Project title: Trustworthy Digital Twins - Maintaining Arguments of Trust for Computational Models of the Heart

Project reference: DT4H_06_2023

Co 1st supervisor: Dr Oleg Aslanidi - School of Biomedical Engineering & Imaging Sciences

Co 1st supervisor: Dr Steffen Zschaler - Department of Informatics

Aim of the project

Cardiovascular disease is a leading cause of morbidity and mortality. It often disrupts electrophysiological mechanisms in the heart leading to cardiac arrhythmias, affecting the blood supply to vital organs, and greatly increasing risks of thrombus formation and stroke. However, the underlying mechanisms are incompletely understood, resulting in suboptimal treatments. Mechanistic and statistical models emerged as powerful tools for improved understanding. These models – digital twins of the heart – are themselves mathematically and computationally complex software. To trust them in clinical contexts, their relation to the real world and validity of predictions need to be carefully considered. Currently, such considerations are left implicit or, at best, expressed in natural language that is difficult to validate automatically.

This project, aims to 1) understand typical argument structures in support of the validity of computational models of the heart, 2) develop a formalisation of and tool support for such arguments that supports semi-automated validation of arguments, and 3) explore usability implications of using such arguments in the context of research and product development.

Project description

Cardiovascular diseases (CVD) are a leading cause of morbidity and mortality, accounting for over 30% of all deaths in recent decades. Mechanisms of such diseases are often linked to the occurrence of cardiac arrhythmias – common pathologies caused by disruptions in electrical activity in the heart. Several mechanisms contribute to arrhythmogenesis, such as disruptions in the propagation of cardiac action potentials (APs) by re-entry. However, the precise spatio-temporal mechanisms of most arrhythmias are still poorly understood, and the success rate of their treatments remains suboptimal. Computational models of cardiac electrophysiology (cardiac digital twins) emerged as a quantitative framework that can integrate clinical and experimental data, dissect complex mechanisms of arrhythmias and improve therapy [1].

AP generation is typically modelled using ordinary (ODEs) and partial (PDEs) differential equations that are solved numerically. To trust such models in clinical contexts, many questions need to be answered, for example: Do the ODEs and PDEs capture all relevant aspects of the real heart? How precisely do they capture these aspects, in particular relating to predictions to be made by the model? What range of parameters are correctly represented by the model and when does the model stop to be valid? How does the implementation of the model in a particular computer program affect its validity? Even more broadly, trustworthiness is also an issue when considering the European Union's General Data

Protection Regulation (GDPR), as any algorithmic decision used in patient care requires an explanation for transparency [3]. Answering these questions requires providing explicit arguments drawing on a wide range of literature, explicit design decisions, calibration and validation experiments, test suites, etc. Moreover, the arguments need to be constantly maintained and evolved as the software model, or the scientific understanding and literature evolve.

Currently, such models of validity are not captured explicitly or, at best, as natural-language arguments which are difficult to validate and maintain over time.

In Software Engineering, engineers of safety-critical systems have long been using structured arguments [2] to capture their assessment of the safety of a given system. More recently, these ideas have been translated informally into the area of computational modelling [6][7]. This project will explore these ideas in the concrete context of cardiovascular modelling and will develop a tool that can support the formal capture of structured arguments, their maintenance and evolution, and their semi-automated validation. To this end, the project will draw on ideas from safety assurance, domain-specific modelling, model-driven engineering, and agile software engineering. The work will follow a design-science methodology, using our cardiovascular models for predicting patient-specific ablation therapy [5] and risk of thrombogenesis [4] as case studies.

References:

1. Boyle P. M., Zghaib T., Zahid S., Ali R. L., Deng D., Franceschi W. H., et al. Computationally guided personalized targeted ablation of persistent atrial fibrillation. *Nat. Biomed. Eng.* 3, 870–879. 2019. doi:10.1038/s41551-019-0437-9
2. Kelly, Tim, and Rob Weaver. The goal structuring notation–a safety argument notation. *Proceedings of the dependable systems and networks 2004 workshop on assurance cases,* 2004.
3. Mourby, M., Ó Cathaoir, K., and Collin, C. B. Transparency of machine-learning in healthcare: The GDPR & European health law. *Comput. Law Secur. Rev.* 43, 105611. 2021. doi:10.1016/j.clsr.2021.105611
4. Qureshi A., Lip G.Y.H., Nordsletten D.A., Williams S.E., Aslanidi O., de Vecchi A. Imaging and biophysical modelling of thrombogenic mechanisms in atrial fibrillation and stroke. *Front. Cardiovasc. Med.* 9:1074562, 2023.
5. Roy, A., Varela, M., Chubb, H., Robert, M. L., Hancox, J. C., Schaeffter, T., Aslanidi O. Identifying locations of re-entrant drivers from patient-specific distribution of fibrosis in the left atrium. *PLoS Comput. Biol.* 16: e1008086, 2020.
6. Stepney, S. Polack, Fional A.C. *Engineering Simulations as Scientific Instruments: A Pattern Language.* Springer, 2018.
7. Zschaler, S., Polack, Fional A.C. Trustworthy agent-based simulation: the case for domain-specific modelling languages. *Softw Syst Model* 22, 455-470, 2023.