear; however, you could use the same credibility factors to tell the V&V story for each sub-model.

3.2.3 Q3 – On Credibility Coverage Gradations

For ensuring coverage of each credibility factor, VV40 proposes a list of choices (from less to most demanding). The standard clarifies that all gradations are only suggestions, and that each model should be treated differently. In addition, the standard leaves to the practitioner the task to convert the grade to scores in the Model Risk Matrix.
These points pose at least 2 issues, in our opinion:

- Risk of inconsistency in the coverage of different models
- Arbitrary choice given to the user: one may be tempted to do less

A solution would be multiplying the worked-out examples, beyond the content of the current VV40 Annex.

- Could consider creating a repository of examples, for users to use as guideline for different models?
- Did you consider a peer-review system to confirm that the model has been adequately addressed? Peer-review journals? Other ways?

[Addressed during meeting, see Minutes in Appendix 3]

3.2.4 Q4 – On Model Risk Matrix

The Model Risk Matrix is a very useful and clear tool. However, by design, it may lead in some situations to a blocking point.
What happen if, no matter what, and for some good reasons, one (or few) credibility factors cannot be covered in a satisfactory manner with respect to the model risk and the established ranking?

In this case, VV40 suggests several approaches (cf. ASME VV40 Figure 7-1), such as conducting additional credibility activities, changing the computational model, reducing the model influence and/or modifying the COU. These options are sometimes not possible. In that case, abandoning the model seems the only choice left by the standard.

But then, VV40 also states (VV40 §5 NOTE): It may be valuable for stakeholders to consider how exceeding or missing a specific credibility factor goal would change the overall credibility of the computational model.
This sentence seems somehow contradict the flow-chart above, and open to some justifiable arbitrations.

Indeed, we may have some credibility factors which we consider sufficiently covered (to the best of our abilities), but, due to the proposed ranking of activities, do not reach the required Risk level. This would not invalidate at once the credibility of the model.
VV40 may benefit from improving this point, offering some more clarification, and some alternative strategies. In particular: do all credibility factor coverages need to meet Risk level equally? Or some derogations are possible, and under what conditions?

[Addressed during meeting, see Minutes in Appendix 3. The following additional feedback was provided remotely by Mr. Bischoff]

VV40 contains no strong statements implying that all credibility factors must be covered to a level matching the risk level, if the implied work for doing so is disproportionate with the coverage level which is considered adequate from a technical standpoint.
In that case, one should document the reasons why the additional work necessary to cover the risk level is not necessary.

3.2.5 Q5 – Medical Devices vs. Drugs

VV40 is intended for numerical models of medical devices, addressing safety and efficacy.
What about Drugs? Is there a major push-back not extending VV40 to drugs interactions with humans?

[Addressed offline with Mrs. MARTIS, see Minutes in Appendix 3]

3.2.6 Q6 – On Model Uncertainty

We wonder why uncertainty is only discussed on VV40 Validation sections.
To our understanding, many uncertainty factors deriving from verification activities contribute to the overall model uncertainty (e.g. time and space discretization uncertainty, solver uncertainty). Such terms are not explicitly mentioned in the model uncertainty sections, which instead focus on model form and model inputs.

[Addressed remotely, feedback from Mr. Bischoff]

Gap in V&V40, that are we intending to address in the revision. UQ from verification should also propagate forward, both to the validation model and the COU model.

3.2.7 Q7 – On Model Applicability

What if we cannot assess whether the gap between comparator and context of use is acceptable? Can we still defend the model?

[Addressed remotely, feedback from Mr. Bischoff]

This relies on the strength of the rationale. However, to me, in order to 'defend' the model, that is the same thing as defending why the 'gap is acceptable' – you cannot do the former without the latter.

3.2.8 Q8 – On Annexes

On VV40 annex examples (very useful and detailed), we noticed that each example addresses only a subset of the credibility factors, and the subset is different from example to example and seems somehow arbitrary.
The feeling is that we are entitled to choose only the credibility factors that we understand and can address, rather than covering all of them.
Some wording may be useful for clarifying that upfront.
We also suggest adding at least one (or some) examples which address systematically all factors.

[Addressed during meeting, see Minutes in Appendix 3]

3.3 VVUQ Working Group within Standardisation Organisations

Following up the various standardisation activities pertaining to the computational modelling, the International Electrotechnical Commission (IEC) Technical Committee on "Medical Equipment, Software and Systems" (TC 62), has formed an ad–hoc working group, the "ahG 11", tasked with drafting a preliminary work item on this topic [1].
SCT consortium partners VPH Institute and InSilicoTrials are proactively following this working group activities, whereby the VVUQ learning and advancements made within WP6 are leveraged.

The goal of the IEC working group, constituted by members of the industry, standardisation bodies (ISO-IEC, ASME VV40), along with members of Avicenna Alliance, is to draft a preliminary work item to establish how to demonstrate the credibility of computational models in medical device applications. This includes all types of models, such as those based on knowledge (mechanistic), as well as those derived from only data, and hybrid approaches. The objective is to involve the principles of VVUQ of the computational models in the same lines as what the SCT project has been championing for. Thus, strive for reliability of computational models, by establishing a new standard on the principles of verification, validation and uncertainty quantification.

As one can presume, the process of drafting a consensus standard is long and involves multiple formal stages. The ahG working group is in forming stage, preparing to present an outline of the eventual standard, envisioned to be presented in the upcoming TC62 plenary meeting at the end of 2025.

SCT consortium is represented in the ahG 11 working group through VPHi and IST. Both partners look forward to bringing the prior work on GSP as well as the SCT's VVUQ approaches and recommendations to this IEC working group. The partners aim to possibly inspire and eventually incorporate the SCT use cases as tangible examples to strengthen the envisioned standard on modelling and simulation.

[1] IEC-TC 62 ahG - Establishing the credibility of computational modelling in the field of medical devices through verification, validation, and uncertainty quantification
https://iec.ch/ords/f?p=103:14:5500319302698::::FSP_ORG_ID,FSP_LANG_ID:51475,25

4. Conclusion

This report and its attachments constitute the Deliverable D6.4 of the SimCardioTest (SCT) project's Work Package 6 (WP6). In this report we summarized the activities engaged with selected standardisation bodies during years 2023 and 2024 addressing standardised guidelines on verification, validation, and uncertainty quantification (VVUQ) of computational models applied to medical devices and drugs.
Two main international documents were selected: ISO/DTS 9491-1 draft technical specification and ASME VV40 international standard.
For what concerns ISO/DTS 9491-1, SCT WP6 contributed with a series comment jointly sent with VPHi through the DIN channel during the ballot of the document. According to the received feedback, such comments will be taken in consideration for the development of the future ISO 9491-1 international standard.
For what concerns ASME VV40, SCT WP6 organised with the help of VPHi a joint meeting with the ASME VV40 technical committee. The meeting was beneficial to both parties. For WP6 it provided positive feedback that the VVUQ work undertaken within WP6 for demonstrating the credibility of the computational models in Use Cases 1/2/3 was carried out consistently within the guidelines of the ASME VV40 standard. For ASME VV40 it provided constructive feedback on the current limitations on the usability of the standard and on the potential axes of improvement, some of which are already currently accounted for in the future revision of the document and the development of related new technical reports.
Building on the valuable learnings and demonstrations of VVUQ within SCT, the VPH community is set to champion the ASME VV40 and Good Simulation Practice to an internationally recognised IEC standard in the years to come.

5. Bibliography

[1] R. Setzu, A. Olivares, J. Mill and others, "Building credibility of computational models in cardiovascular medicine through verification and validation," *EDMA,* vol. 19, pp. 86-90, 2024.

[2] ISO/DTS 9491-1, "Biotechnology — Predictive computational models in personalized medicine research — Part 1: Constructing, verifying and validating models," ISO, 2023.

[3] ASME V&V40, "Assessing Credibility of Computational Modeling Through Verification and Validation: Application to Medical Devices," ASME, New York, 2018.

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

[5] M. Viceconti and L. Emili, Toward Good Simulation Practice: Best Practices for the Use of Computational Modelling and Simulation in the Regulatory Process of Biomedical Products, Synthesis Lectures on Biomedical Engineering - Springer, 2024.

Appendix 1 – Comments for ISO/DTS 9491-1 Ballot

Table 2: List of VPHi/SCT comments in the frame of ISO/DTS 9491-1 ballot

Clause/ Position	Comments	Proposed Change
Subtitle	[general] The subtitle is not appropriate as most of the document deals with data preparation for model construction, which is interesting and should be better highlighted in the title. The topic of verification should be either developed in the text or completely removed from the title since it is currently absent from the text (the word "verification" itself is only cited 2-3 times, in side-notes or external references)	Remove "verifying form the sub-title" For example: "Part1: Data preparation for constructing models and model validation" Else, clarify and expand the role of "verifying" computational models within the scope of this standard.
Introduc -tion 5th Paragra -ph	[general] The sentence states: "This document presents modelling requirements and recommendations for research in the field of personalized medicine, especially with focus on collaborative research, such that health-related data can be optimally used for translational research and personalized medicine worldwide." It should set the context and the scope of the document as it concludes the introduction, but it could be misleading for several reasons: • It gives the impression the document is prescriptive, even though it is a Technical Specification and not a standard yet. Setting the tone and objectives of the document would be valuable for external readers. • The term 'computational modelling' should be used in full as "modelling" alone encompasses many other fields • It states that the purpose is to use data for "translational research and personalised medicine." although the title of the document focuses on usage for model construction, verification and validation, which is different. • It suggests the document covers all types of computational modelling approaches for personalised medicine although the content of the document is largely oriented towards the molecular, cellular and drug field and overlooks other types of modelling relevant for personalised medicine (3D organ models, computerised assisted design, etc.)	Replace with: "This document proposes recommendations and points at important standardisation needs for what concerns computational modelling for non-clinical research in the field of personalised medicine, especially with focus on collaborative research, such that health-related data can be optimally used for computational modelling research and personalised medicine worldwide. The recommendations are primarily oriented towards biotechnologies (e.g. biomolecular and cellular research) although the vocation is to broaden the scope."

Clause/ Position	Comments	Proposed Change
1 2nd	[general] The sentence states: "This document does not apply to computational models used in clinical, diagnostic or therapeutic purposes." This sentence seems to be contradicted several times across the document body. For instance, in Table 3 - "Cellular systems biology models" R6 states that: "User-friendly graphical interfaces should be developed to ease the use of models in clinics." Similar ambiguous claims can be found across the document. *In addition, it could be made clearer whether applications such as computational modelling for health product development (industrial purpose) are also part of that exception or not.	Verify consistency of statement in section 1 across the document. Remove unjustified exceptions. When exceptions are justified, add a note as this does not contradict the overall document scope statement.
3.1 1st Paragra -ph	[technical] The definition of AI is provided as" Artificial intelligence <system> capability to acquire, process, create and apply knowledge, held in the form of a model, to conduct one or more given tasks". The definition of AI does not align well with other publicly recognized definitions, such as the ones provided in the EU AI Act or by the OECD. In particular, the current definition is not satisfying because AI is a system (as categorised in the current document), hence it can't be a 'capability' at the same time. In addition, the terms computer, computational or machine are clearly missing. The concept of possible autonomy is also key in the definition of AI systems.	Revise definition with: "Machine-based system capable to acquire, process, create and apply knowledge, held in the form of a computational model, to conduct one or more given tasks with varying levels of autonomy. The capacity to learn for an algorithm may arise from different processes such as supervised or non-supervised learning and reinforcement learning."
3.5 1st Paragra -ph	[technical] The definition of computational model is given as "description of a system in a mathematical expression and/or graphical form highlighting objects and their interfaces" It is commonly accepted that a computational model is more than just a mathematical expression, it is also the implementation of the mathematical expression in a computer. It can't be dissociated from the set of algorithms allowing the dynamic study of the system. Otherwise, one would talk about "mathematical model" not "computational model". Here, the scope is mainly on the biological/ biomedical side, hence the biological nature of the system could be specified.	Revise definition as follows: "computational model in silico model description of a biological system in a mathematical expression and/or graphical form that is implemented and studied with a computer and highlighting objects and their interactions."

Clause/ Position	Comments	Proposed Change
3.6 1st Paragra-ph	[technical] This definition of data-driven model tends to exclude a large variety of data-driven models that do not rely on tests or investigated processes. E.g data-driven models using real world data, clinical-routine data, which is accounted for in the rest of the document, actually.	Revise definition as follows: “data-driven model model developed through the use of data derived from tests, from the output of investigated processes or from real world data or routinely acquired primary care data. “
4.2.1 Figure 1	[technical] A step for model building/training is missing, although it is a critical part between Data integration and Model Simulation	Include model building/training in the Figure
4.2.1 1st Paragra-ph	[general] The reference cited for FDA (ref [16]) could be updated as a new version of the FDA report was released in 2022. see: https://www.fda.gov/science-research/about-science-research-fda/modeling-simulation-fda	Add reference: https://www.fda.gov/science-research/about-science-research-fda/modeling-simulation-fda
4.2.1 1st Paragra-ph	[general] The sentence states: “Computational models are integrated in different fields in medicine and drug development expanding from disease modelling, molecular biomarker research to assessment of drug efficacy and safety.” In this paragraph but also overall in the text, the content is largely oriented towards Biotechnologies and Pharmaceutical applications while it regrettably ignores MedTech applications even though the title and the general introduction are very general and seem to promise to encompass the whole spectrum. Therefore, we recommend to either restrict the scope and specify the title at the beginning or to enrich the document to be more inclusive and include a larger variety of examples. In most cases replacing “drug” by “health product” will already be more inclusive. See example of suggestion for current paragraph. (Other examples of possible changes are listed in other comments below)	Revise the document to be more inclusive with MedThec applications for personalised medicine. Consider replacing “drug” by “health product”, where appropriate. For example, replace the cited sentence by: “Computational models are integrated in different fields in medicine and health product development expanding from disease modelling, molecular and physiological biomarker research to assessment of drug and medical device efficacy and safety.”
4.2.2.1 1st Paragra-ph	[general] Both data-driven and mechanisms-based models are mentioned here. The term data-driven model was defined in Clause 3 but mechanisms-based model was not. It should be added since it represents a large part of the models covered in the current document.	Add definition for mechanism-based models in Clause 3.
4.2.2.1	[technical] The sentence states: “Data-driven approaches require sufficiently rich and quantitative time-course data to train and to validate the model”	Remove ‘time-course”: Data-driven approaches require sufficiently rich and quantitative

Clause/ Position	Comments	Proposed Change
2nd Paragra -ph	This statement is important; however, it should be acknowledged that many data driven models do not require time-course data. Actually, some algorithms are not designed to handle time-course data and can NOT exploit that type of information.	data to train and to validate the model.
4.2.2.1 2nd	[editorial] The sentence is: "Due to its often black-box nature, the model validation process in data-driven approaches relies on performance tests against known results." The grammar of the sentence should be revised to make clearer what 'its' refers to in the following part "Due to its often black-box nature". The black-box nature should refer to the data-driven models, although in the current form it seems to refer to the validation process.	Correct with: "Due to the often black-box nature of data driven approaches, the model validation process relies on performance tests against known results".
4.2.3.2 2nd Bullet point	[general] We understand that the list of challenges in the entire clause 4.2 may not aim to be exhaustive, but we recommend generalising it as much as possible in order to avoid restricting the application of the current standard to purely genetic and molecular applications. For example, in the sentence: "Limited replication of genetic associations and poor application of diverse populations (e.g. too poorly represented to be of interest for specific analyses), specifically of mixed or non-European ancestry.", limited replication of experiments and lack of representativity of diverse populations is a common challenge for several types of experimental data, not only genetic associations. Generalising it would greatly increase the impact of the document. We strongly encourage the authors to revise the entire document with that perspective in mind.	Revise Clause 4 to generalise concepts and avoid restricting the document to genetic applications, where applicable. For example, correct the cited sentence with: "Limited replication of measurements and analyses and poor application of diverse populations (e.g. too poorly represented to be of interest for specific analyses), specifically of mixed or non-European ancestry."
4.2.4.2 1st Paragra -ph - 2nd bullet point	[general] "Developing transparent and quality-controlled workflows for molecular data generation and interpretation in clinical settings." The statement would still be true, informative and would be more inclusive by removing the term "molecular".	Developing transparent and quality-controlled workflows for data generation and interpretation in clinical settings.
4.2.5 Sub-clause title	[technical] The title could be improved to better represent the content of the paragraph, which is not only on PK/PD modelling but also encompasses all sorts of pharmacometrics models (QSP, etc.).	Replace title by: "Pharmacometrics and in silico trial simulations".

Clause/ Position	Comments	Proposed Change
4.2.5.1 1st Paragra -ph	[general] The section would gain from adding a description of the the main usage of PBPK modelling.	Add sentence: “PBPK models are commonly used for interspecies extrapolations and drug-drug interactions modelling.“
4.2.5.1 1st Paragra -ph	[general] Alongside PBPK, it may be interesting to also mentioned IVIVC (In Vitro/In Vivo Correlations).	Add sentence: IVIVC (In Vitro/In Vivo Correlations) methods can be used to extrapolate in-vivo absorption of drugs from in-vitro data. Add reference to: Guidance for Industry Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations https://www.fda.gov/regulatory-information/search-fda-guidance-documents/extended-release-oral-dosage-forms-development-evaluation-and-application-vitroin-in-vivo-correlations Add reference to: Guideline on the pharmacokinetic and clinical evaluation of modified release dosage forms EMA/CHMP/EWP/280/96 Rev1 Appendix III: in vitro in vivo correlation.
4.2.5.2 3rd Bullet point	[general] The sentence states: “Reporting of results is very heterogeneous and inconsistent”. This is correct; however, it would be valuable to mention relevant existing initiatives and documents supporting researchers in that endeavour, to draw an up-to-date view of the current state of the art.	Add reference to: GUIDELINE ON REPORTING THE RESULTS OF POPULATION PHARMACOKINETIC ANALYSES (Doc. Ref. CHMP/EWP/185990/06) https://www.ema.europa.eu/en/reporting-results-population-pharmacokinetic-analyses-scientific-guideline Add reference to: Guideline on the reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation EMA/CHMP/458101/2016

Clause/ Position	Comments	Proposed Change
		https://www.ema.europa.eu/en/reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation-scientific-guideline
4.3.2 1st Paragra-ph	[technical] Some important challenges are missing from the list. Consider adding the suggestions listed in the 'proposed changes.	Add as challenge: • Lack of domain/data-specific standard method for data pre-processing • Lack of standard and GDPR compliant workflow for personal health data access and processing. (E.g common workflow language, etc.) • Lack of training, awareness and empowerment for existing standards and workflows
4.4.1 1st Paragra-ph	[general] In line with the earlier comments on sub-clause 4.2.3.2, 4.2.3.4 to generalise / be inclusive, the statement would still be valid, after replacing "life sciences", by a general term like "Healthcare". Also, rephrasing could improve readability.	Rewrite as: "Computational models in healthcare and in particular research on personalized medicine, are increasingly incorporating rich and varied data sets to capture multiple aspects of the modelled phenomenon."
4.4.2 1st Paragra-ph	[general] This entire sub-section on 'Sampling Data', covers all aspects of "biological specimens. It is well written; however, the preceding and subsequent sections on 'standards and 'formatting', remain broad on all data types (including medical images) and not just specimens. While the criticality of 'sampling' for biological specimens is completely understandable, it seems to take a deep dive on one topic, without any pre-text. This disrupts the broader flow of the document.	Add: a broad outline on 'Sampling', before detailing on specimens.
4.4.2 1st Paragra-ph	[technical] Some recommendations could be added regarding the use of standard methods in measurements. Indeed, the lack of repeatability of many measurement methods hampers the acquisition of reliable data that are key for model construction (e.g. tissue mechanical properties, some enzyme kinetic constants, etc). See suggestion.	Add: "Measurement methods for analysing the samples should follow standard approaches as much as possible. For instance, characterisation of biological tissues (e.g. mechanical resistance) should be done following community consensus approaches

Clause/ Position	Comments	Proposed Change
		in order for the data to be reliable and accurate enough for modelling purposes (ref: Famaey, N., & Fehervary, H. (2022). C4Bio: Community Challenge towards Consensus on Characterization of Biological tissue. https://c4bio.eu/)"
4.4.2 Table 1	[technical] A critical information that should be mentioned in the category "Information about the specimen, collection from the donor or patient and processing" is the anatomical location where the sample was taken from (at organ level but even relative position or x,y,z coordinates, if applicable) preferably following existing standards.	Add Anatomical location where the sample was taken from (at organ level but even relative position or x,y,z coordinates or genetic locus, if applicable) preferably following existing standards.
4.6.1 1st Paragra -ph	[editorial] Suspected repetition in sentence: "which are accurate and confident in predictions, both in terms of accuracy and confidence in predictions".	Erase "both in terms of accuracy and confidence in predictions". If the rephrasing does not reflect the initial purpose of the sentence, please consider alternative rewording to clarify its meaning.
4.6.1 1st Paragra -ph	[technical] The entire section 4.6.1 should be enriched as it should be the core of the current technical specification. As of now, the topic of model validation is too superficially addressed. For instance, the reference to V&V40 or existing European guidelines could be better emphasised and explained here and some considerations about uncertainty quantification may be added.	Enrich the section on validation. Replace sentence: "There are guidelines and methods for validating models, which are accurate and confident in predictions." by: "Existing domain-specific guidelines and methods for validating models should be leveraged to assess the credibility of any model and ensure the accuracy and confidence of the predictions. Of interest is the ASME VV-40-2018 "Verification and Validation in Computational Modelling of Medical Devices", being a risk-based framework stating that the level of validation should be commensurate to the risk associated with using the model in a pre-specified context of used.

Clause/ Position	Comments	Proposed Change
		An important phase in any validation process is to characterise and quantify the model uncertainty (ref: 10.1016/j.compbiomed.2022.106407, 10.1098/rsif.2021.0864), which is often classified as aleatoric and epistemic uncertainty. Existing standards from other engineering fields on that matter may be applicable. “
4.6.2 Table 3	[general] The recommendations done in Table 3 are intended for model validation. However, it is unclear how they all can serve the process of validating the model itself. In fact, some recommendations (see below) address aspects which are clearly outside the scope of validation itself, which is intended as “comparison between the output of the calibrated model and the measured data”, as defined in 3.15. In some cases, the recommendations are addressed to the standard practitioners, but concern other aspects of the model than the model validation. For instance, “Cellular systems biology models - R6” recommends that “User-friendly graphical interfaces should be developed”. This recommendation seems more adequate for the overall model usability than model validation. In some cases, the recommendations are not clearly addressed to the standard practitioner but are rather addressed to the whole community involving all stakeholders, and are more appropriate as recommendations in a tribune, than as guidance for the standard practitioner. This is the case, for instance, of “Disease course and therapy response prediction - R5 and R6”.	Clarify the scope of Table 3, as table of recommendations for a robust validation to the standard practitioner. And leave all recommendations destined to the community, or recommendations beyond the scope of model validation outside of this table. Consider creating different tables for grouping the recommendations which are clearly outside the scope of Table 3.
4.6.2 Table 3	[technical] Pharmacokinetic/-dynamic modelling and in silico trial simulations - R3 - NOTE The need for V&V40-like initiative in the context of disease description and drug development is rightly emphasised. However, text suggest that such initiative have never been considered, which is incorrect as can be seen from the following white paper that explored the applicability of V&V40 in the context of drug development: “Scientific and regulatory evaluation of mechanistic in silico drug and disease models in drug development:	Mention existing initiatives and replace sentence “In the current situation, a similar initiative oriented to disease description and drug development would be of great value.” by: “Some initiatives have proposed to use a similar framework oriented to disease description and drug development, through the use of a credibility matrix (ref: Scientific and

Clause/ Position	Comments	Proposed Change
	Building model credibility" - CPT: Pharmacometrics & Systems Pharmacology Volume 10, Issue 8 p. 804-825." https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/psp4.12669	regulatory evaluation of mechanistic in silico drug and disease models in drug development: Building model credibility" - CPT: Pharmacometrics & Systems Pharmacology Volume 10, Issue 8 p. 804-825)."
4.6.2 Table 3	[editorial] Missing horizontal separator in table between R2 and R3 in "Pharmacokinetic/-dynamic modelling and in silico trial simulations"	Add horizontal separator
4.6.2 Table 3	[editorial] Recommendation R3 in "Pharmacokinetic/-dynamic modelling and in silico trial simulations" is unclear, and possibly contains a typo, or a fragmented sentence.	In the sentence: "It should be ensured that the standards (to be) used for model development, evaluation, reporting and related decision-making are commonly acknowledged by all the involved parties (…) are relevant for all the types of models that can be used." remove the first underlined "are". If the rephrasing does not reflect the initial purpose of the sentence, please consider alternative rewording to clarify its meaning.
4.6.2 Table 3	[editorial] The sentence of recommendation R1 of "Artificial intelligence models" states "Model validation should involve a three-phase process". This is not a viable standalone sentence and does not read as a plausible recommendation per-se.	Replace by "Model validation should involve the three-phase process, as detailed in following recommendations R2-R4", or similar.
4.6.2 Table 3	[technical] The sentence of recommendation R4 of "Artificial intelligence models" is vague. The meaning of the sentence could be made more concrete. For example, if the meaning is to include a third phase of model validation that involves running a clinical validation of the tool this should be made more explicit.	Proposed rewording: "If beneficial with respect to the AI model context of use, a phase 3 validation should include a clinical validation of the AI prediction points by relevant clinicians." If the rephrasing does not reflect the initial purpose of the sentence, please consider alternative rewording to clarify its intended meaning.
4.6.2 Table 3 Artificial	[technical] The concept of validation phases is interesting. However, when looking closer to recommendation R3 it suggests that AI models should be compared to clinical decision	Revise R1: AI model validation may follow three-phases depending on the nature of the model.

Clause/ Position	Comments	Proposed Change
Intelligen-ce models R3	making (and is supported by a reference on AI-aided clinical decision making), which is surprising for two reasons. First, because all AI models are not necessarily meant to provide clinical decision support, therefore it should be made clearer that the recommended phases may not always be applicable. Second, the scope section (Clause 1) of the current document states that "This document does not apply to computational models used in clinical, diagnostic or therapeutic purposes." We recommend revising the document for consistency between the stated scope and the content.	
4.7.1 1st Paragra-ph	[general] The sentence "Model simulation brings the models to life" appears misplaced in a standard or technical specification as its scope is unclear and it does not add technical value.	Please replace the sentence by a more appropriate definition model simulation.
4.7.1 1st paragra-ph	[technical] Concerning the mentioned mathematical concepts: "from graph theory to dynamical systems theory". It is difficult to imagine that these two examples are the two extremes of the spectrum that is implied.	Replace by a list of examples: "(e.g. graph theory, dynamical systems theory)".
4.7.1 1st Paragra-ph	[technical] The sentence states: "Consequently, model simulation, computational costs and the resolution of results is model-specific with models having different levels of abstraction, and predictive power that can be categorized into three main levels: topological, constraint-based and kinetic modelling". We appreciate the effort of categorising the various simulation methods that may exist. However, this section mistakenly suggests that the three cited categories are an exhaustive list of the current state of the art. This is obviously an extremely restricted view of the field, and we strongly recommend revising that section. In addition, the construction of the above sentence suggest that the categories pertain to "levels of abstraction, and predictive power" rather than "simulation methods".	Refer to accepted standards or ontologies describing simulation methods. and enrich the section. The cited sentence could be reworded as: "Consequently, model simulation, computational costs and the resolution of results is model-specific with models having different levels of abstraction, and predictive power. Model simulation methods can be categorised in various ways, such as topological, constraint-based or kinetic modelling, which are three categories typically used for modelling of biochemical systems." The rest of the section and Table 4 caption should be revised, accordingly.

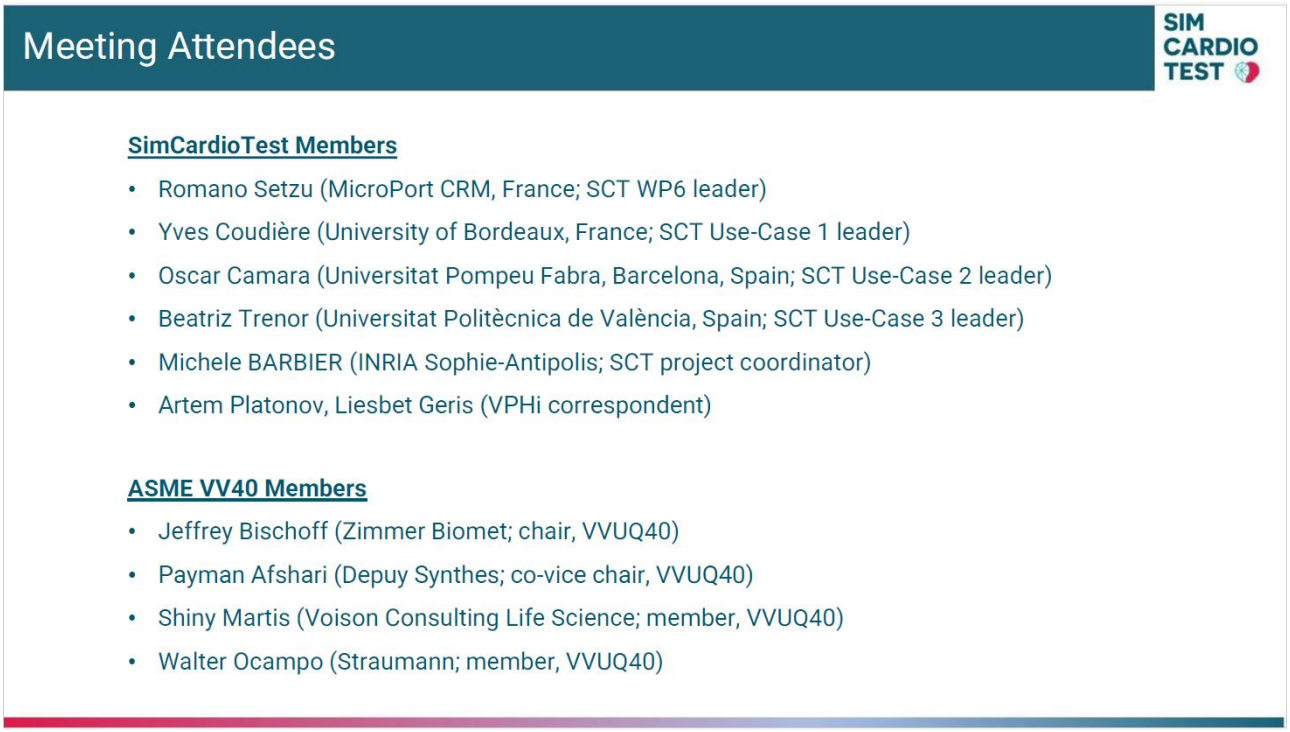
Clause/ Position	Comments	Proposed Change
4.7.1 1st paragra-ph	[general] Concerning the last sentence: "Data for the three main methods and their simulation…" The term "Data" should be clarified or replaced by a more appropriate term. It seems that the word "Data" in this sentence generically refers to "A description".	Replace "Data for the three main methods and their simulation…" With: "A description for the three main methods and their simulation…" Else clarify what it is meant by the word "Data" in the context of this clause.
4.7.1 NOTE 1, 2, 3	[technical] I would argue that NOTES 1 to 3 are very important in the frame of this clause, because they define the three methods for model simulation.	Proposal to make NOTES 1 to 3 as paragraphs of the clause rather than simple notes. In addition, use bold or italic text to highlight the three mentioned methods.
4.8 1st paragra-ph	[editorial] The sentence "…and shared in a way that make them accessible…" contains a typo. "make" should be "makes".	Rectify typo.
4.9.2 Bullet List	[editorial] A similar list of recommendations appears in Table 3 above.	Harmonize the presentation of the list of recommendations as in Table 3 above. Choose an appropriate table caption.
4.9.2 2nd bullet in list	[editorial] The word "fora" may be difficult to understand for non-English speakers.	Replace "Fora" with "Forums", easier to understand.
4.9.2 5th bullet in list	[technical] Recommendation R5 should be made clearer. The "whereas" conjunction seems misplaced. "Whereas…" reads "in contrast or comparison with the fact that…".	Rephrase R5 to: "Independent of the model type, the rigor in model evaluation should be the same, and all stakeholders (…) should use the same valid tools." If the rephrasing does not reflect the initial purpose of the sentence, please consider alternative rewording to clarify the connection between the two sentences connected by "whereas".
4.9.2 5th bullet in list	[technical] The sentence states: "Independent of the model type, the rigor in model evaluation should be the same", The sentence should be clarified as in the current state, it seems unrealistic and conflicting with the commonly accepted framework. In particular, it is unrealistic that an organisation will apply the same depth of validation for a model used in preliminary drug discovery as for a model	Revise sentence as follows: "All models should be rigorously evaluated. The rigour in model evaluation and credibility assessment should be defined according to the context in which the model is used and commensurate

Clause/ Position	Comments	Proposed Change
	used in complement or support of a clinical trial. As a matter of fact, both FDA and EMA recommend that the level of model credibility assessment should be dependent on the context of use and commensurate with the risk associated with using the specific model.	with the risk associated with using the model in that context".
4.10 4th paragra-ph	[general] The lack of transparency of AI driven models is a known limitation in the purpose of model validation. That said, the recommendation derived from this statement, that such models are less ethically suitable for deployment in medicine, seems too strong and categorical. We can already witness and anticipate that AI-augmented diagnostic tools will have an important role in diagnostic in the future. Why would we confine them to research and development because of their black-box form? In addition, it is unclear why this lack of transparency would make them less "ethically suitable". It is arguable that in this situation the ethics would rely more on the quality of data on which the models are based/trained/validated rather than the model operation itself.	Reconsider this statement, and if agreed that it is overshot, please consider removal, or rephrase to nuance it further.
Annex A Table A.3	[technical] We understand that the authors did not aim to provide an exhaustive list of all standards in the Annex A, but we strongly recommend adding more examples of data standards commonly used for 3D or multi-physics modelling, with the objective to increase the scope of the current document. For example, the STEP-file is commonly used to represent 3D objects.	Add a row: Standard: STEP-file Description: STEP-files are widely used data exchange form in computer-aided design (CAD) and to represent 3D objects since they contain three-dimensional model data for a wide variety of design tasks. The format of a STEP-file is defined in ISO 10303-21 Clear Text Encoding of the Exchange Structure. They are frequently used as input to represent organ and biological structure shapes in computational models.
B.1 1st paragra-ph	[editorial] The sentence "The genetic variance of a disease is a combination of small effects of multiple variants across the allele frequency spectrum" is repeated twice.	Remove repetition.
B.2 1st paragra-ph	[general] The sub-clause is missing an introductive sentence clarifying the object of the clause.	Add an introductive sentence such as: "Artificial Intelligence models can be grouped in three categories: supervised, semi-supervised, and unsupervised learning models".

Clause/ Position	Comments	Proposed Change
Whole docum-ent	[general] In general, there are 4 main aspects across the document which need improvement: 1. There is a mix between guidance to the standard practitioner and proposal guidelines for the whole community which make it difficult to use the document. 2. There is ambiguity on the role that model verification plays in this document. It is part of the title, but never explicitly developed. 3. There is ambiguity on the Context of Use of the models concerned by this document. The document scope claims that the models only concern predictive medicine research, and all clinical uses are out of scope. This point seems contradicted across the document. 4. The types of models accounted for in the current document overlook a large part of the spectrum that would be relevant in the context of personalised medicine.	To fix the cited general issues: 1. Review the document in the perspective of standard practitioner (guidance to develop and validate models for medical research), and in the perspective for the whole community (guidance to develop new guideline to foster progress in this field). Where needed, rewrite and reorganise text such as these two scopes do not overlap. 2. Clarify the role that model verification plays in this standard. 3. Clarify the context of use of the models concerned by this standard: solely "predictive medicine research" or beyond (including clinical applications). 4. Generalise statements that are applicable beyond genetic data and biochemical models and introduce more examples from alternative modelling technologies.

Appendix 2 – Slide Deck support of ASME VV40 discussion

Slide Deck 1 – General Introduction



Goal of the Meeting

- For SimCardioTest to get some constructive feedback on the pertinence of our interpretation of the VV40 standard
- To give some feedback to VV40 on the standard usability, based on our experience

Minutes of the meeting will be compiled and included as part of the SCT project deliverable (upon review and agreement)

Agenda

- General Introduction (Romano SETZU) – 15 min
- UC1 – V&V approach and results (Yves COUDIERE) – 10 min + 5 min QA
- UC2 – V&V approach and results (Oscar CAMARA) – 10 min + 5 min QA
- UC3 – V&V approach and results (Beatriz TRENOR) – 10 min + 5 min QA
- Questions / Discussion (All) – 30 min

SimCardioTest project in a nutshell

Project Goal
To demonstrate feasibility, effectiveness and benefits of in-silico trials for cardiac devices & drugs

How
We develop a standardized and secure cloud-based platform where in-silico trials will be run

Then
We test the platform effectiveness on 3 cardiac use-cases:
- UC1: Active Implants
- UC2: Passive Implants
- UC3: Drugs

Standardised Cloud-based Platform
WP1

For each Use-Case
- We develop a computational model
- We assess model credibility according to VV40 guidelines
- We run in-silico trials on the validated model using the cloud platform

4 years project
Ongoing till end 2024 (+6M ?)

SCT – WP6 Scope

Demonstrating Model Credibility through VVUQ is paramount to the SimCardioTest project

- To **verify** that the simulation software is achieving reasonable numerical errors
- To **validate** that the simulated quantities of interest are realistic enough for the context of use
- To **quantify the uncertainty** in the prediction of these quantities of interest

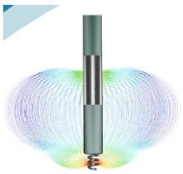
SCT V&V activities well advanced,
But still ongoing

VVUQ on Selected Questions of Interest

For each Use Case, we focus VVUQ on one selected Question of Interest

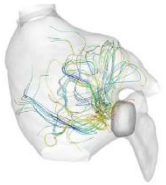
Use Case 1 – Lead Electrical Performance

• What are the stimulation pulse characteristics (voltage amplitude in V and pulse duration in ms) required for a bradycardia lead in bipolar (tip/ring) mode to capture (stimulate) healthy cardiac tissue?



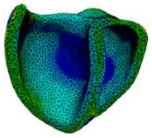
Use Case 2 – LAAO Safety

• Does covering of the pulmonary ridge with a LAAO device (plug or pacifier) relates with the likelihood low blood flow (<0,2m/s) velocities around the device and induce the device-related thrombus (DRT)?



Use Case 3 – Safety: TdP Risk

• What is the maximum concentration/dose regimen of a drug to assure TdP-related safety in a population of healthy subjects?



WP6 – Harmonized V&V approach

Significant Effort in VVUQ Harmonization
Based on ASME VV40 guidelines

- Challenge due to very different Use-Case scopes:
  - Active Devices, Passive Devices, Drugs

For selected Questions of Interest, we specified:
- Context of Use
- Model Risk
- Comparator for Validation

- We ensured a proper coverage of **all** credibility factors in ASME VV40

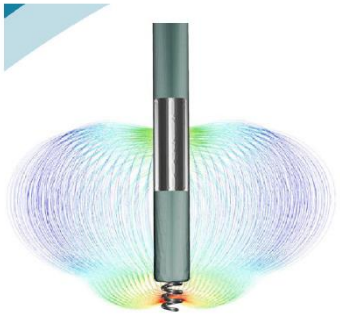
Model influence	high	3	4	5
	medium	2	3	4
	low	1	2	3
		low	medium	high
		Decision consequence		

Credibility Factor Coverage Level (per ASME VV40)		1	2	3	4	5
Code Verification: Software Quality Assurance	I - V					
Code Verification: Numerical Code Verification - NCV	I - V					
Calculation Verification - Discretization Error	I - V					
Calculation Verification - Numerical Solver Error	I - V					
Calculation Verification - Use Error	I - V					
Validation - Model [Form]	I - V					
Validation - Model [Inputs]	I - V					
Validation - Comparator [Test Samples]	I - V					
Validation - Comparator [Test Conditions]	I - V					
Validation - Assessment [Input Parameters]	I - V					
Validation - Assessment [Output Comparison]	I - V					
Applicability: Relevance of the Quantities of Interest	I - V					
Applicability: Relevance of the Validation Activities to the COU	I - V					

UC1 – Credibility Factors Coverage

Model Risk: 2/5
- Decision Consequence: Low
- Model Influence: Medium

Credibility Factor Coverage Level (per ASME VV40)		1	2	3	4	5
			x			
Code Verification: Software Quality Assurance	V		x			
Code Verification: Numerical Code Verification - NCV	V		x			
*Calculation Verification - Discretization Error **	V		x			
*Calculation Verification - Numerical Solver Error **	V		x			
Calculation Verification - Use Error	IV		x			
Validation - Model [Form]	III		x			
*Validation - Model [Inputs] **	III		x			
*Validation - Comparator [Test Samples] **	IV		x			
*Validation - Comparator [Test Conditions] **	II		x			
*Validation - Assessment [Input Parameters] **	III		x			
Validation - Assessment [Output Comparison]	N.A.		x			
*Applicability: Relevance of the Quantities of Interest **	V		x			
*Applicability: Relevance of the Validation Activities to the COU **	V		x			



* ongoing V&V activities
N/A: V&V activities requiring completion to be evaluated

UC2 – Credibility Factors Coverage

Model Risk

Credibility Factor Coverage Level (per ASME VV40)		1	2	3 (COU 2)	4 (COU 1)	5
Code Verification: Software Quality Assurance	V			x	x	
Code Verification: Numerical Code Verification - NCV	V			x	x	
Calculation Verification - Discretization Error	V			x	x	
Calculation Verification - Numerical Solver Error	V			x	x	
*Calculation Verification - Use Error **	III			x	x	
Validation - Model [Form]	III					x
Validation - Model [Inputs]	III					x
Validation - Comparator [Test Samples]	III					x
Validation - Comparator [Test Conditions]	I					x
Validation - Assessment [Input Parameters]	III					x
*Validation - Assessment [Output Comparison] **	IV					x
Applicability: Relevance of the Quantities of Interest	V					x
Applicability: Relevance of the Validation Activities to the COU	IV					x
Validation - Model [Form]	III			x		
Validation - Model [Inputs]	III			x		
*Validation - Comparator [Test Samples] **	I			x		
Validation - Comparator [Test Conditions]	III			x		
Validation - Assessment [Input Parameters]	V			x		
*Validation - Assessment [Output Comparison] **	V			x		
Applicability: Relevance of the Quantities of Interest	V			x		
*Applicability: Relevance of the Validation Activities to the COU **	IV			x		

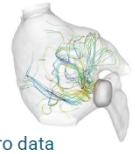
Model Risk for **COU 1**: 4/5
- Decision Consequence: Medium
- Model Influence: High

Model Risk for **COU 2**: 3/5
- Decision Consequence: Medium
- Model Influence: Medium

2 COUs were defined due to the difficulty in conducting in-vitro experiments covering large datasets.

Credibility scores can be improved by:
- Running deeper Uncertainty Analyses
- Enriching experimental datasets used for Validation

COU 1 : QI addressed by simulations only

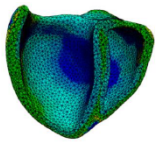


COU 2 : QI addressed by both simulations and in-vitro data

* ongoing V&V activities

UC3 – Credibility Factors Coverage

Model Risk: 3/5
- Decision Consequence: Medium
- Model Influence: Medium



Credibility Factor Coverage Level (per ASME VV40)		1	2	3	4	5
				x		
Code Verification: Software Quality Assurance	III			x		
*Code Verification: Numerical Code Verification - NCV **	IV			x		
*Calculation Verification - Discretization Error **	III			x		
*Calculation Verification - Numerical Solver Error **	III			x		
*Calculation Verification - Use Error **	IV			x		
*Validation - Model [Form] **	III			x		
*Validation - Model [Inputs] **	III			x		
*Validation - Comparator [Test Samples] **	III			x		
*Validation - Comparator [Test Conditions] **	II			x		
*Validation - Assessment [Input Parameters] **	III			x		
*Validation - Assessment [Output Comparison] **	III			x		
*Applicability: Relevance of the Quantities of Interest **	III			x		
*Applicability: Relevance of the Validation Activities to the COU **	II			x		

* ongoing V&V activities

NOTE:
- Low credibility levels are due to scarcity of comparator conditions for EP model (i.e. retrospective clinical study data)
- However: EP conditions for COU are less relevant than sample type (drug) to assess TdP risk

We consider model predictions still sufficiently credible for decision-making

Modular Model:
PK model
0D model
3D model

SCT VVUQ Achievements

VVUQ project deliverables

All documents are public and will be soon available (undergoing editorial review)

Peer-Review papers on specific V&V topics related to Use Cases are also under writing

Verification Report:
7 docs / 231 pages


EU Horizon 2020 Research & Innovation Program
Digital transformation in Health and Care
SC1-DTH-06-2020
Grant Agreement No. 101016496

SimCardioTest - Simulation of Cardiac Devices & Drugs for in-silico Testing and Certification

Technical Report
D 6.1: VERIFICATION & UNCERTAINTY QUANTIFICATION FOR THE USE CASES OF WP2-5

Work Package 6 (WP6)
VERIFICATION, VALIDATION, UNCERTAINTY QUANTIFICATION & CERTIFICATION

Task Lead: UBx, France
WP Lead: MPC, France

Validation Report:
4 docs / 151 pages


EU Horizon 2020 Research & Innovation Program
Digital transformation in Health and Care
SC1-DTH-06-2020
Grant Agreement No. 101016496

SimCardioTest - Simulation of Cardiac Devices & Drugs for in-silico Testing and Certification

Technical Report
D 6.2: VALIDATION OF THE MODEL PREDICTIONS FOR THE USE CASES OF WP2-5

Work Package 6 (WP6)
VERIFICATION, VALIDATION, UNCERTAINTY QUANTIFICATION & CERTIFICATION

Task Lead: MPC, France
WP Lead: MPC, France

Slide Deck 2 – UC1

SIM CARDIO TEST

European Commission | Digital transformation in Health and Care
SC1-DTH-06-2020
Grant Agreement No. 101016496

SimCardioTest

Simulation of Cardiac Devices & Drugs for in-silico Testing and Certification
--
UC1 : pacing leads & catheters

Stuttgart, VPH – September 5th, 2024

Y. Coudière
Université de Bordeaux

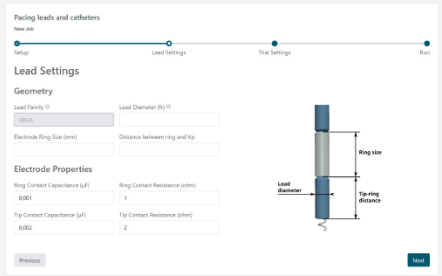
université de BORDEAUX

IHU LIRYC ELECTROPHYSIOLOGY AND HEART MODELING INSTITUTE

Question of interest for electrical capture

SIM CARDIO TEST

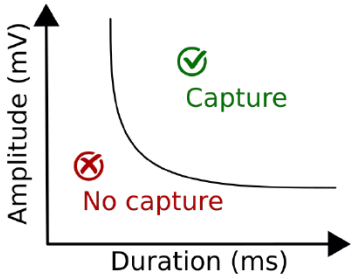
QoI : What are the stimulation pulse characteristics (voltage amplitude in V and pulse duration in ms) required for a bradycardia lead in bipolar (tip/ring) mode to capture (stimulate) healthy cardiac tissue ?



Input : design parameters

Computer model and simulation

Verification and validation

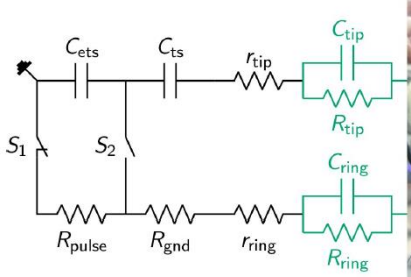


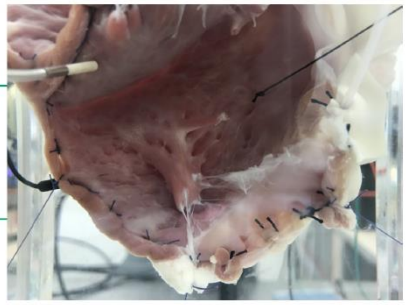
Output : Lapicque curve, etc

CoU : As it would be measured during threshol detection tests during animal experiments

Model Risk: 2/5, Decision Consequence: Low - Model Influence: Medium

Formulation of the 3D model, implemented in CEPS





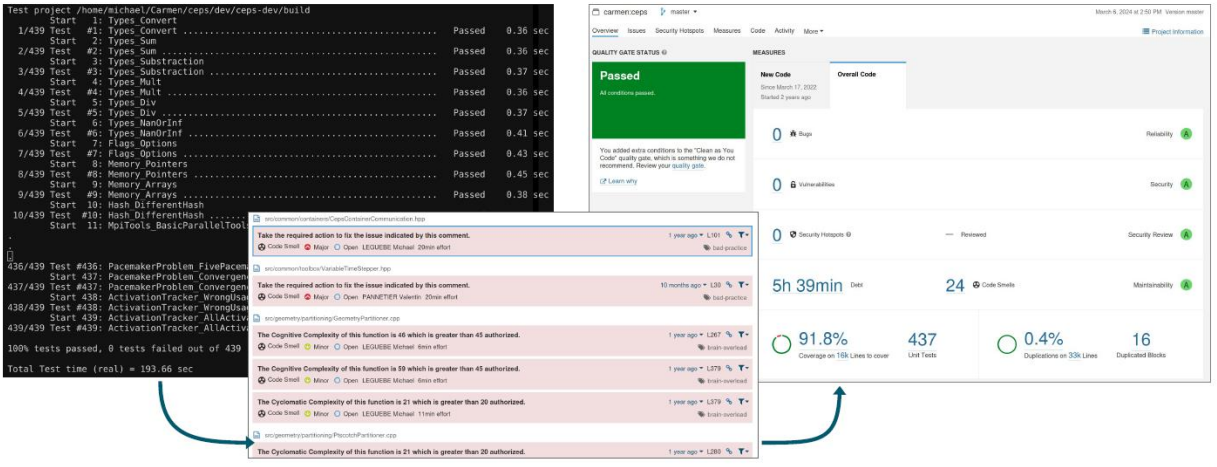
$$\tau_{tip}\frac{dU_{ring}}{dt}+U_{ring}=\tau\frac{dU_{tip}}{dt},$$
$$\tau_{tip}=r_{tip}C\ll\tau=RC,$$
$$R=R_{pulse}+R_{gnd}+r_{ring},$$
$$\frac{1}{C}=\frac{1}{C_{ets}}+\frac{1}{cts}$$

$$-\operatorname{div}(\sigma_i\nabla u_i)=-c_m\partial_t v_m-I_{ion}(v_m,h)\quad\text{in }\Omega_M$$
$$-\operatorname{div}(\sigma\nabla u)=\begin{cases}c_m\partial_t v_m+I_{ion}(v_m,h) & \text{in }\Omega_M\\ 0 & \text{in }\Omega_B\end{cases}$$
$$\partial_t h+g_{ion}(v_m,h)=0\quad\text{in }\Omega_M$$

$$j_i=-\sigma\nabla u\cdot n_i=c_i\partial_t(U_i-u_{|\Gamma_i})+\frac{1}{r_i}(U_i-u_{|\Gamma_i}),\quad i\in\{\text{tip},\text{ring}\}$$

Verification: Software Quality Assurance

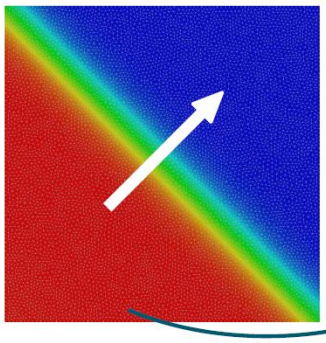
- An unusual effort, especially for small research teams
- Unit tests, with results passed to SonarQube linter

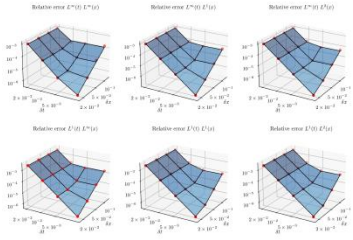


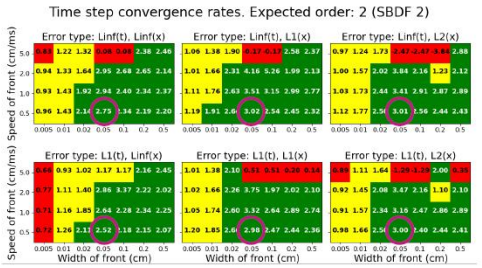
4/12

Verification: Numerical code verification

- Small problems with analytic solution (Laplacian, Heat…)
  - Manufactured solution that share properties of cardiac APs: pseudo wavefront propagation (i.e. heat equation with provided specific source term)
  - Error is easy to measure in that case, CEPS runs an automatic convergence study with few additional parameters







Errors wrt time step, mesh

Convergence rates vs solution profile

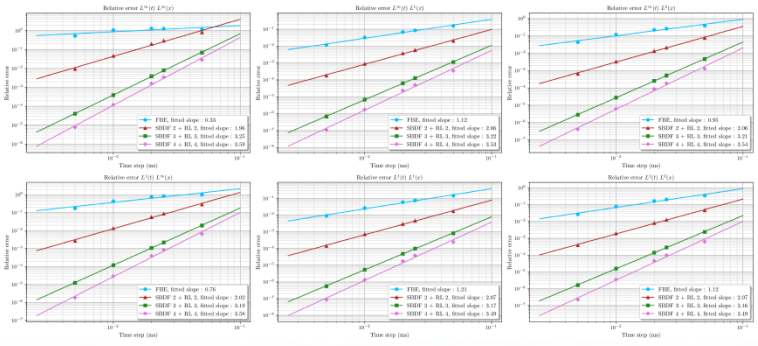
○ : cardiac context

5/12

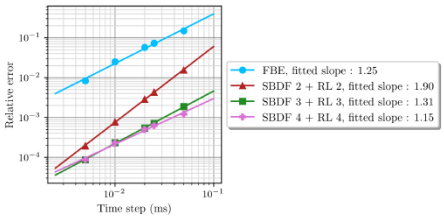
Verification: Numerical errors

- Cardiac problems w/o analytic solution:
  - Error is measured wrt to a reference solution (fine mesh, very small time step)
  - Limited to continuous ionic models (Beeler-Reuter is ok, but not Ten Tuscher !) and smooth volumic stimulations to get high convergence orders.
  - Convergence can also be checked on post-treatment of solution: activation map
- Bidomain results

Convergence of solution (potentials)



Convergence of activation map
(order 1 due to simplistic detection of AP)



6/12

Verification: Numerical errors

- Influence of numerical parameters on Lapicque curve
  - limitation: pacemakers introduce discontinuity => convergence ?
  - very restrictive linear solver tolerances for evaluation of high order methods : sometimes the solver does not converge...

Thresholds for SBDF 2 time scheme, rtol=1.e-12, rmesh=1			
stim duration	dt = 0.01	dt = 0.005	dt = 0.002
0.25	3.00	3.00	2.50
0.50	2.00	2.00	1.50
1.00	1.25	1.25	1.00
1.50	1.00	1.00	1.00
2.00	1.00	1.00	0.75

our "usual" 10 µs time step

Thresholds for FBE time scheme, rtol=1.e-12,rmesh=1.			
stim duration	dt = 0.01	dt = 0.005	dt = 0.002
0.25	2.50	2.50	2.50
0.50	1.50	1.50	1.50
1.00	1.00	1.00	1.00
1.50	1.00	1.00	1.00
2.00	0.75	0.75	0.75

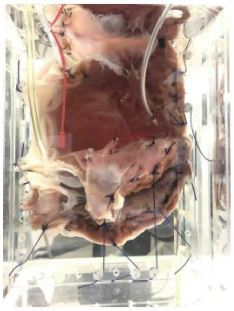
Thresholds for FBE time scheme, dt=0.01, rmesh=1.				
stim duration	rtol=1e-4	rtol=1e-6	rtol=1e-8	rtol=1e-10
0.25	0.25	3.00	3.00	3.00
0.50	0.25	2.00	2.00	2.00
1.00	1.25	1.25	1.25	1.25
1.50	0.75	1.00	1.00	1.00
2.00	1.00	1.00	1.00	1.00

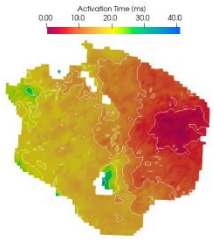
Thresholds for FBE time scheme, rtol=1.e-12, dt=0.01			
stim duration	rmesh=1	rmesh=1.5	rmesh=2
0.25	2.50	2.50	2.50
0.50	1.50	1.50	1.50
1.00	1.00	1.00	1.00
1.50	1.00	1.00	1.00
2.00	0.75	0.75	0.75

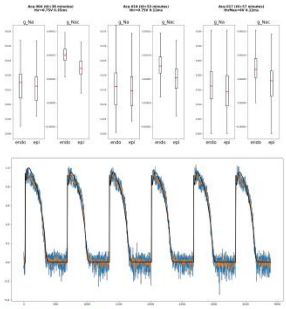
rmesh = 1.0 : 269k cells
rmesh = 1.5 : 800k cells
rmesh = 2.0 : 1,762k cells

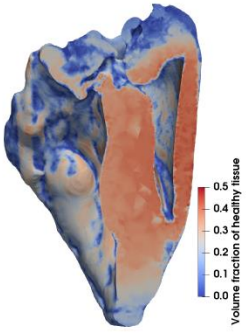
7/12

Validation activities









- Contact parameters → done, bath experiments (cf D2.3)
- Ionic parameters → optical signals, pixel-wise, done, sheep #1 and #3
- Conductivity coefficients → 9.4T MRI (~300 µm) + activation + activation maps, on-going
- Still processing data

We could design and run our own experiments

Validation metric

- Lapicque curves close to experimental Lapicque regions

Validation: test sample and test conditions

SIM CARDIO TEST

Plan of the experimental activities

- 4 healthy + 4 scar animals
- 3 locations, moreless close to the scar center
- 2 leads (Microport and a concurrent)

Reality issues, intrinsic complexity of experiment on living samples

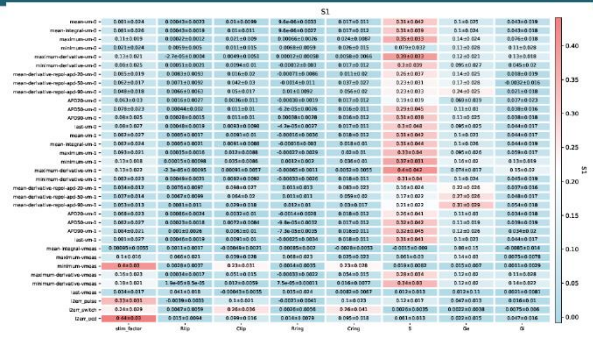
- Animal dies unexpectedly
- Leads brake → same model, but not same device
- Design for the concurrent lead not known

Post processing difficulties

- Quality of the signals
- Multi-modality, multi-resolution images (structure, morphology, electrical function)

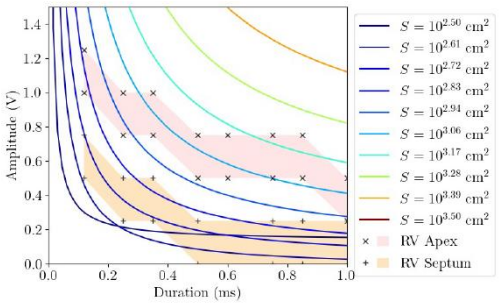
Validation : sensitivity, comparison

SIM CARDIO TEST



Sobol indices

output used to detect capture w/r model parameters



Comparator : Lapicque curve

Computational limitation

- Currently, doable only for a simplified 0D model
- In any case, surrogate modelling seems to be needed (complex PDE model with many parameters)

Applicability

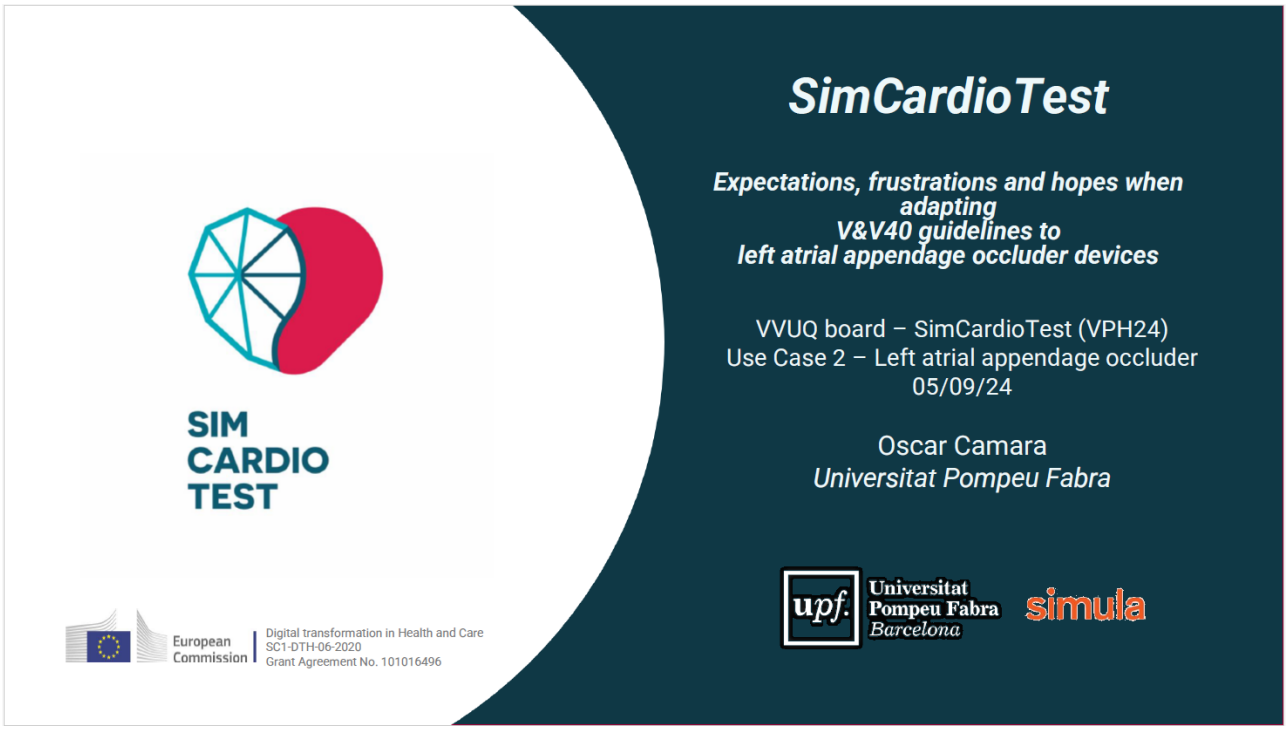
Limitation - Context of use is animal experiments (pre-clinical)

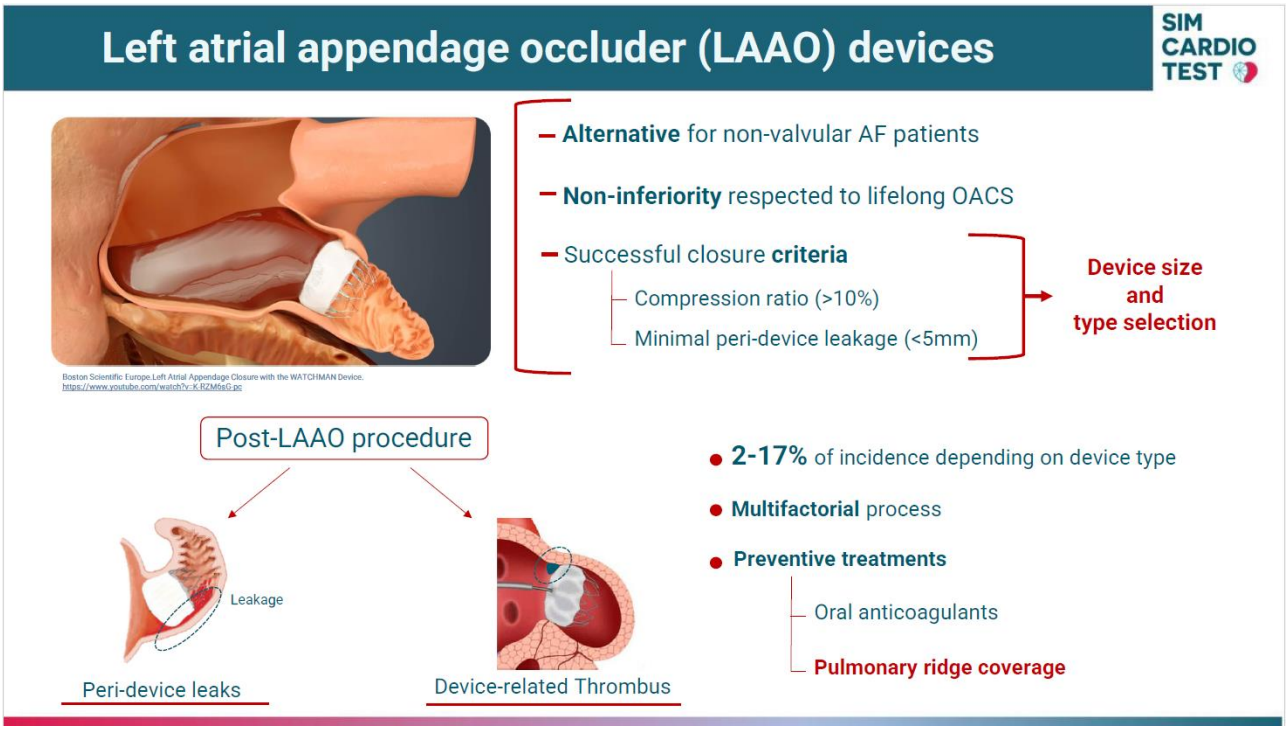
- Small number of amplitudes / voltages : technical limitation of the device → incomplete answer to the QoI ?
- Small animal sample : adapted to the reproducible animal model

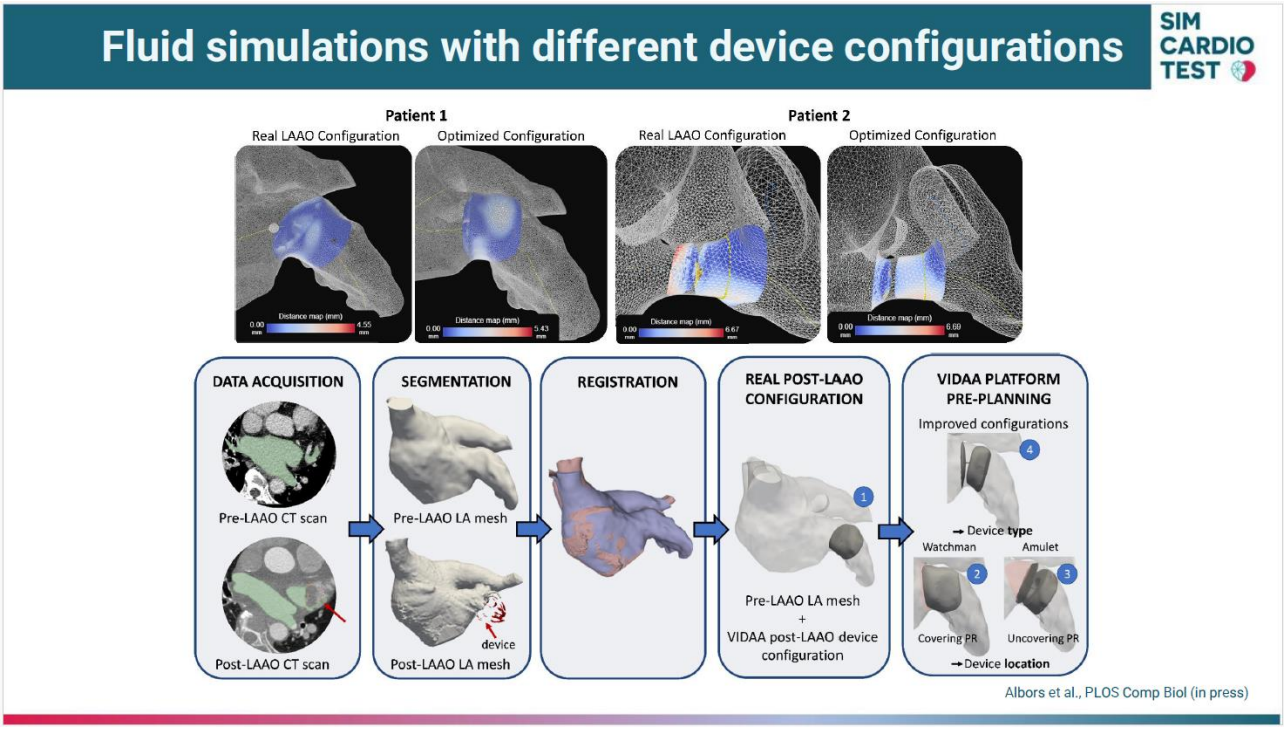
And for in-silico trials ?

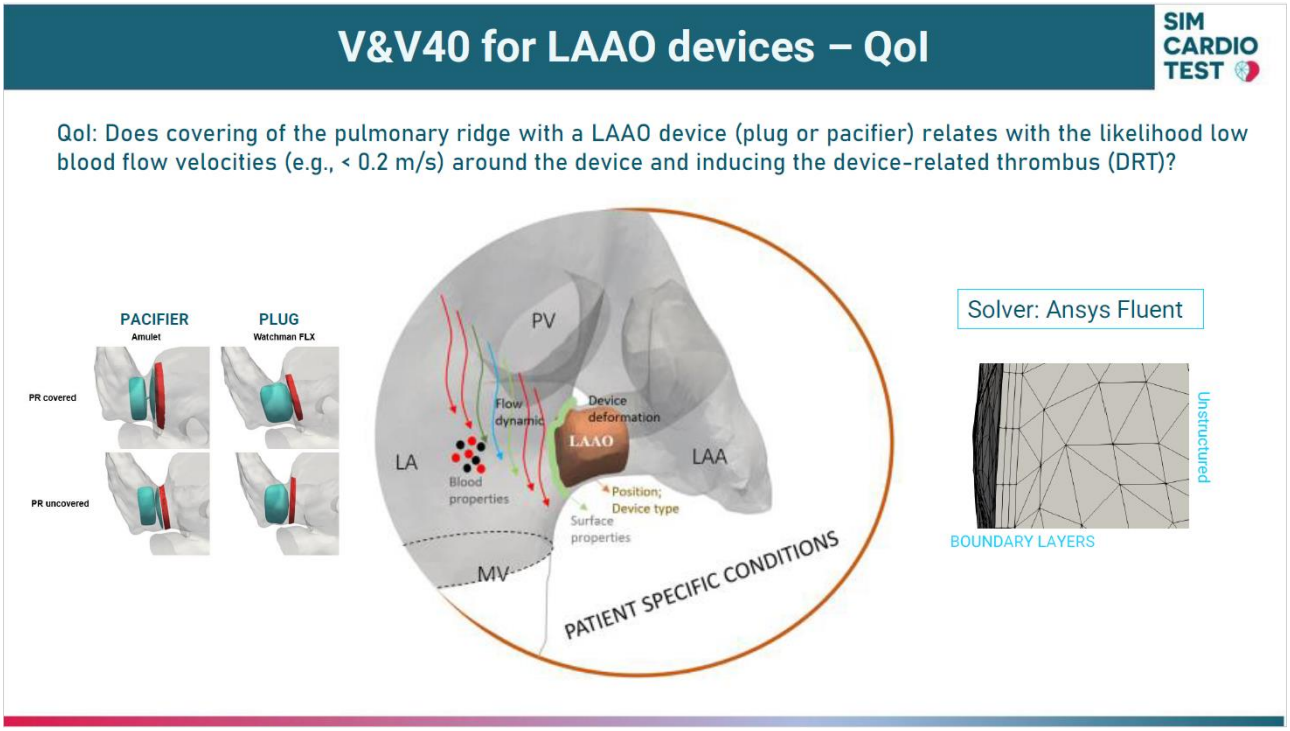
- Translation to human: sheep is a well known replacement for human, our facilities has access to a well defined, reproducible model
- Extrapolation to human values:
  - values for conductivity coefficients and ionic conductances from the literature
  - optical data and 9.4T MR images of human samples, available at Liryc (obtained in a different context)
  - experiment on human samples

Slide Deck 3 – UC2






Fluid simulations with different device configurations
SIM CARDIO TEST
Patient 1
Real LAAO Configuration
Optimized Configuration
Patient 2
Real LAAO Configuration
Optimized Configuration
DATA ACQUISITION
Pre-LAAO CT scan
Post-LAAO CT scan
SEGMENTATION
Pre-LAAO LA mesh
device
Post-LAAO LA mesh
REGISTRATION
REAL POST-LAAO CONFIGURATION
Pre-LAAO LA mesh
+
VIDAA post-LAAO device configuration
VIDAA PLATFORM PRE-PLANNING
Improved configurations
Device type
Watchman
Amulet
Covering PR
Uncovering PR
Device location
Albors et al., PLOS Comp Biol (in press)


V&V40 for LAAO devices – QoI
SIM CARDIO TEST
QoI: Does covering of the pulmonary ridge with a LAAO device (plug or pacifier) relates with the likelihood low blood flow velocities (e.g., < 0.2 m/s) around the device and inducing the device-related thrombus (DRT)?
PACIFIER
Amulet
PLUG
Watchman FLX
PR covered
PR uncovered
PV
Flow dynamic
Device deformation
LAAO
LA
LAA
Blood properties
Position; Device type
Surface properties
MV
PATIENT SPECIFIC CONDITIONS
Solver: Ansys Fluent
Unstructured
BOUNDARY LAYERS

V&V40 for LAAO devices – COU & Model Risk

COU1 - Performance evaluation with computational fluid simulations only: Computational modeling is used to identify low blood flow velocities near the device, placed in a proximal or distal position (e.g., covering or not the PR) with both device types (i.e., plug and pacifier). There is no supporting data from in-vitro testing available for assessing the performance of the occluder devices.

COU2 - Performance evaluation with computational fluid simulations and in-vitro data: In addition in-silico experiments, in-vitro testing is conducted to create additional evidence on whether the covering of the PR is critical for DRT with both types of device.

Model Influence			
High	3	4 (CoU1)	5
Medium	2	3 (CoU2)	4
Low	1	2	3
	Low	Medium	High

Decision consequence

V&V40 for LAAO devices – Quantities of Interest

Quantities of Interest (QoI)

To evaluate the quantities depending on the question of interest, first it is important to select the zone and time of interest.

By zone:

- Region of 5 mm surrounding the implanted device (see Figure 1)
- Region of LAA corresponding to cases without devices
- Output region. Control flow and velocities around the MV
- Inlet regions and centre of LA to control the pressures.

By beat time:

- Total beat
- Systole
- Diastole
- 5-10 beats

The **quantities of interest** are listed below:

QOI 1:

Average and maximum of blood flow (Q)
Average and maximum velocities
Endothelial cell activation potential (ECAP)
Platelet adhesion zone
Particles tracking
HR
Cardiac output (CO)
Lambda 2
Q-criterion

Quantity	Abbreviation/Symbol	Definition	Unit
Time averaged wall shear stress	TAWSS	$\frac{1}{T}\int_0^T \lvert\tau\rvert\, dt$	[Pa]
Oscillatory shear index	OSI	$\frac{1}{2}\left(1-\frac{\left\lvert\int_0^T \tau\, dt\right\rvert}{\int_0^T \lvert\tau\rvert\, dt}\right)$	[-]
Relative residence time	RRT	$\frac{1}{(1-2\cdot \mathrm{OSI})\cdot \mathrm{TAWSS}}$	[1/Pa]
Endothelial cell activation potential	ECAP	$\frac{\mathrm{OSI}}{\mathrm{TAWSS}}$	[1/Pa]

V&V40 for LAAO devices – Ansys Fluent code verification

SIM CARDIO TEST

ANSYS FLUENT

Software Quality Assurance (SQA) and Numerical Code Verification (NCV)

- Benchmark test cases replicate previously established results and comparing the error with analytical solutions → level of credibility appropriate for COUs

- Ansys fluid dynamics verification manual
  https://dokumen.tips/reader/f/ansys-fluid-dynamics-verification-manual

- Ansys certifications (highest levels of NVC)
  https://www.ansys.com/company-information/quality-assurance
  https://www.ansys.com/content/dam/company/quality-assurance/2021-2024-ansys-inc-iso-9001-2015-dqs.pdf

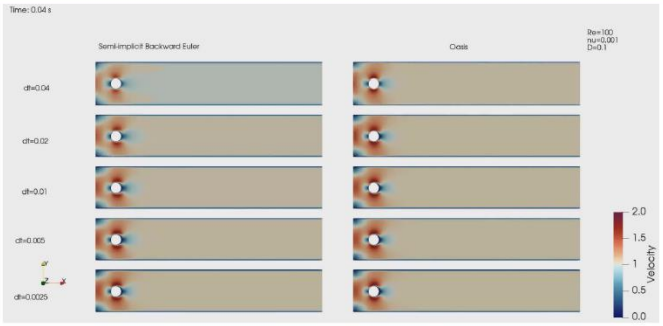
- Using Ansys (provides the necessary insights to test and optimize new products), help accelerate regulatory approval but medical approaches under an elevated risk may need additional benchmarks
  - Comparison with Oasis numerical solution in benchmark tests
  - Quantification of discretisation error with mesh convergence and accuracy order studies

V&V40 for LAAO devices – space/time convergence studies

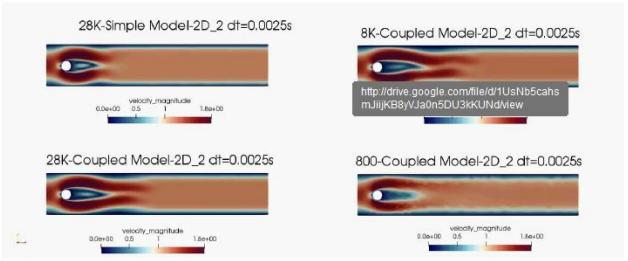
SIM CARDIO TEST

simula

Comparison of Ansys-like and Oasis solvers with different time-steps



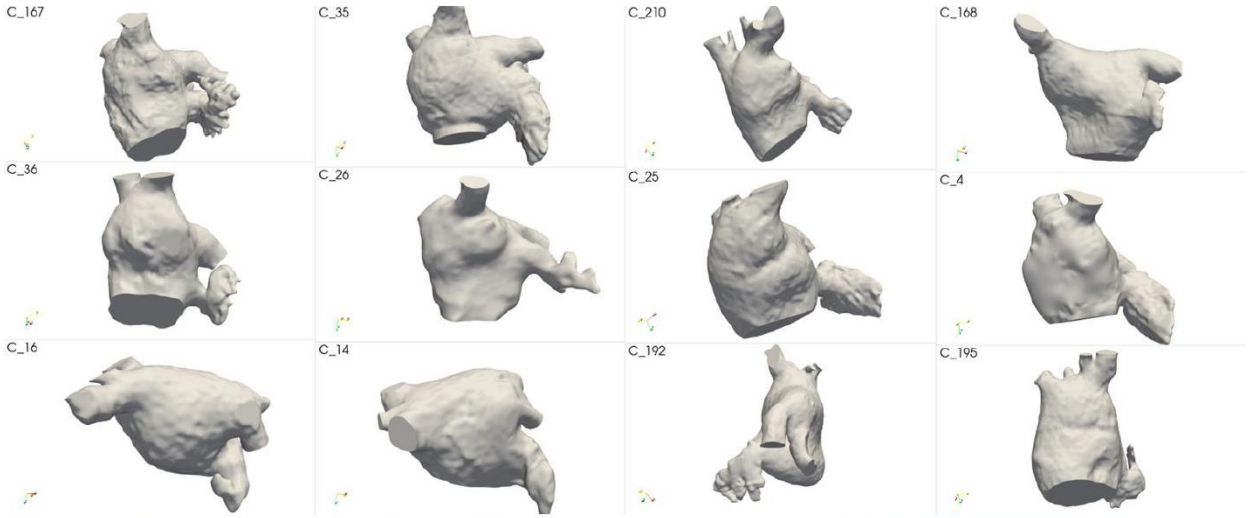
Ansys results with different spatial resolutions and models



V&V40 for LAAO devices – clinical data for verification

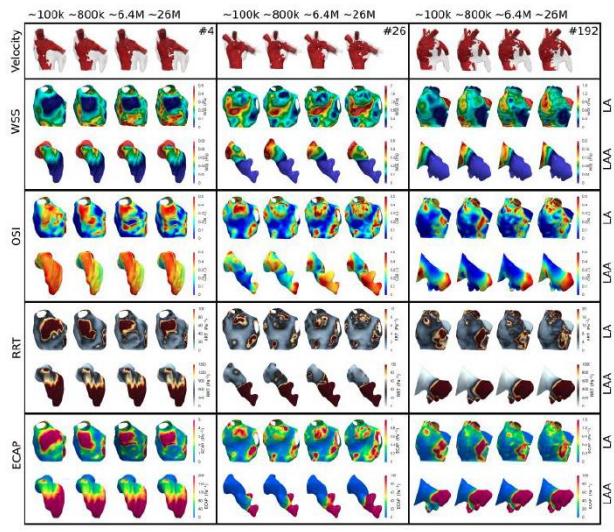
SIM CARDIO TEST

- 12 models selected for study

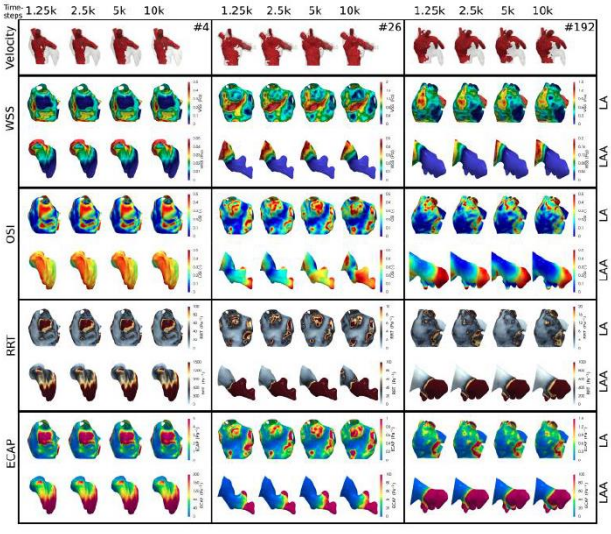


V&V40 for LAAO devices – convergence studies

SIM CARDIO TEST

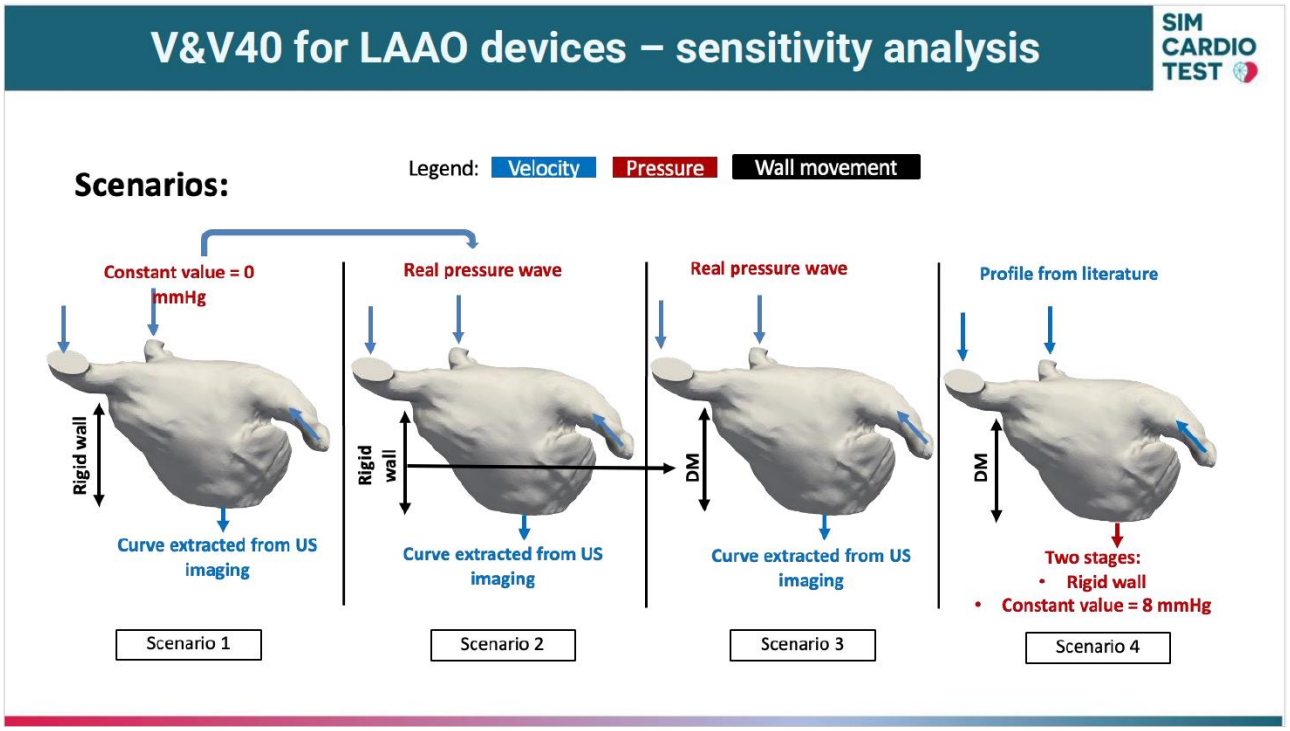


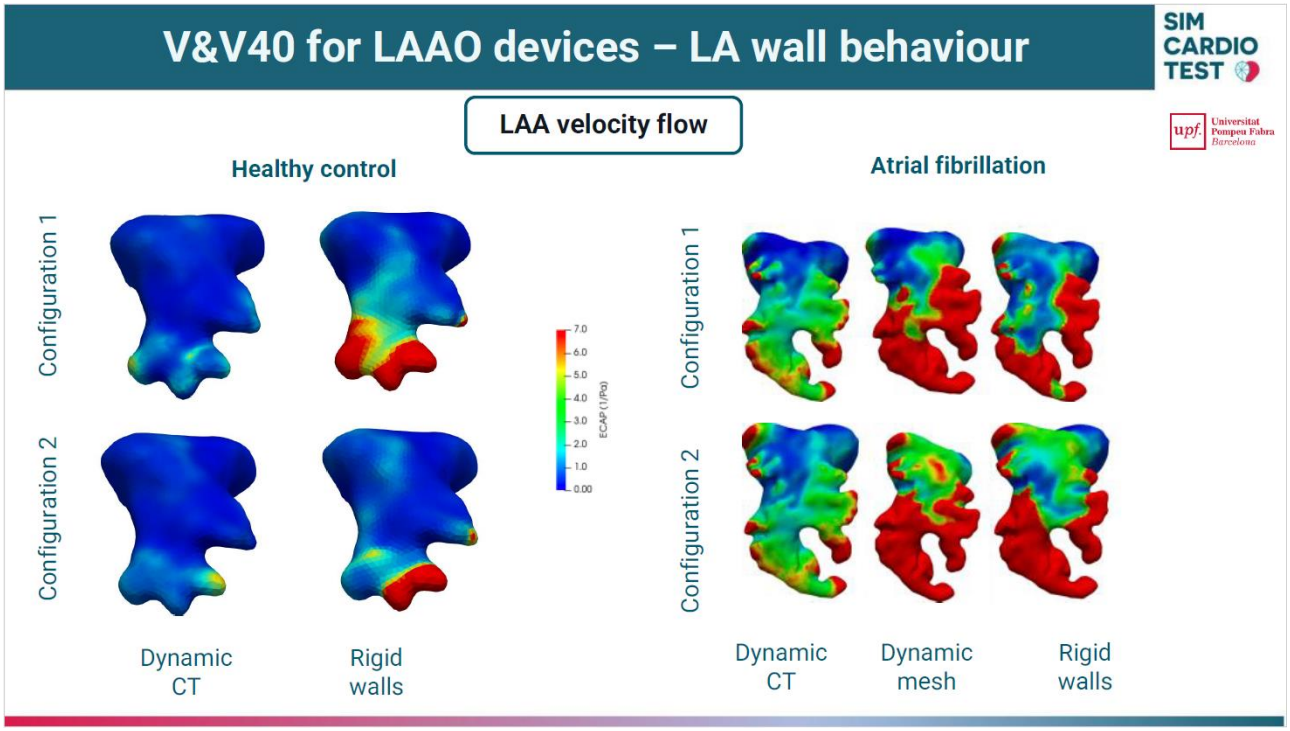
Mesh convergence



Time-step convergence

Khalili et al., Int J Numer Method Biomed Eng 2024


V&V40 for LAAO devices – sensitivity analysis
SIM CARDIO TEST
Scenarios:
Legend: Velocity Pressure Wall movement
Constant value = 0 mmHg
Rigid wall
Curve extracted from US imaging
Scenario 1
Real pressure wave
Rigid wall
Curve extracted from US imaging
Scenario 2
Real pressure wave
DM
Curve extracted from US imaging
Scenario 3
Profile from literature
DM
Two stages:
• Rigid wall
• Constant value = 8 mmHg
Scenario 4


V&V40 for LAAO devices – LA wall behaviour
SIM CARDIO TEST
upf. Universitat Pompeu Fabra Barcelona
LAA velocity flow
Healthy control
Atrial fibrillation
Configuration 1
Configuration 2
ECAP (1/Pa)
Dynamic CT
Rigid walls
Dynamic CT
Dynamic mesh
Rigid walls

V&V40 for LAAO devices – simulations with devices

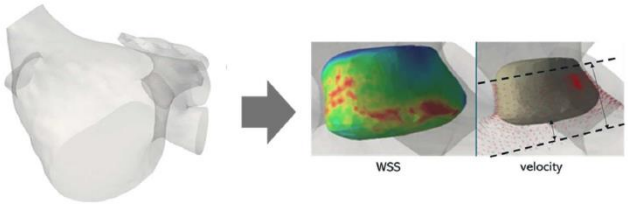


Table 8: Time convergence study for a plug-type occluder device. WSS: wall shear stress.

Time Step convergence Δt [s]	0.1	0.05	0.01	0.005	0.001
with Plug Device (Δx = 1.83)					
Average WSS [Pa]	0.9823	1.3514	1.655	1.6685	1.680
Errors based on WSS [%]	41%	19.64%	1.4%	0.6%	-
Average velocity magnitude [m/s]	0.09511	0.06532	0.0721	0.0733	0.07723
Errors based on velocity [%]	23%	15%	6.6%	5%	-

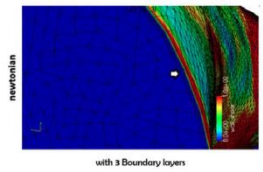
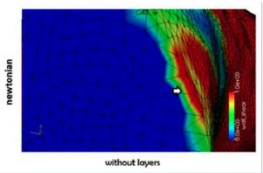


Table 9: Mesh convergence study for a plug-type occluder device. WSS: wall shear stress.

Mesh Convergence h =Δx [mm]	3.17	2.44	1.83	1.55
with Plug Device (simulation→ Δt = 0.01s)	150k	320k	750k	1.5M
Avg. WSS [Pa]	1.040	1.204	0.853	0.892
Errors [%]	17%	34.5%	4.5%	-
Avg. Velocity Magnitude [m/s]	0.0475	0.0778	0.0721	0.0732
Errors [%]	35%	6.2%	1.5%	-



Validation with in-vitro experiments

SEGMENTATION → 3D Printing → MOCKUP

UPF PROTOCOLS

- CT images
- Source CT system
- Isotropic voxel sizes (e.g., ranged between 0.37-0.5 mm; 512 x 512 x [270-403] slices
- Using region growing and thresholding in 3D-slicer software
-Taubin smoothing

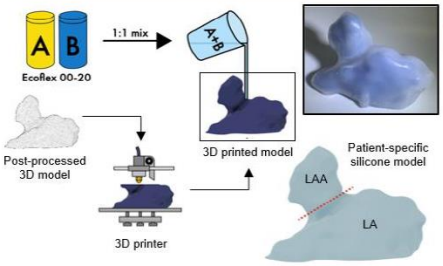
MIT PROTOCOLS

- Object Connex 500
- Silicone casting (Ecoflex00-20)
- Pneumatic artificial muscles (PAMs)

MIT PROTOCOLS

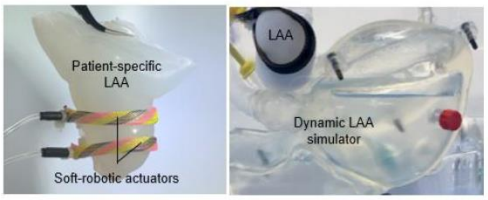
-Pressure measured in the LA/LAA
- Flow inlet was controlled
-Flow velocities can be measured with US or PVI

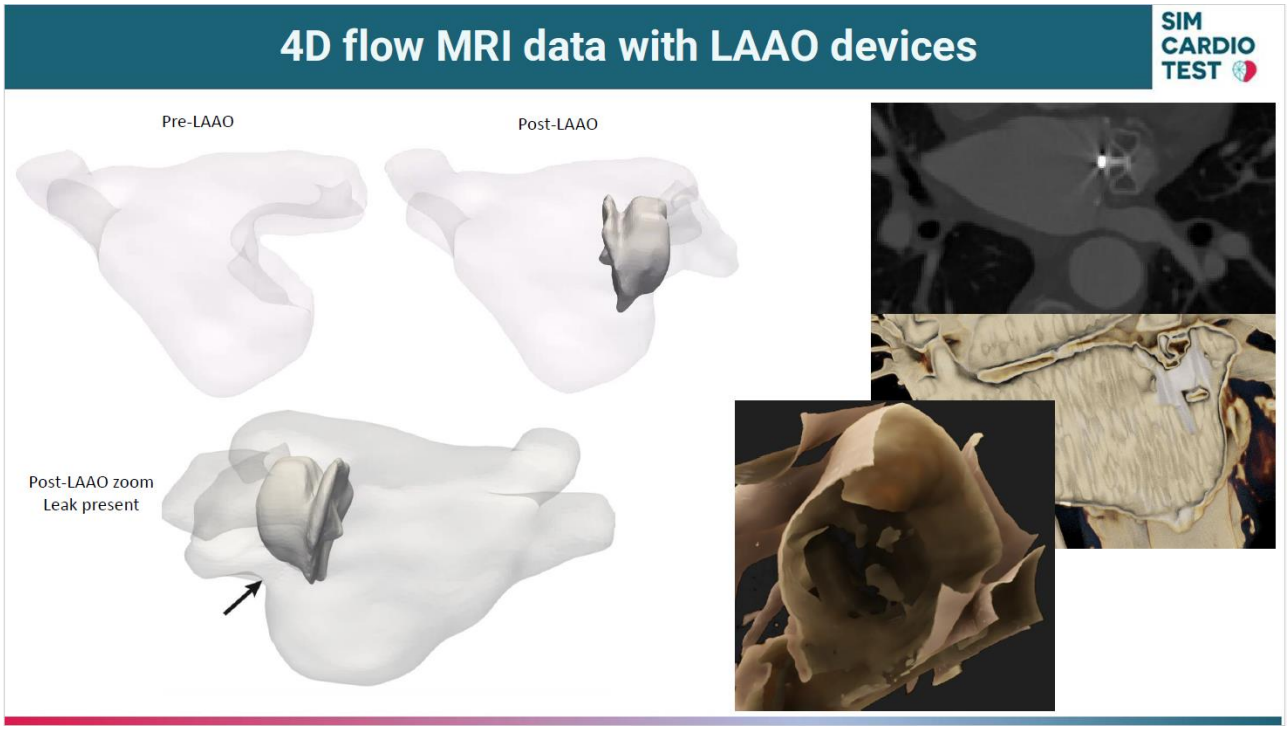
3. 3D Printing & Silicone Casting

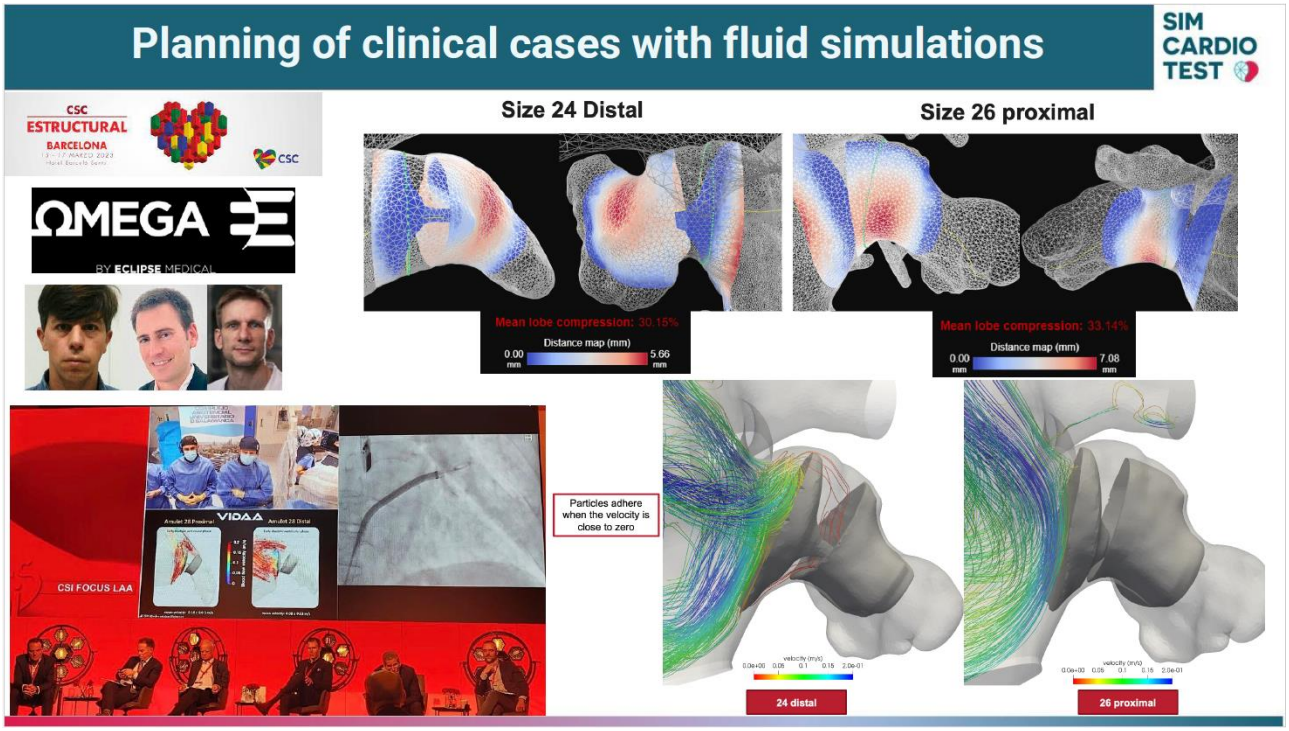


Massachusetts Institute of Technology
Harvard–MIT Health Sciences & Technology
imes
MITMECHE

4. Insertion into Dynamic Simulator




4D flow MRI data with LAAO devices
SIM CARDIO TEST
Pre-LAAO
Post-LAAO
Post-LAAO zoom
Leak present

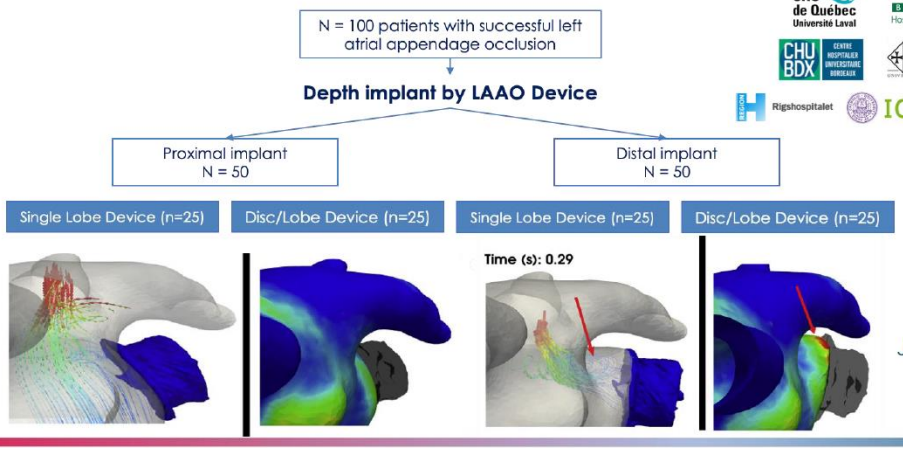

Planning of clinical cases with fluid simulations
SIM CARDIO TEST
CSC ESTRUCTURAL BARCELONA
OMEGA
BY ECLIPSE MEDICAL
Size 24 Distal
Size 26 proximal
Mean lobe compression: 30.15%
Distance map (mm)
0.00
5.66
Mean lobe compression: 33.14%
Distance map (mm)
0.00
7.08
VIDAA
CSI FOCUS LAA
Particles adhere when the velocity is close to zero
24 distal
26 proximal

Real in silico trial study → IDEAL study

SIM CARDIO TEST

Objectives:

- Differences in flow dynamics between proximal and distal implant.
- Differences in flow dynamics between single lobe vs disc/lobe devices in distal implant.
- Differences in flow dynamics between DRT vs no-DRT.



Cepas-Guillén et al., JACC Cardiovasc Interven 2023

Expectations, frustrations, hopes of V&V40 on LAAO devices

SIM CARDIO TEST

Expectations

- Required for clinical/industrial translation of models
- Valid exercise for any computational technique (V&V70 for AI)
- Certain common ground with GDPR and AI-focused EU manifestos
- Trust from clinical and industrial partners

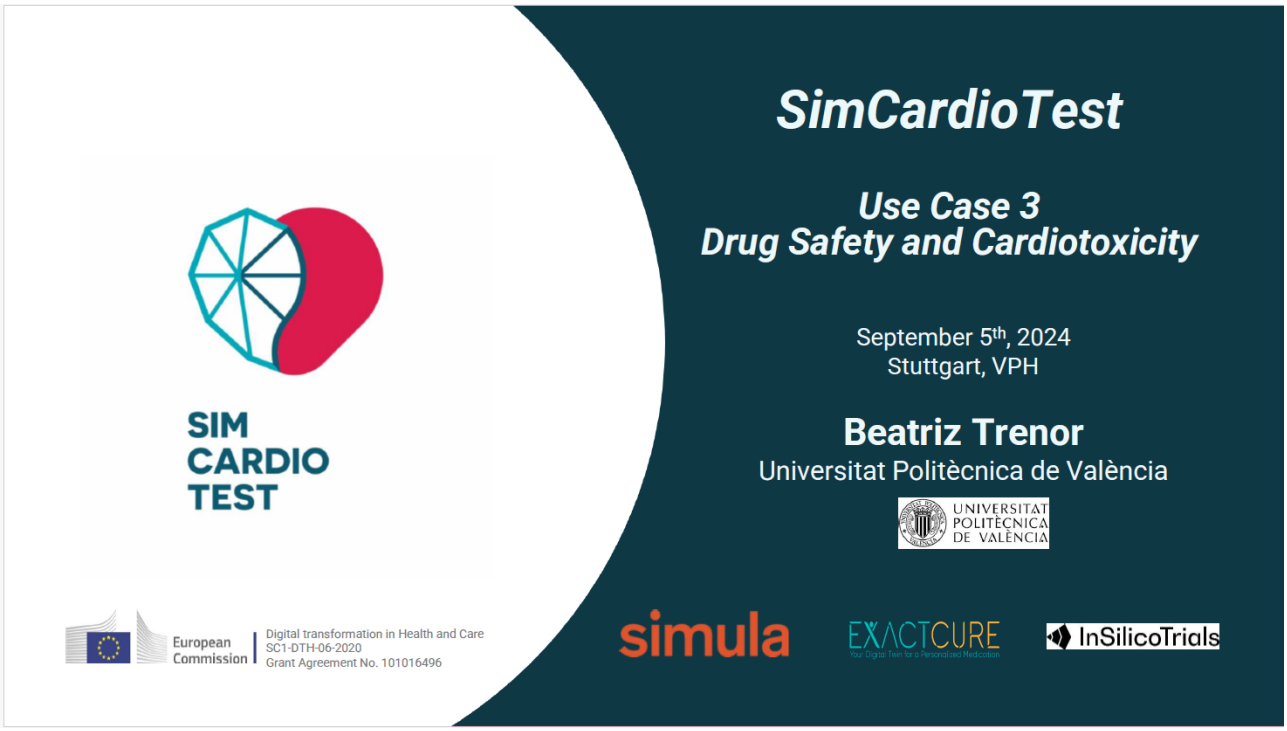
Frustrations

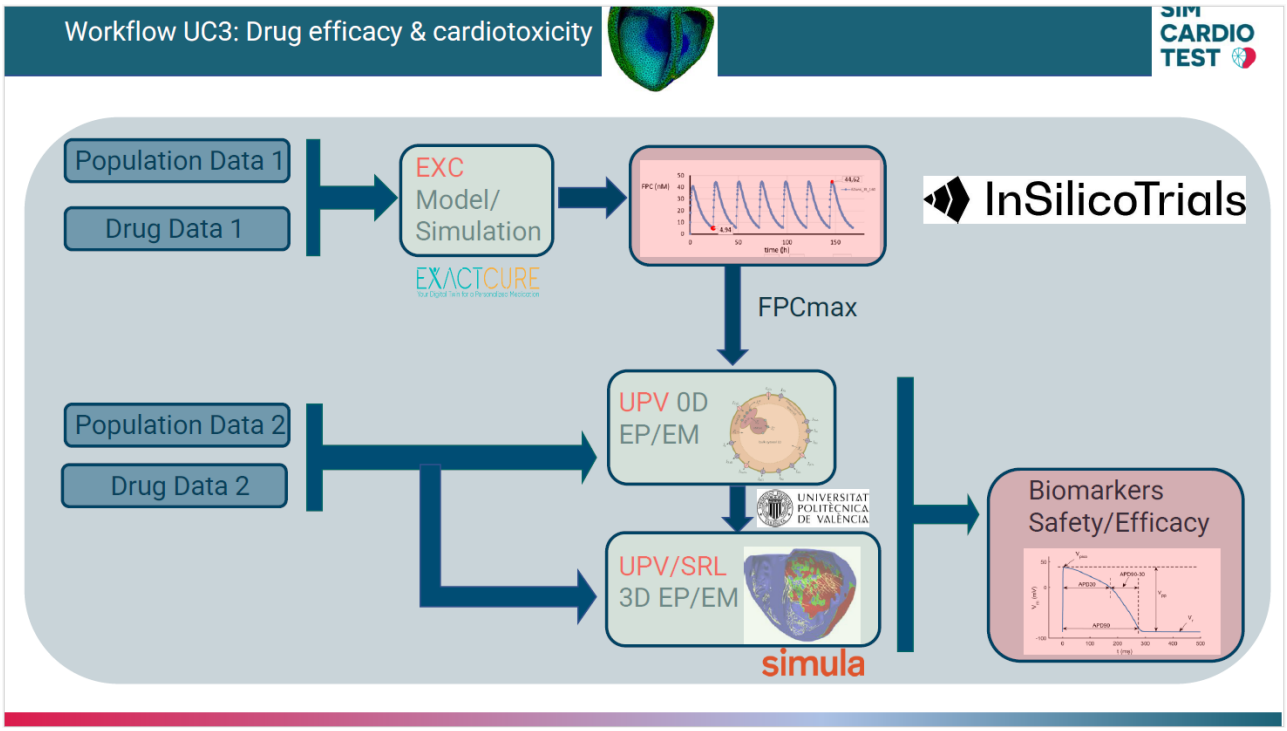
- Not the most exciting tasks, it can be endless; difficult to publish (but useful)
- Patient-specific meshes & personalised boundary conditions
- Computational times/infrastructure for translation to clinical routine
- Data availability/sharing
- Trust from clinical and industrial partners

Hopes

- Benchmarks and standards for (LA) fluid simulations; data/code availability/sharing
- Integration with machine(deep) learning / reduced-order methods

Slide Deck 4 – UC3





Question of Interest

SIM CARDIO TEST

- What is the maximum concentration/dose regimen of a drug to assure TdP-related safety in a population of healthy subjects?

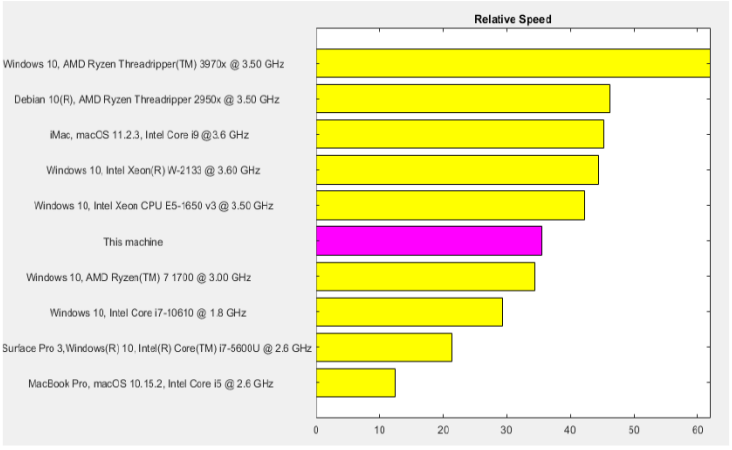
Context of Use
Perform in-silico trials that guide in selecting drugs and doses without TdP-risk to improve clinical trials

Model influence	high	3	4	5
	medium	2	3 COU	4
	low	1	2	3
		low	medium	high
		Decision consequence		

Code Verification

SIM CARDIO TEST

SOFTWARE QUALITY ASSURANCE



MATLAB®

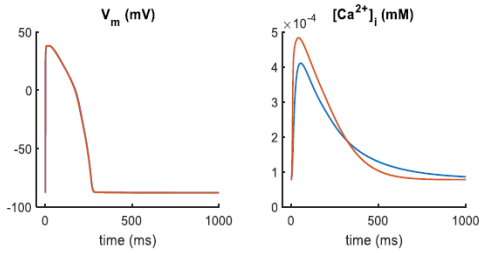
2021b Release run on local machine:

- Windows 10
- Intel® CORE™ i7-6700 @3.4 GHz
- x64 processor
- 8 logical cores
- 32 GB RAM

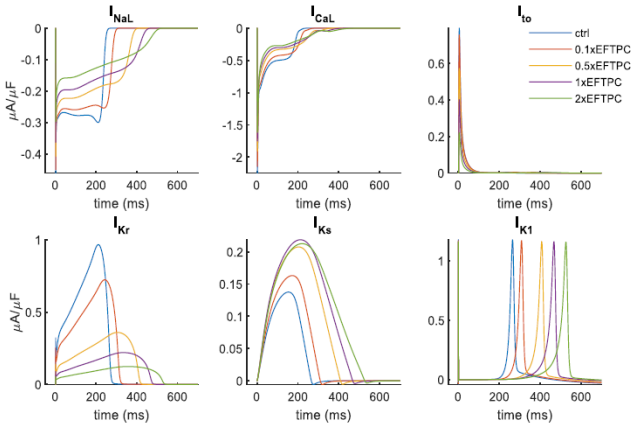
- Benchmark tasks

Code Verification

Numerical Code verification



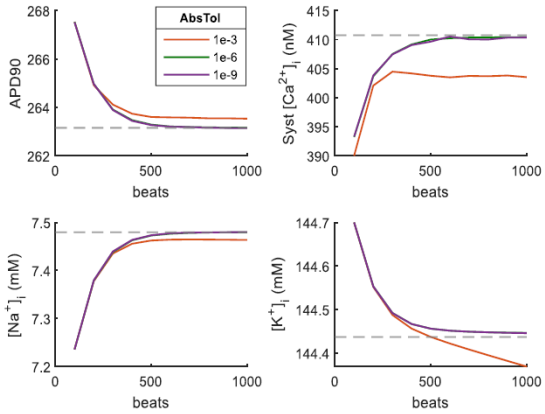
- Qualitative assessment of cellular signals over time
- Biomarkers in range



- Modulation of ion currents with a drug model (example with Quinidine)

Calculation Verification

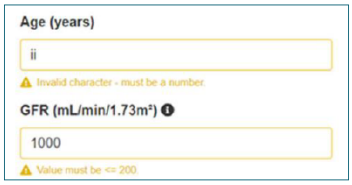
Discretization and numerical solver error



- Effect of tolerance error on convergence

Use error

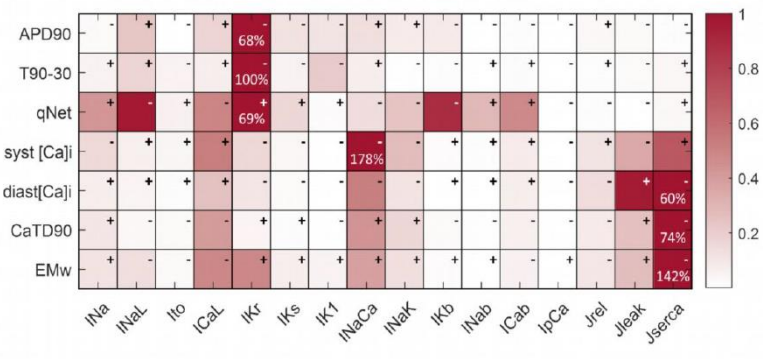
- Internal peer-review
- Typographical errors controlled in the IST platform

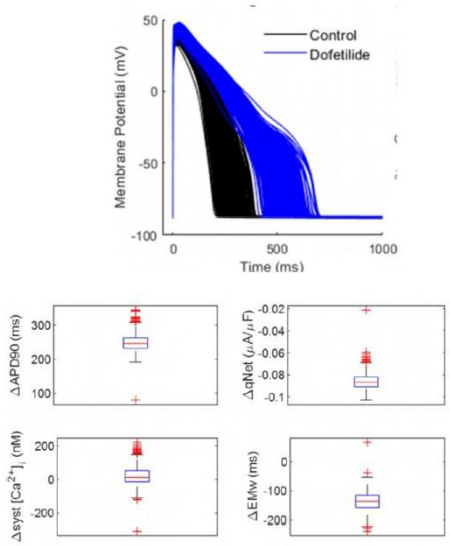


Type	ode15s
Step size	variable
Absolute tolerance	1e-6
Relative tolerance	1e-3

Validation. Computational model

MODEL FORM AND INPUTS

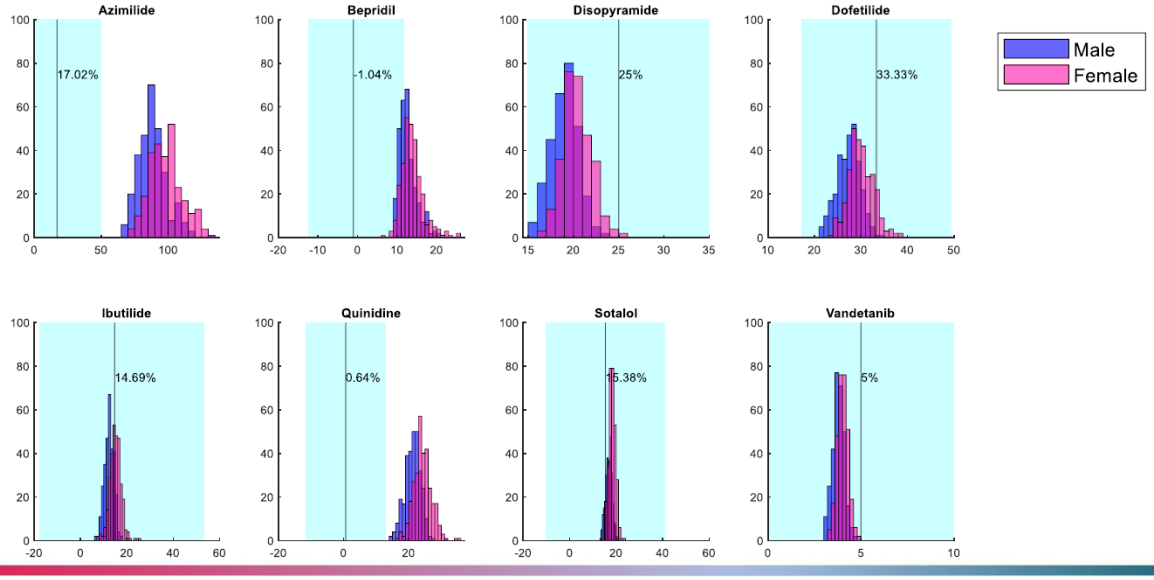




Validation. Comparator & Assessment

AGREEMENT WITH COMPARATOR

High-risk drugs % APD variation

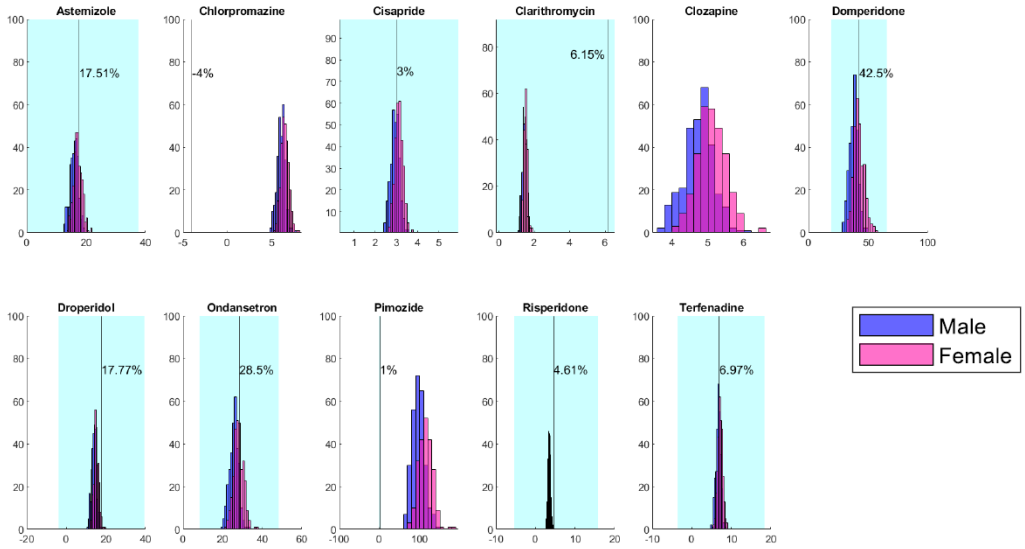


Validation. Comparator & Assessment

SIM CARDIO TEST

AGREEMENT WITH COMPARATOR

Intermediate-risk drugs

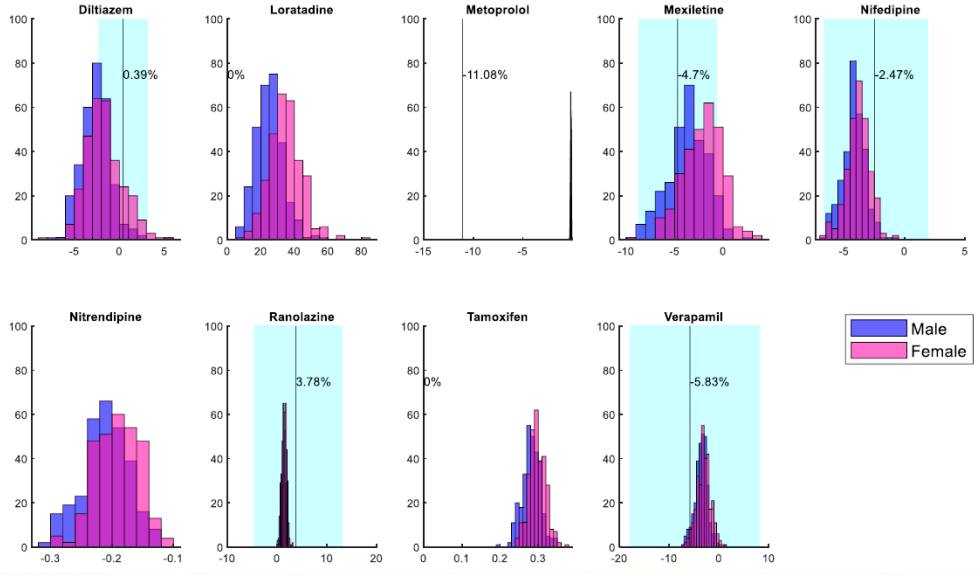


Validation. Comparator & Assessment

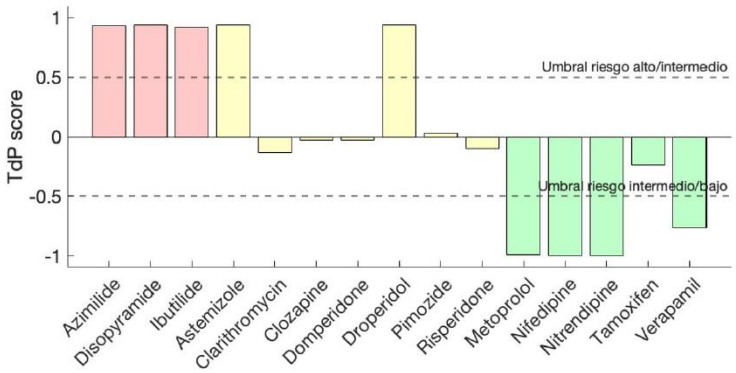
SIM CARDIO TEST

AGREEMENT WITH COMPARATOR

Low-risk drugs

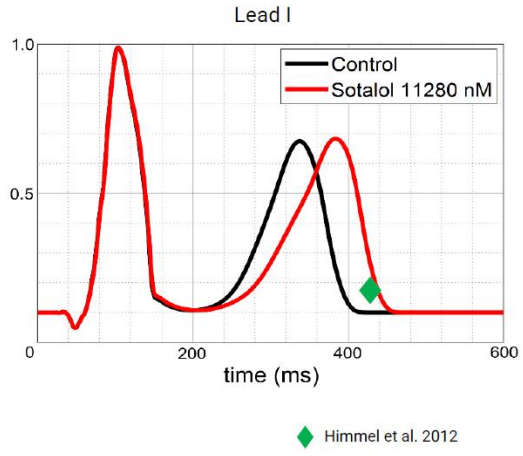


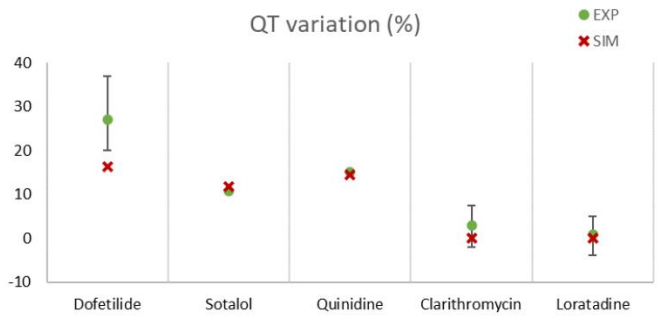
Validation. Classification accuracy



	High	Int.	Low
Pred. High	3	2	0
Pred. Int.	0	5	1
Pred. Low	0	0	4
Accuracy:	80%		
MCE:	0.2		

Validation. 3D simulations.





Male vs Female ?

Applicability

Available data limits the applicability of the model to:

- Assessment of the general population, without differentiating according to sex.

- TdP-Risk classification of drugs according to the EFTPC

Appendix 3 – Minutes from ASME VV40 Meeting on September 5th, 2024

Introduction

The following appendix compiles the minutes of the meeting held between SimCardioTest WP6 stakeholders and some ASME VV40 members in the frame of SimCardioTest D6.4 activities.
The intent of the meeting was to get some constructive feedback on the pertinence of WP6 interpretation of the VV40 standard, and to give some feedback to VV40 on the standard usability, based on the SCT experience.

All opinions expressed in this section by ASME VV40 committee members are their personal opinions and do not represent the official perspective of the ASME organisation nor of their respective companies.

Meeting Details

- Date: Thursday, September 5th, 2024
- Location: Stuttgart, Keplerstrasse 17, room 10.017
- Time: 9:00am – 10:30
- (discussion continued up to 11:15)

Attendees

From SimCardioTest consortium:

- Romano SETZU (MicroPort CRM, France; SCT WP6 leader)
- Yves COUDIÈRE (University of Bordeaux, France; SCT Use-Case 1 leader)
- Oscar CAMARA (Universitat Pompeu Fabra, Barcelona, Spain; SCT Use-Case 2 leader)
- Beatriz TRENOR (Universitat Politècnica de València, Spain; SCT Use-Case 3 leader)
- Michele BARBIER (INRIA Sophia-Antipolis; SCT project coordinator)
- Artem PLATONOV, Liesbet Geris (VPHi correspondent)
- Alessia BARETTA (InSilicoTrials, Italy; SCT member)
- Javier Saiz (Universitat Politècnica de València, Spain; SCT Use-Case 3 researcher)

From ASME VV40 committee:

- Jeffrey BISCHOFF (Zimmer Biomet; chair, VVUQ40)
- Payman AFSHARI (Depuy Synthes; co-vice chair, VVUQ40)
- Walter OCAMPO (Straumann; member, VVUQ40)

NOTE: Shiny MARTIS (Voisin Consulting Life Science; member, VVUQ40) was unable to join the meeting; however, she contributed to the discussion by providing answers to some specific questions (see Follow Up Discussions section below).

Meeting Agenda

- General Introduction (Romano SETZU)
- UC1 – V&V approach and results (Yves COUDIERE)
- UC2 – V&V approach and results (Oscar CAMARA)
- UC3 – V&V approach and results (Beatriz TRENOR)
- Questions / Discussion (All)

Short VVUQ40 Introduction

Mr. Bischoff's introduction on VVUQ40 committee and document.
The committee currently counts about a hundred participants, of whose about 20-30 people are actively involved.
The standard is currently under revision and will be probably released under the next couple of years.
The revision draft will be circulated maybe next year.
In addition, a couple of Technical Documents related to VVUQ40 are currently under draft, and will be circulated, including a technical document containing full worked-out examples of credibility assessment according to VVUQ40 guidelines.
The standard has some known limitations. For instance, it is currently too much centred on benchtop and in-vitro validation; therefore, it needs to be adapted for addressing more patient-specific modelling (such as better addressing clinical data as comparator sources).
In addition to VV40, other ASME standards are available or under development:

- VVUQ20 (2004) on fluid simulations (broad scope)
- VVUQ70 (2019) on machine learning (broad scope)
- VVUQ80 (just started) on computational modelling of pharmaceutical products

General Introduction

Mr. Setzu's presentation. See Slide Deck 1 in Appendix 2.

Question from Mr. Bischoff
Model risk was assessed within the context of use intended by the people involved in designing and implementing the model credibility assessment through VVUQ activities. The model is then intended to be uploaded in the InSilicoTrials cloud platform to enable users to run custom in silico trials using such model.
1) How does the risk assessment apply to the final user, knowing that a priori a third-party user may target a different context of use?
2) Likewise, how does SCT ensure that the model applicability assessment is still valid for the platform final users? They may evaluate the model beyond the bounds for which the model has been assessed as credible.

Offline answer from WP6
During the meeting the questions were noted, but not discussed in depth due to time constraints.
Both questions, somehow related, merit some deeper consideration. First, we need to distinguish between two situations:

1) A final user intends to upload its own model in the InSilicoTrials cloud platform
2) A final user intends to use one of the models developed within SCT project

We anticipate that Case 1 will cover most situations. This case is beyond the scope of Mr. Bischoff's question, which instead addresses the credibility assessment of the numerical models developed by the 3 Use Cases.
That said, InSilicoTrials platform may provide informative material to third party users including guidance toward simulation good practices and anticipating expectations from regulators in case their model is intended to inform decisions impacting patients' health.

For what concerns Case 2, we agree that we cannot anticipate to what purpose the model outcome may be used, and, depending on the specific final context of use, the final model risk may be greater than the risk retained to design and implement the VVUQ process.
To mitigate such risk, the InSilicoTrials platform may include the following strategies:
1) Provide clear guidance to the end users on the Context of Use for which the model was initially developed, the associated risk profile, and the applicability of the model to such Context of Use.
2) Include a disclaimer note for the user allowing to proceed using the model in case:

- their Context of Use is not covered by the developer's defined Context of Use,
- their risk profile may be higher than the original model risk,
- or their applicability ranges are not covered by the initial applicability

3) Include in the user inputs/output some control on the data ranges (already in place) which allow to identify if the magnitude of the data inputs and model outputs are in the range of applicability of the model intended by the model developers.
4) The End Users may want to run independent VVUQ on their specific Context of Use using any of the Use Case models. In case SCT intends to allow this scenario (Feasibility To Be Discussed), the InSilicoTrials platform should provide, upon request, full reports of the VVUQ process. The End Users may rely on such information to design and implement their own VVUQ plan. Some results may be directly applicable to their VVUQ plan, upon their justification.

In general, SCT cannot commit on the success of a potential regulatory submission of End Users who would like to build a submission including any of the Use Case models and their credibility assessment. Such models were implemented as demonstrators for the effectiveness of the InSilicoTrials cloud platform and were not meant for supporting regulatory submissions of any industrial client.

UC1 – V&V approach and results

Mr. Coudière's presentation. See Slide Deck 2 in Appendix 2.

Question from Mr. Bischoff

- Did you use a specific SQA standard when developing the verification strategy? (such as ISO/IEC 25010)

Mr. Coudière: No specific SQA standard was implemented. The software used to run the model was fully developed in-house. The workload implications are significant, and not commensurate with our

structure, resources, and the timeframe of the project. We relied on wide-known and accepted tools (such as SonarQube) to implement our SQA activities.

Question from Mr. Bischoff

- Before a release, is there an audit process to verify?

Mr. Coudière: Not specifically. It remains a proof of concept, not a software commercially released.

Question from Mr. Bischoff

- For verification are there any parameters allowing users to verify independently?

Mr. Coudière: The software is completely available (the solver is open source). Any user may run the code verification tests provided with the software, or independent code verification if needed. If the users cannot do some tasks, they would fully inherit the verification work already performed.

Question from Mr. Coudière

- On VV40 perspective on using Surrogate Models for expanding sensitivity analysis.

Mr. Coudière pointed out that, in order to run extensive sensitivity analysis, a surrogate model (less computational demanding, and faster results) was used instead of the full model.
We questioned VV40 if they have a specific position concerning this point: the possibility of running sensitivity analysis and uncertainty analysis on surrogate models rather than the original models.
The point was acknowledged by VV40, as it is common to many domains dealing with VVUQ, without specific guidance.

Question from Mr. Bischoff

- How do the heart anatomies used to build and validate the model relate to the final model users in the IST platform?

Mr. Coudière: The model was built using a generic healthy heart as reference, because the underlying idea consists in placing the pacing lead in healthy tissues in order to stimulate the patient. These tissues will deteriorate over time. The model is then validated against eight different heart anatomies, of which some are pathologic conditions. However, we agree that the IST platform users may use the model in contexts of use which may still differ from the one used for validation. This would create additional uncertainty.

UC2 – V&V approach and results

Mr. Camara's presentation. See Slide Deck 3 in Appendix 2.

Question from Mr. Bischoff

- Is your question of interest focused on safety or efficacy?

Mr. Camara: It is actually both.

Question from Mr. Bischoff

- How the model performs against retrospective clinical data?

Mr. Camara: Several cases were already tested. We noticed in one case contradictory results with data over three months, but the model performs well over longer periods. There may be an intrinsic

issue of quality of available clinical data on the short term, which highlights the complexity of the clinical process we are addressing.

Question from Mr. Bischoff

- Companies A and B may like or dislike the model prediction on a specific device design based on whether its result concurs with their opinion (confirmation bias issue). Did you witness instead cases where their opinion has been flipped by simulation results?

Mr. Camara: Yes, as a matter of facts it already happened that following simulation advice, companies changed their opinion on specific designs. To some extent, the model proved to help companies to identify what they should do to improve their devices.

General advice from Mr. Bischoff

You should avoid a dogmatic use of the risk-coverage matrix (such that if you have low/medium/or high-risk application that every factor must meet individually that risk level).
The way you structured the coverage level (l/m/h) for each of the credibility factors should be based on how you can reasonably categorize different activities (e.g. around the validation). The risk may require doing extra V&V activities which do not add value to assess the credibility of the model.
Recent FDA credibility guidance and future VVUQ40 revision clarify this same message: do not engage additional V&V activity solely based on risk consideration, if technically not justified. In that case, just document the reason why a lower coverage is sufficient for that specific credibility factor.

UC3 – V&V approach and results

Mrs. Trenor's presentation. See Slide Deck 4 in Appendix 2.

Question from Mr. Bischoff

- Did you have benchmark solutions (numerical) for ion current functions?

Mrs. Trenor: There is no simple benchmark solution for our precise applications. However, for cellular electrophysiology, good models exist (https://www.cellml.org/). Whenever a new cellular model is published, it goes there. However, care should be taken when using such database, rather than taking such model as granted truth.

Question from Mr. Bischoff

- Considered that some drugs may pass validation, and some drugs won't, are you going to publish on the IST platform only drugs which will pass validation?

Mrs. Trenor: Yes.

Question from Mr. Bischoff

- What is, in your experience, the biggest difficulty in applying VVUQ40 standard to modelling of drugs. Is it the lack of knowledge, or rather V&V is too big, and people are going to limit to one part of it?

Mrs. Trenor: While dealing with drug modelling, we had to discard many clinical studies not fitting our simulations. We were limited to choosing the most reliable and comparable to our case.

General Questions / Discussion

General consideration from Mr. Bischoff

A lot of simulations we do in medical-devices industry is meant to support a standardised test method. A standardised test method is most often a single method, with a very precise scope and parameter space. The numerical model is intended to test the device against such test methods in a virtual space. Therefore, there is no point of validating the model outside the scope of the test method: the validation test conditions are set by the standard. Because we rely on a standardised method, we can accept lower coverage than what required by the model risk, as long as it is justified by the standardised test method.

Question from Mr. Camara to VV40

- What can be the role of academics to make VVUQ awareness progress?

Mr. Bischoff: The effort for physical testing is very significant and expensive so there is a lot of general interest in in silico. FDA is moving towards regulatory science developing toolkits to provide option to help generating evidence (cf. Regulatory Science Tools RST-catalogue, https://cdrh-rst.fda.gov/). FDA has a general interest in creating programs for academia helping to provide such tools that the industry can use. Additional feedback from FDA on model VVUQ can be collected through this program.
In addition, other ways to help raising awareness on VV40 and VVUQ across academia and industry may include:

- Becoming VV40 member
- Participating in standards review through public ballots
- Participating to FDA's MDIC (Medical Device Innovation Consortium) meetings (the next is planned end 2025)

Question from Mr. Camara to VV40

- What is VV40 experience on regulatory submission of VVUQ for specific COU?

Mr. Bischoff: As manufacturers, we usually had a good review experience. Some reviewers may be less familiar with VVUQ guidance, so additional effort may be needed. This will change from case to case, but in general a file well structured around VV40 guidelines is adequate to support submission.

Follow up Discussions

Offline discussion on the possibility of extending VV40 credibility assessment guidelines to computational model of drugs.

The following questions and answers summarize an offline discussion with Mrs. Martis on the possibility of extending VV40 credibility assessment guidelines to computational model of drugs.

Question from WP6

- Can we use VV40 to address credibility assessment of drugs models?

Mrs. Martis: Yes, in principle.
However, in my opinion, some additional work, including additional uncertainty quantification, is necessary to consider the specificity of clinical data used as model comparators (as opposed to in-vitro data).

Currently, VV40 only addresses the medical device perspective when studying the interaction with the test environment: both model and comparator uncertainty sources derive from the knowledge of the intrinsic behaviour of the device and of its interactions with the test environment.
On the other hand, in the case of a drug, the patient knowns and unknowns become predominant.
For instance, Patient Comorbidities should be adequately addressed.
The question of interest should clearly address both drug's desirable and adverse effects for both healthy patients and patients with comorbidities.
In addition, Model Risk should be reassessed considering comorbidities, leading to a much granular model risk to account for all these possibilities, or, in alternative, having different risks for different classes of patients.
Best Responders should be clearly identified. [Romano: sorry, this point was too technical, I didn't fully catch your suggestion.]

Question from WP6

- Is VV40 ever going to include guidance on drugs models VVUQ?

Mrs. Martis: There is nothing formally precluding from extending VV40 guidelines to drug models on a voluntary basis. FDA guidance on Credibility doesn't preclude it either.
However, being VV40 edited by a mechanical association focused on medical devices, drugs may never be included on the main document.
An alternative would be working on a dedicated ASME technical report (a case study) which, starting from the premises of VV40, expands on drugs models (without the ambition to becoming a binding document).

Question from WP6

- Availability of adequate clinical data is a big limitation to robust V&V. Much data may be discarded because not usable with the current model. What do you think?

Mrs. Martis: sometimes, available data is so scarce that we need to find a way to make a reasonable use of what we have, in the best interest of the patients.
A possibility is expanding your dataset using "model enhanced clinical data".
By that I mean the following:
Once you have a model on which you gained reasonable credibility through your applicable data subset, you could modify your model to fit the constraints and outputs of the additional available data, and then the comparison of the modified model with the new data can be used to strengthen your confidence on the model.

This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101016496