

A Computational Model-Based Framework to Plan Clinical Experiments – an Application to Vascular Adaptation Biology

Stefano Casarin^{1,2,3}, Scott A. Berceci^{4,5}, and Marc Garbey^{1,2,3}

¹ LASIE UMR 7356 CNRS, University of La Rochelle, La Rochelle, France

² Center for Computational Surgery, Houston Methodist Research Institute, Houston, TX, USA

³ Department of Surgery, Houston Methodist Hospital, Houston, TX, USA

⁴ Department of Surgery, University of Florida, Gainesville, FL, USA

⁵ Malcom Randall VAMC, Gainesville, FL, USA

scasarin@houstonmethodist.org

Abstract. Several computational models have been developed in order to improve the outcome of Vein Graft Bypasses in response to arterial occlusions and they all share a common property: their accuracy relies on a winning choice of the coefficients' value related to biological functions that drive them. Our goal is to optimize the retrieval of these unknown coefficients on the base of experimental data and accordingly, as biological experiments are noisy in terms of statistical analysis and the models are typically stochastic and complex, this work wants first to elucidate which experimental measurements might be sufficient to retrieve the targeted coefficients and second how many specimens would constitute a good dataset to guarantee a sufficient level of accuracy. Since experiments are often costly and time consuming, the planning stage is critical to the success of the operation and, on the base of this consideration, the present work shows how, thanks to an *ad hoc* use of a computational model of vascular adaptation, it is possible to estimate in advance the entity and the quantity of resources needed in order to efficiently reproduce the experimental reality.

Keywords: Agent Based Model - Experiment Planning - Virtual Dataset.

1 Introduction

An unfavorable postsurgical vascular adaptation is the main cause of the long-term failure of Vein Graft Bypasses (VGBs), with an incidence rate of 12% after just a month from the original intervention [1].

To improve the current outcome, our group of investigators, in addition to others [2], studied vascular adaptation both on a clinical [3] and a computational base [4-6], from deterministic differential equation models to especially Agent-Based Models (ABMs) that use a so-called bottom up approach depending on the level of spatio-temporal and biological information needed.

Our hypothesis is that the ABM [5] mimics the fundamental biological functions of the vein graft's adaptation (see Fig. 1) with high credibility, being indeed able to replicate the basic observations driving Intimal Hyperplasia (IH) and Wall Remodeling (WR) that are globally recognized as the leading events for the restenosis [1].

We used standard methods both to restrict the analysis to a minimum set of coefficients and to reach a high level of confidence with our model, respectively by performing a non-linear stability analysis [7], and by cross-validating the ABM on phenomenological models [5,8].

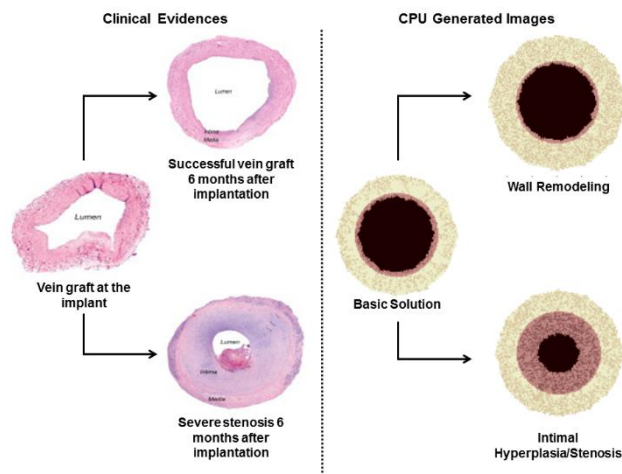


Fig. 1. Agent-Based Model. An ABM (on the right) replicates the clinical evidences (left) of vein bypasses long-term follow-up [5].

The overall goal is to retrieve the fundamental parameters of vascular adaptation on a base of experimental data, a task that implies to perform a large number of experiments. This need is actual both in vascular field as several times demonstrated by different group of investigators [9-11], but also more in general in the world of computational models developed to replicate pathophysiological phenomena [12].

Before starting the actual experiments, two fundamental questions concern the kind of measurements and the number of subjects needed in order to achieve a realistic outcome.

Considering that the ABM model is stochastic and assuming that it is as noisy as the biological system that replicates, our basic method consists into using the ABM to mimic the experiments by generating a virtual experimental data set.

In this way, we can ask ourselves how large the dataset should be and what kind of measurements will be good enough to recover the unknown biological parameters. This latter is typically a non-trivial exercise, which complexity strongly depends on the kind of the measurements needed.

Finally, the performance of the proposed technique relates on an accurate setup of the ABM's driving coefficients and especially on an effective choice of their range of

perturbation, as fluctuations of the ABM can lead to large output variations that would be hard to catch with the proposed inverse problem.

Both the required features are guaranteed by an accurate study of the experimental data preceding the model's implementation [3].

2 Methods

According to the conceptual scheme of Fig. 2, the method is based on a twofold usage of our ABM [5] that acts both as virtual experimental dataset generator (labeled as ABM1) and as a true computational model (ABM2).

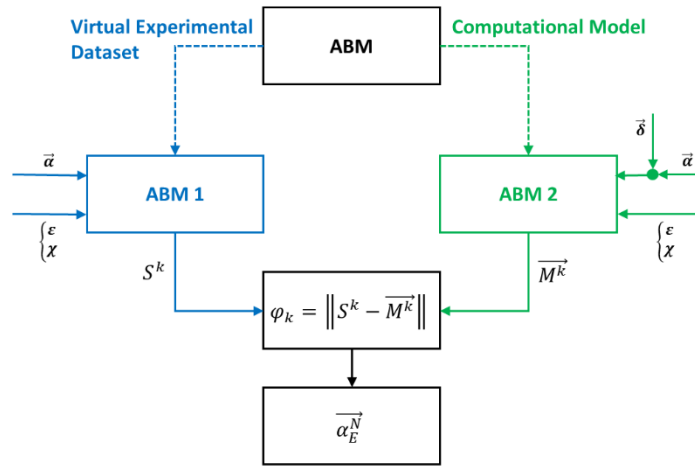


Fig. 2. Conceptual scheme. A twofold usage of an ABM generates both the reference measurements (S^k) intended as experimental data to be replicated, and the perturbed output (\vec{M}^k) to be compared with the reference in order to solve the inverse problem and to plan the experiment.

First, the left part of Fig. 2 refers to the generation of the virtual experimental data. The ABM works as a black box driven in input by hemodynamic perturbations, ε and χ respectively referring to shear stress and wall tension, and by constant coefficients corresponding to the level of activity of the cellular events of interest, i.e. vector $\vec{\alpha}$. ABM1's output is the virtual measurement S^k which relation follows:

$$S^k = \sum_{n=1}^N S_n^k, \quad (1)$$

where k is the label identifying the specific measurement (e.g. $k=1$ for intimal area, $k=2$ for medial area, *et cetera*) and N is the number of ABM1 independent simulations that would be the specimen's size in a true experimental setup.

Second, in the right part of Fig. 2 the ABM is used as a computational model and for the purpose of this work, a browsing of the coefficients' space was performed in order

to solve the inverse problem, i.e. starting from the same setup of ABM1, a vector $\vec{\delta} = \{-0.5, -0.375, \dots, 0, \dots, 0.5\}$ was applied to the driving coefficients defined with $\vec{\alpha}$. At each element of $\vec{\delta}$ corresponds a new input/output of ABM2. ABM1's output is a constant measurement S^k , while ABM2's output is a vector of measurements \vec{M}^k , where the size of \vec{M}^k is equal to the size of $\vec{\delta}$, and it corresponds to $P=9$.

To solve the inverse problem consists in comparing S^k with \vec{M}^k , i.e. in minimizing the objective function (φ_k) defined as follows:

$$\varphi_k = \sqrt{\frac{1}{P} \sum_{p=1}^P (S^k - M_p^k)^2} * 100. \quad (2)$$

If the entity of the measurements (k) is suitable enough, and if the number of samples (N) guarantees a robust analysis, the result of the minimization of φ_k , indicated with $\vec{\alpha}_E^N$, will correspond to the same values defined with vector $\vec{\alpha}$, i.e. $\vec{\alpha} = \vec{\alpha}_E^N$.

A counter-proof will be to generate another independent virtual dataset by running the ABM on the base of N simulations and by setting it with $\vec{\alpha} = \vec{\alpha}_E^N$. In this way the percentile error Ξ_k between S^k , the old reference, and S_N^k , the new reference, is

$$\Xi_k = \frac{|S^k - S_N^k|}{S^k} * 100. \quad (3)$$

When the Ξ_k vs. N plot goes below an acceptable error threshold, N is suitable to build an experimental setup, and at the same way, k is an effective measurement.

In this work, IH and WR were chosen as case studies and they were addressed both separately and coupled.

2.1 Intimal Hyperplasia

IH is mostly driven by Smooth Muscular Cells (SMCs) division within intima fostered by a decrease in shear stress from the baseline condition, representing the vein at implant. As ABM is stochastic, the related formula corresponds to a probability distribution, true for WR too, which writes:

$$P_{div}^{int} = \alpha_1 \left(1 + \alpha_3 \frac{\tau(t) - \tau_0(1+\varepsilon)}{\bar{\tau}} \right), \quad (4)$$

where $\alpha_1 = \frac{1}{T_{matrix}} = 0.5$ defines the baseline of cellular division, $\tau(t)$ is the shear stress recorded at t -th time point, τ_0 is the baseline value of the shear stress, $\bar{\tau} = 0.25$ is a normalization constant to maintain the probability discrete value between 0 and 1, $\varepsilon = 0.5$ represents a 50% drop in shear stress to promote SMC division, and finally $\alpha_3 = 0.2$ is the constant coefficient that drives SMC intimal division.

The measurement chosen to test the experimental setup is the intimal area, labeled as $k = 1$, and the accuracy of the inverse problem solution will be tested, in terms of percentile error, with an increasing number of samples starting with $N=10$ up to $N=100$.

To solve the inverse problem for IH means to be able to retrieve for α_3 the value of 0.2 by minimizing the objective function φ_1 corresponding to the RMS between the

output of ABM1, S^1 and ABM2, \overline{M}^1 . As last remark, the addressed problem will be one-dimensional as only a single coefficient will be retrieve out of the input vector $\vec{\alpha}$.

2.2 Wall Remodeling

The framework for WR is similar to the one of IH, with the difference that WR is fostered by SMC division within media favored by an increase of wall tension from the baseline condition. The related mathematical formula writes

$$P_{div}^{med} = \alpha_1 \left(1 + \alpha_4 \frac{\sigma(t) - \sigma_0(1+\chi)}{\bar{\sigma}} \right), \quad (5)$$

where $\sigma(t)$ is the wall tension recorded at t-th time point, σ_0 is the baseline value of wall tension, $\bar{\sigma} = 40$ is the normalization constant, $\chi = -0.005$ represents a 0.5% increase in wall tension to promote SMC division, and finally $\alpha_4 = 0.3$ is the constant coefficient that drives SMC medial division.

The measurement chosen is the medial area, labeled k=2, while the number of samples stays between N=10 and N=100.

To solve the inverse problem for WR means to be able to retrieve for α_4 the value of 0.3 by minimizing the objective function φ_2 obtained with the RMS between the output of ABM1, S^2 and ABM2, \overline{M}^2 . As for IH and for the same reasons, the problem is still one-dimensional.

2.3 Coupled Intimal Hyperplasia and Wall Remodeling

A more realistic implementation considers the simultaneous study of IH and WR which maintains the same setup for both the cellular events, described with Eq. (4) and Eq. (5).

The measurements chosen to solve the inverse problem are still intimal area and medial area respectively, however, if with a single cellular event the inverse problem was one-dimensional (φ_1 and φ_2), with two cellular events acting simultaneously the problem becomes two-dimensional ($\varphi_{1,2}$ and $\varphi_{2,1}$). By maintaining the same vector $\vec{\delta}$ and by applying it both to α_3 and α_4 , the output of ABM2 will not be a simple vector of measurements, but a matrix of measurements where the i-th row includes the values for a fixed α_4 at the variation of α_3 , while the j-th column includes the values for a fixed α_3 at the variation of α_4 . Accordingly, each of the two coefficients will be retrieved by minimizing a 2D-surface instead of a 1D-curve.

In general, by adding more cellular events to the analysis, the complexity and the dimensionality of the inverse problem increase proportionally with the closeness to the physiological reality.

3 Results and Discussions

The virtual experiments performed are described with the density map reported in Fig. 3. In each panel the number of SMCs across the graft wall is represented against the

normalized wall thickness, making so both intima (in red) and media (in blue) share the same unitary dimension. The density map is reported as a histogram with 5 steps of discretization for each compartment.

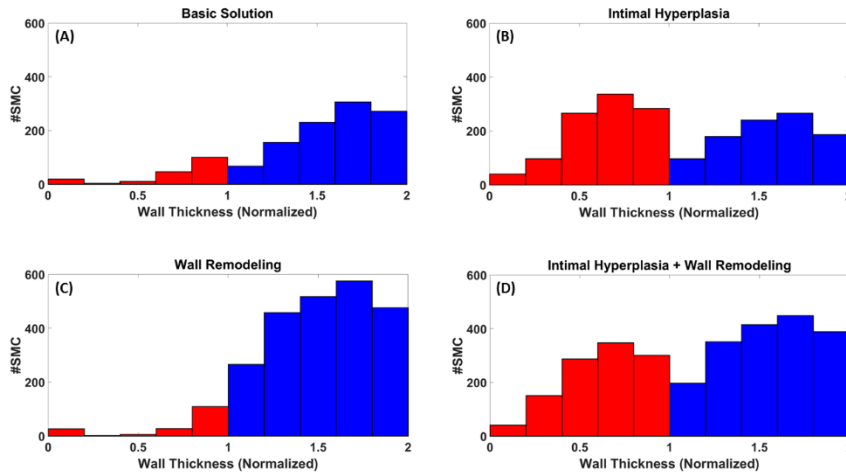


Fig. 3. Virtual Experiments. Starting from a basic solution (A), representing a healthy vein at the implant, ABMI is setup in order to simulate the hyperplasia of tunica intima (B) and tunica media (C), both acting singularly and coupled (D).

The analysis of the density map is coherent with the well-known physiological reality: Fig. 3(A) represents the ABM baseline condition that corresponds to a vein at the implant. By fostering intimal SMCs division, a noticeable increase in intimal SMCs number is recorded in Fig. 3(B), well representing IH, while no substantial variation is recorded in the number of medial SMCs. Similar considerations for Fig. 3(C), where SMCs division within tunica media generates an increase in SMCs medial number, while no variation from the baseline is recorded for intimal SMCs, typical of the WR phenomenon. Results appreciated in Fig. 3(B) and 3(C) are coherent with the single cellular event nature of the experiments performed so far, while a different subject is valid for Fig. 3(D), where SMCs division is active both in intima and in media, doing so both the number of SMCs in intima and in media are prone to increase.

Despite the setups to simulate the WR in Fig. 3(D) and Fig. 3(B) are the same, a lower increase in medial cell numbers is recorded in Fig. 3(D) than in 3(B). This is due to the fact that IH and WR act simultaneously in the last experiment performed, and accordingly, the IH phenomenon accelerates the wall relaxation toward the baseline value and this reduces the SMCs division in media. This is the perfect example of how cellular events can generate noise at experimental level when they act simultaneously, making the identification process way harder, in the specific case by affecting the identification of α_4 .

The inverse problem to singularly identify α_3 was successfully solved and the related analysis is reported in Fig. 4.

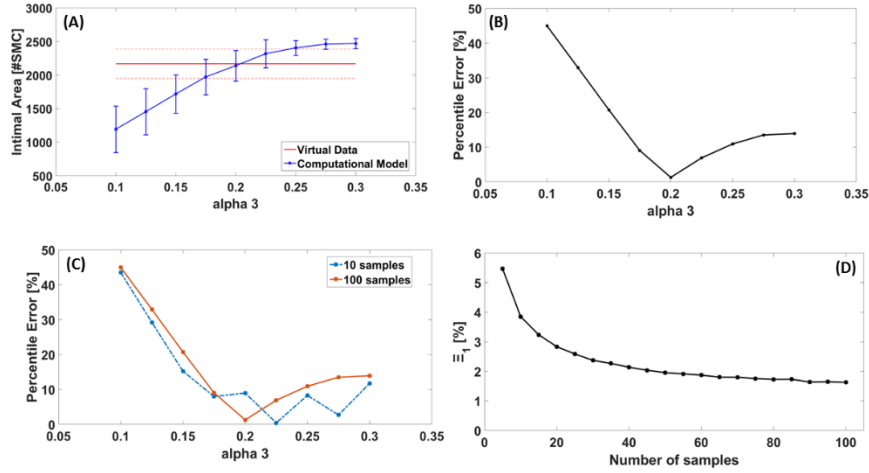


Fig. 4. Intimal Hyperplasia. The interval of confidence (A) and the actual value (B) of α_3 are retrieved on the base of $N = 100$ samples by measuring the intimal area. The accuracy of the retrieval is evaluated by comparing the error using 10 and 100 samples (C). Finally, the accuracy of the fitted model is studied against the number of samples used during the retrieval process (D).

By monitoring the intimal area as the mean value over 100 samples, a precise identification of $\alpha_3 = 0.2$ was obtained. This is clear from Fig. 4(A) where the intimal area as output of ABM1 (S^1) is plotted in solid red, along with its standard deviation in dashed red, against the intimal area as output of ABM2 (\bar{M}^1) in solid blue, again along with its standard deviation. Fact that the perturbation level-dependent blue curve intersects the constant red curve right in correspondence of $\alpha_3 = 0.2$ is a clear indicator of the goodness of our analysis, also corroborated by Fig. 4(B), where the PRMS function between S^1 and \bar{M}^1 (φ_1) is precisely minimized in correspondence of $\alpha_3 = 0.2$. Furthermore, from the comparison between the minimization of φ_1 by using $N = 10$ (blue dashed line) and $N = 100$ (solid red), reported in Fig. 4(C), it is easy to deduce how the number of samples used is pivotal in order to precisely identify the right value of α_3 . One can indeed simply appreciate how, by using only 10 samples, φ_1 is not uniquely minimized and not even in correspondence of the right value. Finally, the non-suitability of a low number of samples is further sustained with Fig. 4(D) that shows the trend of Ξ_1 against N . If an error lower than 2% is considered as acceptable, then at least 50 samples are needed to obtain a realistic analysis, a number that increases if a lower error is desired.

Finally, according with the successful minimization result, the intimal area can be considered as a solid biological measurement to prepare the experimental setup, and as its measuring does not pose any significant issue, other potential measurements have not been investigated.

Very similar considerations can be retrieved from the identification of α_4 , which is reported in Fig. 5.

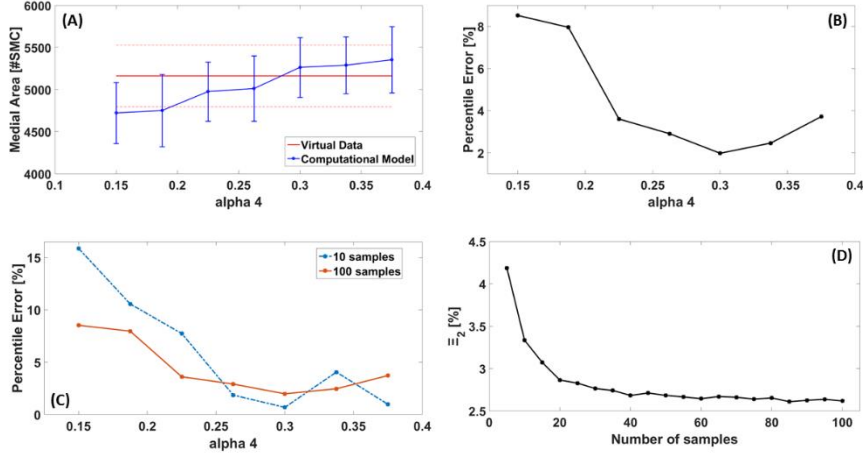


Fig. 5. Wall Remodeling. The interval of confidence (A) and the actual value (B) of α_4 are retrieved on the base of $N=100$ samples by measuring the medial area. The accuracy of the retrieval is evaluated by comparing the error using 10 and 100 samples (C). Finally, the accuracy of the temporal simulation is studied against the number of samples used during the retrieval process (D).

The results are here presented with the same fashion of Fig. 4. The accuracy of this identification may be not as high as the one appreciated for α_3 , but it is still reasonable enough to be considered as suitable for a modeling purpose. Indeed, if on one side from Fig. 5(A) emerges that the computational model curve (blue) does not perfectly intersect the virtual dataset curve (red) in $\alpha_4 = 0.3$, on the other side it is clear from Fig. 5(B) how φ_2 is still nicely minimized right around 0.3, even though not with the same sharpness of φ_1 in correspondence of $\alpha_3 = 0.2$ (see Fig. 4(B)).

What already appreciated for α_3 about the increased confidence along with an increased number of samples is confirmed for α_4 by observing Fig. 5(C), where φ_2 does not reach a global minimum when minimized on the base of 10 samples. The analysis of Fig. 5(D) not surprisingly confirms the need of a conspicuous sample numbers to ensure the reproduction of the experimental data free from a high error. It is interesting to notice how not even a high number of samples like $N = 100$ ensures to reach a $\Xi_2 < 2\%$ like instead happened for Ξ_1 . However, a sort of plateau is reached with $N = 60$ that can be considered as a number beyond that the goodness of the analysis does not significantly increase.

Accordingly, if someone assumes a $\Xi_2 \approx 2.6$ as an acceptable error, the medial area can be considered as an appropriate measure to setup the experiment and for the same reason reported for intimal area, other measurements have not been considered in this analysis.

Finally, the inverse problem solution is reported in Fig. 6 for the coupled IH and WR, where a qualitative more than a quantitative analysis is appreciable. It is important to remark how to solve the inverse problem for the coupled IH and WR means to ideally

minimize a surface around a straight line instead of a curve around a single point. Also, the number of samples was chosen to be here $N = 60$, which is the highest between $N = 50$ and $N = 60$ to singularly identify α_3 and α_4 respectively.

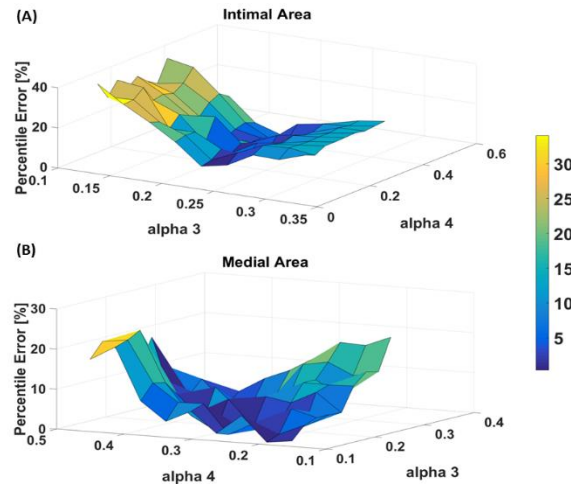


Fig. 6. Intimal Hyperplasia and Wall Remodeling. The value of α_3 (A) by intimal area measurement and α_4 (B) by medial area measurement are retrieved by letting intimal hyperplasia and wall remodeling acting simultaneously.

From Fig. 6(A), the 2D surface corresponding to $\varphi_{1,2}$ is minimized in correspondence of the straight line where $\alpha_3 = 0.2$. This is a remarkable result, indeed even though the level of complexity of the system has increased of one order of magnitude, by maintaining the same experimental setup, a satisfying result is still achievable. However, the same cannot be stated about the minimization of $\varphi_{2,1}$ (see Fig. 2(B)), for which a clear minimization is not visible around the straight line identified by $\alpha_4 = 0.3$. This implies that, in order to setup an experiment able to replicate the coupled IH and WR mechanisms, the number of samples must be incremented to reach a better convergence, or that the medial area is not a suitable measurement to correctly tune α_4 , or even both. This was actually to be predicted, indeed, as anticipated while discussing Fig. 3, it was clear how the IH had some kind of noisy influence on the WR, while the *vice versa* was not valid. The resulting limited variation in the number of medial SMCs, at least for the coupled IH and WR respect to the single WR, makes that a different measurement, still representative of WR but not influenced by IH, might be more suitable to setup α_4 .

To summarize, the analysis presented showed how, by measuring intimal and medial thickness on the base of 50 and 60 samples, a realistic setup for a model that replicates IH and WR separately is reached and in both cases this task does not pose particular criticalities. However, the same measurements are not suitable enough to inform the model in order to replicate IH and WR when they act simultaneously. Therefore, to design an effective experimental setup cannot exempt itself from an analysis of spatial SMCs density maps across the graft's wall, a procedure undoubtedly more resources-

consuming than simply measuring the thickness of the compartments. This necessity becomes even more actual when thinking about cellular motility, not considered for simplicity in this analysis, that also plays a key role in vascular adaptation [2] and that adds to the system an additional level of noise that requires a more complex measurements, and probably also a larger number of samples to be overpassed.

4 Conclusions

This work represents an agile framework to design the experimental setup needed to obtain computational models able to reproduce as accurately as possible the biological reality. With an *ad hoc* use of our ABM [5] the minimum number of samples needed for the analysis can be easily predicted and, not less important, the suitability of the biological measurements verified *a priori*.

This can represent a great turning point in the field of computational models, because it will give the researchers a feeling of the resources needed before conducting the actual experiment, limiting in this way any kind of wastefulness in terms of equipment and of resources more in general, especially animal resources.

The next steps will be i) to reach a satisfying identification for the coupled cellular events, where the real challenge will be to identify the right biological measurement to rightly tune α_4 , and ii) to improve the framework by getting it even closer to the physiological reality. This will be done by adding more cellular events, like cellular apoptosis, ExtraCellular Matrix (ECM) synthesis, and especially cellular motility.

References

1. Goldman S, et al. Long-term patency of saphenous vein and left internal mammary arteries grafts after coronary artery bypass surgery: results from a Department of Veterans Affairs Cooperative Study. *J. Am. Coll. Cardiol.* 2004; 44:2149-56
2. Manini S, Passera K, Huberts W, Botti L, Remuzzi A. Computational model for simulation of vascular adaptation following vascular access surgery in haemodialysis patients. *Comput. Methods Biomech. Biomed. Engin.* 2014; 17:1358-67
3. Klein B, et al. Hemodynamic Influence on Smooth Muscle Cell Kinetics and Phenotype During Early Vein Graft Adaptation. *Ann. Biomed. Eng.* 2017; 45:644-655
4. Garbey M, et al. A dynamical system that describes vein graft adaptation and failure. *J. Theor. Biol.* 2013; 336:209-20
5. Garbey M, et al. Vascular Adaptation: Pattern Formation and Cross Validation between an Agent Based Model and a Dynamical System. *J. Theor. Biol.* 2017; 429:149-163
6. Casarin S, et al. Linking Gene Dynamics to Intimal Hyperplasia – A Predictive Model of Vein Graft Adaptation. *Procedia Computer Science* 2017; 108:1842-51
7. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* 2008; 254:178-196.
8. Owens CD, Gasper WJ, Rahman AS, Conte MS. Vein graft failure. *J. Vasc. Surg* 2015; 61:2013-216
9. Casarin S, Berceci SA, Garbey M. Linking gene dynamics to vascular hyperplasia – Toward a predictive model of vein graft adaptation. *PLoS ONE* 12(11): e0187606.

10. Manini S, Passera K, Huberts W, Botti L, Antiga L, Remuzzi A. Computational model for simulation of vascular adaptation following vascular access surgery in haemodialysis patients. *Computer Methods in Biomechanics and Biomedical Engineering* 2014; 17: 1358-1367
11. Pries AR, Secomb TW, Gaehtgens P. Structural Autoregulation of Terminal Vascular Beds. *Hypertension* 1999; 33:153-161
12. Poleszczuk J, Macklin P, Enderling H. Agent-Based Modeling of Cancer Stem Cell Driven Solid Tumor Growth. *Methods Mol. Biol.* 2016; 1516: 335-346