Experimental Modeling and Identification of Cardiac Biomarkers Release in Acute Myocardial Infarction

Anna Procopio[1,⋆], Salvatore De Rosa[2,⋆], Míriam R. García[3], Alessio Merola[1], Jolanda Sabatino[2], Alessia De Luca[2], Ciro Indolfi[2], Francesco Amato[1,§], and Carlo Cosentino[1,2,§,*]

[1]Biomechatronics Lab, Dept. of Experimental and Clinical Medicine, Università degli Studi Magna Græcia, 88100 Catanzaro, Italy

[2]Interventional Cardiology Unit, Dept. of Medical and Surgical Sciences, Dept. of Experimental and Clinical Medicine, Università degli Studi Magna Græcia, 88100 Catanzaro, Italy

[3]Bioprocess Engineering Group. IIM-CSIC, Vigo, Spain

[⋆]Equal contribution

[§]Joint senior authorship

[*]carlo.cosentino@unicz.it

Abstract

Cardiovascular diseases represent, to date, the major cause of mortality worldwide. Diagnosis of the most frequent of such disease, Acute Myocardial Infarction (AMI), requires the evaluation of time-series measurement of specific cardiac biomarkers concentration. The aim of this work is to provide the clinicians with a quantitative tool to analyze such time-series, with the final goal of enhancing the diagnostic and prognostic procedures. The proposed approach is based on a novel dynamical model, which synthetically describes the basic mechanisms underlying cardiac troponin (cTnT) release into the plasma after the onset of AMI. Leveraging tools of system identification and a dataset of AMI patients treated at our University Hospital, the model has been assessed as an effective tool to quantify the characteristic release curves observed under different conditions. Furthermore, it has been shown how the devised approach is also suitable in those cases where only partial measurements are available to the clinician, to recover important clinical information. Finally, an Optimal Experimental Design (OED) analysis has been performed in order to gain insights on how to optimize the experimental data collection phase, with potentially relevant implications on both the quality and cost of the diagnosis procedure.

Index terms— System identification; biological models; cardiac biomarkers; acute myocardial infarction; identifiability; optimal experimental design.

1 Introduction

Circulating biomarkers of myocardial damage are the base for diagnosis of AMI and for the stratification of patients' risk. Among the different types of biomarkers, tissue-specific cardiac troponins T(cTnT) and I(cTnI) are the most widely used for the identification of myocardial injury, and specific diagnostic guidelines for their use have been implemented within the current clinical practice [1], [2], [3]. Clinical guidelines are mostly based on whether or not the cTnT concentration level is above a certain threshold, which is determined by means of a statistical analysis of a healthy population sample: the diagnosis of AMI requires that the

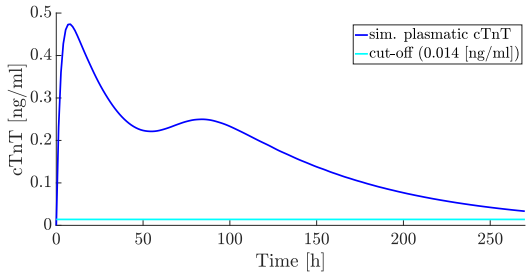


Figure 1: Typical biphasic pattern of the time-evolution of cTnT concentration in the plasma following an AMI.

cTnT concentration is above the 99th percentile of the upper reference limit (URL), in combination with other evidences of myocardial ischemia, as detected, either by clinical symptoms, electrocardiographic (ECG) or imaging evidences. The cTnT concentration level is measured upon clinical suspect and repeated in a time-series, following the most recent diagnostic protocols, with typical sampling intervals varying between 6 and 24 hours over the course of time; therefore, the experimental cTnT data for an AMI patient takes the form of a time-series response.

The release curve in the bloodstream after the onset of AMI typically exhibits a biphasic pattern: the concentration of cTnT, which is generally undetectable in healthy patients, undergoes an initial rapid increase, with a peak after 12-24 hours, followed by a plateau phase, lasting about 48 hours and possibly comprising a second peak, and eventually decreases exponentially to undetectable levels in one-two weeks. A typical time-course of plasma cTnT concentration caused by an AMI is showed in Fig.1.

The shape of the release curve, along with the inter-individual variability (determined also by different treatment strategies and co-morbidities), possible missing samples (especially in the first hours), experimental noise and, in some cases, the uncertainty on the exact time of occurrence of the ischemic event, renders the clinical interpretation of the experimental data quite difficult, especially. To further complicate the problem, one should also take into account that there are other forms of myocardial damage that may result in elevated cTnT concentration levels (e.g., cardiac trauma, myocarditis, severe pulmunary embolism or hypertension). Therefore, there is an urgent need of more specific and quantitative characterization methods for the diagnosis of AMI from cTnT experimental observations.

Considering the points above, it is clear how a simple threshold-based analysis would completely neglect the dynamic nature of the release process, along with all its valuable intrinsic information, which may be of great help in the characterization of the specific pathologic case under observation. The alternative strategy, to overcome this limitation and achieve a quantitative and personalized characterization of the patient pathology, is to base the analysis of the experimental data on a dynamical model of the release process. Ideally, such a model should recapitulate the basic mechanisms of the release kinetics, though allowing enough degrees of freedom to accommodate the different evolutions observed within the AMI patients population. This calls for the use of a parametrized model, whose parameters can be fitted to the specific release curve of each patient, thus enabling the extraction of a discrete number of clinically relevant, personalized and quantitatively gradable features. Thus, the first main contribution of the present work is the design of a novel mathematical model to predict the release of cardiac troponin T (cTnT) subsequent to AMI. Many analysis techniques are based on dynamical models in several fields of medicine and biology, ranging from the reverse-engineering of gene regulatory networks [4], [5], and the characterization of cell differentiation [6], to oncology [7], [8], [9] and cardiology [10]. However, to the best of our knowledge there is no previous work presenting a dynamical model of cardiac troponin release. Preliminary versions of the present work have been presented in [11] and [12]. The present work is refined and extended along several directions: the clinical dataset has been considerably expanded, comprising cTnT time-series of 85 patients with elevation of the ST segment of the ECG wave, which enhances the statistical value of the results. The system identification step has been refined, by implementing a robust regression method, to cope with outliers. Furthermore,

in this work the devised model is leveraged to compute Optimal Experimental Design (OED): realizing the importance of fine-tuning the estimated model to get reliable information even in the presence of few measurements, we have exploited the methods of OED to investigate the optimal allocation of experimental measurements over a reference time window.

The development of a reliable cTnT release model and identification procedure is a valuable step towards AMI personalized diagnosis and treatment. For instance, hereinafter we investigate two possible ways to leverage the proposed mode-based approach: i) to extrapolate information about the cTnT release peak, and ii) to optimize the blood sampling procedure.

The work is structured as follows: Section II provides a short explanation on the clinical and pathological concepts treated; in Section III we introduce the proposed mathematical model and the experimental dataset, as well as the techniques used for identifiability analysis, parameters identification, and optimal experimental design. The results of such analysis are reported in Section IV, whereas in Section V discussion and conclusions are given.

2 Background

One of the most common cardiovascular diseases, acute myocardial infarction (AMI), represents the first cause of death worldwide. Usually, an AMI begins with a coronary block, as a result of a thrombosis. This block causes the necrosis of cardiac tissue due to insufficient or missed perfusion of the myocardium. Acute myocardial infarction (AMI) encompasses two different conditions, depending on the ECG signal:

- STEMI: ST-Elevation Acute Myocardial Infarction (ST is a section of the ECG wave); the obstruction completely blocks the blood flow downstream of it and the ECG exhibits an elevated value in the ST interval;
- NSTEMI: Non ST-Elevation Acute Myocardial Infarction; in this case the obstruction is partial, the ECG signal does not show ST-segment elevation and may even be normal, in some cases. In these patients the analysis of the cTnT levels is even more crucial to identify an AMI [2].

STEMI and NSTEMI present some common features, including the presence of myocardial necrosis, although, however, each of these two clinical syndromes show specific characteristics, requiring different treatment strategies. Considering the lack of previous work on the mathematical modeling of these phenomena, and the much higher heterogeneity of the release curves presented by NSTEMI patients, in the present work we focused on the STEMI case, whereas the NSTEMI one will be the subject of future investigation. The reason why we focused on STEMI was the possibility to determine the time of symptoms' onset more precisely. This has important implications for the quality of the model to predict the age of the AMI.

Cardiac troponin was identified as the optimal biomarker for the following characteristics: i) high specificity for cardiac tissue; ii) high sensitivity to the cardiac damage; iii) wide observation window, which facilitates the collection of samples and the elaboration of the information needed to delineate an accurate clinical profile. In particular, the cTnT is only a part of the cardiac troponin complex, composed by: i) troponin C (cTnC), which binds the calcium, favoring later muscle contraction. However, cTnC, being very similar to its skeletal isoform, has low specificity for cardiac tissue; ii) troponin I (cTnI), which inhibits the interaction between actin and myosin; iii) troponin T (cTnT), which allows the complex to bind to tropomyosin. cTnT exhibits all of the features of an ideal cardiac disease biomarker; therefore, it is considered the gold standard for the diagnosis of AMI:

- Sarcomere, where cTnT is linked to tropomyosin;
- Cytosol, which contains unbound cTnT;
- Plasma, where the cTnT concentration, under physiological conditions, is negligible.

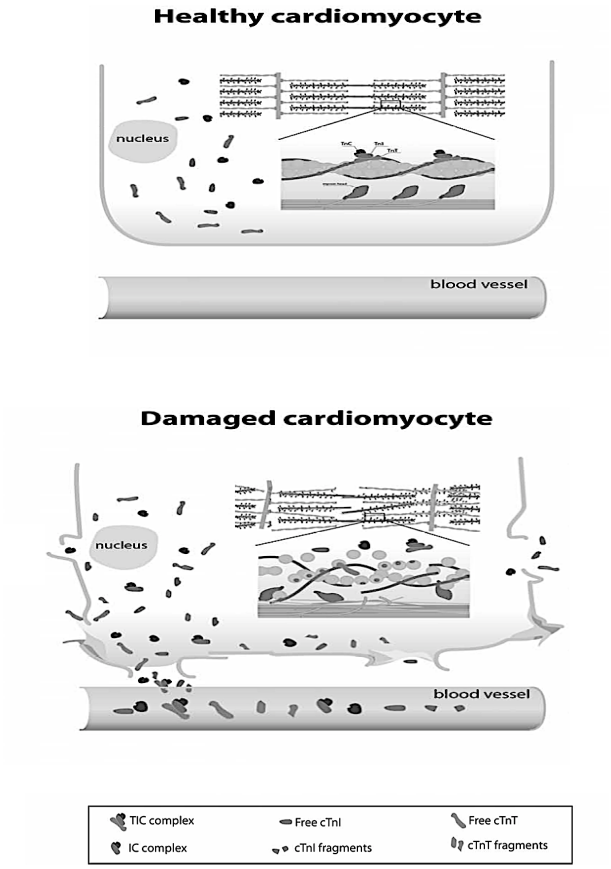


Figure 2: Compartmentalization of cTnT under physiological and pathological conditions (reproduced from [13]).

Fig. 2 shows cTnT compartmentalization in healthy and damaged cardiomyocytes. After the onset of AMI, the affected cardiac cells undergo necrosis, which causes a rupture of the outer cell membrane and of the sarcoplasmic reticulum, with the consequent release of the unbound cTnT molecules. The cTnT concentration in the plasma peaks to a maximum value, after 12-24 hours post-AMI, subsequently decreasing to reach a plateau phase, and in most cases a secondary peak in the subsequent days (see Fig. 1). One of the criteria to confirm the diagnosis of AMI is that plasma cTnT concentration exceeds the 99^{th} Upper Reference Limit, calculated on a reference population.

It is also important to take into account that the STEMI patients may be subjected to one or both of the following treatments: i) Early percutaneous revascularization, to restore the normal bloodflow downstream of an obstruction, by means of balloon dilation and or stent implantation of the occluded or stenotic vessel; ii) thrombolytic therapy, that consist in administration of drugs able to dissolve the blood clots in the occluded vessels. This group of patients usually undergoes coronary angiography at variable time after thrombolytic treatment, depending on patient's clinical characteristics and efficacy of the thrombolysis. Of course, the two treatments have an impact on troponin T release kinetics, affecting the correct estimation of the parameters, the resultant model and, therefore, the correct prognosis/diagnosis.

If we look at the current analysis techniques, we can see how these suffer from several problems, such as: 1) the raw experimental measurements, in most cases, do not allow the clinicians to identify significant values, needed to quantitatively assess the extent of the cardiac damage, like the peak values and the corresponding time point; 2) the exact time of the occurrence of the AMI is often uncertain as well, due

to either unconscious patient or neglected symptoms. In the next sections, we will present our strategy to address these issues, which is based on: i) the development of a dynamical model for the prediction of cTnT release from patient-specific experimental data and ii) the use of the model to compute patient-specific values of clinical interest.

3 Methods

3.1 Mathematical Model of cTnT Release

Cardiac troponin T is normally located into three distinct compartments: the sarcomere, the cytoplasm and the plasma. Therefore, the linchpin out of the model is the description of the fluxes of cTnT between these three compartments. To this regard, notice that the sarcomere is contained within the cell, therefore the model will not comprise any direct flux between the sarcomere and plasma compartments, but only between sarcomere and cytoplasm, and between cytoplasm and plasma.

Since cTnT levels into the sarcomere and cytoplasm compartments are not measurable, the model will be evaluated only with respect to its capability to fit the plasma concentration measurements. The release curve presents a biphasic pattern, which is generated by the following dynamics:

i) fast dynamics due to early and rapid leakage of the cTnT molecules contained in the cytoplasm, following the rupture of the outer cell membrane;

ii) slower dynamics due to the later release of cTnT molecules bound to the sarcomere; the basic assumption here is that the disassembly of the sarcomeric reticulum, with the consequent unbinding of the cTnT molecules, requires the persistent cell damage with respect to the membrane rupture.

The slower dynamics seem to be caused by the action of particular lytic enzymes that, ultimately, attack the subunit T of cardiac troponin, allowing it to detach from tropomyosin.

The state variables of our model are chosen to be the cTnT concentrations in the three compartment: C_s, C_c, and C_p for sarcomere, cytosol and plasma concentration, respectively. Usual simplifying assumptions are made: the concentration in the compartment is homogeneous and the flux between compartments is considered mono-dimensional. The structure of our model has been formulated in the form of the following ODE system

$$\begin{cases} \dot{C}_s(t) = -J_{sc}(t)\Gamma(t) \\ \dot{C}_c(t) = J_{sc}(t)\Gamma(t) - J_{cp}(t) \\ \dot{C}_p(t) = J_{cp}(t) - J_{pm}(t) \end{cases} \tag{1}$$

where the functions J_{sc}, J_{cp}, J_{pm} model the diffusion process of cTnT between the compartments and $\Gamma(t)$ is a function that modulates the sarcomeric rupture (and thus the flux between sarcomere and cytosol) over time.

Using Fick's first law of diffusion [14], we assume the flux between compartments to be proportional to the concentration difference, thus the diffusion terms take the following forms:

- $J_{sc} := D_{sc}\,(C_s(t) - C_c(t))$, the diffusion flux of cTnT between sarcomere and cytosol compartments, the positive flux direction is oriented from the sarcomere to the cytosol;
- $J_{cp} := D_{cp}\,(C_c(t) - C_p(t))$, the diffusion flux of cTnT between cytosol and plasma compartment, the positive flux direction is oriented from the cytosol to the plasma;
- $J_{pm} := \alpha_{cp}\,C_p(t)$, the clearance rate of cTnT from the circulatory system.

The parameters D_{sc}, D_{cp} represent the diffusion coefficients between sarcomere and cytosol, and cytosol and plasma, respectively, α_{cp} is the clearance rate from the plasma.

Note that in system (1) the diffusion process of cTnT between sarcomere and cytosol is obtained by multiplying the Fickian diffusion term J_{sc} by the modulator function Γ. Our choice for the modulator function is

$$\Gamma(t) \quad = \frac{t^3}{(t^3+T^3)} \tag{2}$$

which takes values between 0 (intact sarcomere, no cTnT release into the cytosol) and 1 (sarcomere completely disassembled, unconstrained release of cTnT into the cytosol). T represents the time threshold for sarcomeric disassembly. As illustrated Fig. 3, function Γ has a sigmoidal shape, with a slow initial growth for $t \ll T$ and a sudden rise to values close to 1 starting around $t = T$. Note that the time of switching between 0 and 1 can be tuned by changing the parameter T. By substituting the expressions of the flux and of the modulator function in system (1), the full model of cTnT release reads

$$\begin{cases} \dot{C}_s(t) = -D_{sc}\,(C_s(t) - C_c(t))\,\Gamma(t) \\ \dot{C}_c(t) = D_{sc}\,(C_s(t) - C_c(t))\,\Gamma(t) - D_{cp}\,(C_c(t) - C_p(t)) \\ \dot{C}_p(t) = D_{cp}\,(C_c(t) - C_p(t)) - \alpha_{cp}\,C_p(t) \end{cases} \tag{3}$$

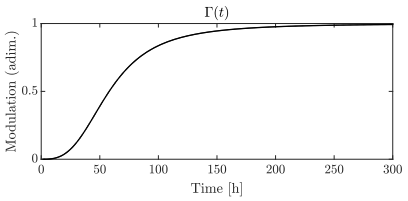


Figure 3: *Time evolution of* $\Gamma(t)$ *function with* $T = 60$*, in the range from 0 to 1.*

3.2 Experimental Dataset and Preprocessing

The experimental dataset was collected by the Interventional Cardiology Unit of the Magna Graecia University Hospital in Catanzaro, Italy, and includes clinical information and biomarkers measurements of 85 patients aged between 18 and 85 years who have been subject to an AMI event. The initial dataset has been examined by the clinical experts to remove patients with low number of acquisitions and/or atypical trend (which may be due to the occurrence of other pathologic events or treatments), leading to the exclusion of 14 patients from the present study. The dataset contains experimental data of STEMI patients with AMI. The measurements were taken using a highly sensitive assays, able to detect even very low levels of troponin in the blood. The *cut – off* value for this assay, calculated on a reference population, is equal to a concentration of 0.014 ng/ml. In general, the first acquisitions were performed every 6 hours from the time of admission for the first one/two days and then every 24 hours. Clearly, this is very dependent on many factors, e.g. symptoms severity, time elapsed from the ischemic event, other clinical evidences. A thrombolytic treatment was administered only to 22 patients, 7 of which did not exhibit an effective response to this first-line treatment and, consequently, underwent a *rescue* percutaneus revascularization. All remaining patients underwent primary percutaneous coronary intervention for myocardial reperfusion.

It is important to take into account that the thrombolytic treatment was always administered to patients at the first Emergency Unit (typically the closest hospital, which is often not the University Hospital); hence, the dataset does not contain acquisitions at time points preceding the admission to the University Hospital structure, as these were not directly available in our institutional database and used different analytical standards. To test the model, we used 60 patients as training dataset and the remaining 11 patients as validation dataset. We have used the training dataset to define the lower and upper bound and in the statistical and sensitivity analysis, while the validation dataset was used to test the hypotheses derived by

analyzing the results obtained from the training dataset. Before the analysis, a preprocessing of data was performed by visual inspection, under the supervision of the clinicians in order to identify potential outliers and correct them if possible.
Appropriate algorithms to identify potential outliers will be necessary in order to automate this step and make efficient the software for the model exploitation in a clinical setting.

3.3 Identifiability Analysis

The devised model comprises six parameters, all of them taking only positive values (negative values would lead to a biologically meaningless model): i) two diffusion coefficients D_{sc} and D_{cp}; ii) a clearance coefficient, α_{cp}; iii) time threshold for sarcomeric disassembly, T, previously introduced; iv) the initial concentrations of cTnT into sarcomere and cytosol or the initial conditions for the state variables, C_{s0} and C_{c0}, respectively. The initial concentration of the cTnT in the plasma is assumed to be negligible, therefore it is set equal to zero. These parameters are not available in the literature nor measurable (except perhaps the clearance rate of plasma cTnT, but at present this value is not available). Therefore, their values must be estimated by fitting the model to each patient experimental available data, that is to dynamic measurements of $C_p(t)$. Before applying a regression algorithm to estimate such parameters, it is important to carry on a thorough identifiability analysis. Indeed, to uniquely recover the unknown value of the parameters from the output evolution, it is necessary that the model is identifiable. The identifiability analysis can yield one of the following responses:

- structurally non-identifiable: it is impossible, due to the structure of the model, to determine the true value of the parameters from input-output experiments; in this case, it might be possible to overcome the limitation by reducing the complexity of the model;
- practically non-identifiable: the number of experimental data is too low or the data are too noisy; in this case a possible strategy could be to improve the number of acquisitions.

There are several methods for determining the structural identifiability of nonlinear models [15, 16]:

- the similarity transformation approach, based on the local state isomorphism theorem that requires to prove that the system is controllable and observable and it is usually computationally prohibitive for non-linear systems;
- the differential algebra methods based on replacing the stimuli-observables behavior of the system by some polynomial or rational mapping;
- the direct test solving directly the identifiability problem from uncontrolled and autonomous systems;
- the power series approaches like the the Taylor series and the generating series approach.

For some of these methods there are software tools available: DAISY [17], GenSSI [18], COMBOS [19] and STRIKE-GOLDD [20]. The choice depends on the features of the model [16]. For example, roughly speaking, DAISY requires the use of numerical calculations for very complex problems and STRIKE-GOLDD is very good for large systems but computes only local identifiability.

GenSSI was selected for our model, as a pure symbolic analysis is possible. GenSSI is based on the generating series approach, which allows to extend the analysis to the entire class of bounded and measurable stimuli when the model is linear in the stimuli.

We should note that our model does not exhibit control inputs, thus the generating series approach is equivalent to the Taylor series approach. The observable $y(t, \Theta) = C_p(t, \Theta)$, which is a function of the parameters $\Theta = [D_{sc}, D_{cp}, \alpha_{cp}, T, C_{s0}, C_{c0}]$, is expanded in a neighborhood of the initial state $[C_{s0}, C_{c0}, 0]$,

$$y(t, \Theta) = y(t_0, \Theta) + t\dot{y}(t_0, \Theta) + \frac{t^2}{2!}\ddot{y}(t_0, \Theta) + ... \tag{4}$$

and defining

$$a_k(\Theta) = \lim_{t \downarrow t_0^+} \frac{d^k}{dt^k} y(t, \Theta), \quad k = 0, 1, 2, ..., k_{max} \tag{5}$$

then a sufficient condition for global structural identifiability is given by

$$a_k(\Theta) = a_k(\Theta^*), \quad k = 0, 1, 2, ..., k_{max} \Longrightarrow \Theta = \Theta^*$$

where k_{max} is the smallest positive integer, such that the symbolic computations give the solution of the parameters (see [16] for a deeper discussion).

The structural identifiability analysis determined that model (3) with $k_{max} = 12$ was non-identifiable. This prompted us to carefully reconsider the structure of the model, eventually realizing that there is some redundancy in the parameters. In this case, the non-identifiability issue can be fixed by computing an equivalent scaled model, to reduce the number of free parameters. In our case this is readily obtained by dividing the left- and right-hand sides of the three differential equations in (3) by the diffusion coefficient D_{sc}, which yields

$$\begin{cases} \dfrac{dC_s}{d\tau} = -(C_s - C_c) \dfrac{\tau^3}{(\tau^3 + T_d^3)} \\ \dfrac{dC_c}{d\tau} = (C_s - C_c) \dfrac{\tau^3}{(\tau^3 + T_d^3)} - a\,(C_c - C_p) \\ \dfrac{dC_p}{d\tau} = a\,(C_c - C_p) - b\,C_p \end{cases} \tag{6}$$

where the parameters of the scaled model are $a = D_{cp}/D_{sc}$, $b = \alpha_{cp}/D_{sc}$, $T_d = T/D_{sc}$ and $\tau = D_{sc}\,t$. The scaling has reduced the number of parameters from six to five ($\Theta = [a, b, T_d, C_{s0}, C_{c0}]$). Now, applying the identifiability analysis, we can conclude that model (6) is locally structurally identifiable, thus it is possible to proceed with the parameters identification phase.

3.4 Parameters Identification

Parameter identification is a crucial step in the analysis of the collected experimental data, especially in the clinical context we deal with, since the proposed model sets out to fit the cTnT release curves of a heterogeneous population of patients, with potentially very different severity grading of ischemic attack and different ages, physical characteristics, co-morbidities and treatment records. Before setting up the identification algorithm, it is important to clarify what is the purpose of the estimated model. In many cases, especially in the control engineering field, the estimated model is intended to be used as an accurate proxy of the behavior of the real system, which can be used to predict the system behavior under different stimuli and initial conditions and/or to design the control law. If the system parameters can be reasonably assumed to be constant, the best strategy to tackle the regression problem is to find the best-fitting model on the overall dataset, i.e., computing a unique point in the parameter space that minimizes a cost function that involves all of the experimental data.

An interesting case, which is often encountered in biological studies, is when the system parameters are not deterministic, but can be assumed to belong to a known probability distribution. In this case, the regression problem can still be tackled using a cost function that includes the whole dataset, by describing the system in the form of a mixed-effects model, i.e., a model with parameters consisting of a constant value plus a random variable [21]. This class of models is usually the best choice when dealing with repeated measurements on the same subjects, e.g. in clinical longitudinal studies. In such cases, indeed, the parameters are expected to randomly vary around a fixed value.

Our problem cannot be classified in one of the above categories: in our setting, the variability of the parameters derives from the different characteristics of each patient and from the other clinical factors. This makes it impossible to estimate a model that satisfactorily fits the experimental data on multiple patients.

Each patient is different from the other in terms of peak values and times, initial concentration rate of climb, length of the plateau phase and settling time.

On the other hand, it is important to consider that our aim is not to model the average response over a population, but rather to analyze the response of each specific patient subjected to AMI and to extrapolate useful information on the basis of patient specific measurements. With this in mind, it is now clear that the parameter identification algorithm has to be applied to each experimental curve separately, thus deriving a unique parameter vector for each patient. Such parameter vector can be thought of as a set of quantitative features that characterizes each subject of our patients sample.

Since the model is nonlinear in the parameters, and it is locally structural identifiable, we had to resort to global optimization methods. Specifically, we tested two different techniques: i) the interior-point method [22], combined with multiple runs with a randomized choice of the initial guess over the parameter space, and ii) the particle-swarm optimization algorithm [23]; the two approach yielded almost identical results. The two global optimization strategies were applied using the corresponding software routines implemented in the Matlab Global Optimization Toolbox. Regarding the *multi-start* optimization procedure, we set the number of initial guesses to 40, which is sufficient to consistently find the global minimum in all of the examined cases.

On the other hand, the result of the fitting procedure heavily depends on the definition of the cost function that is minimized by the global optimization algorithm. Note that the typical release curve to be interpolated is characterized by an initial peak, which, in most cases, attains much higher values than the rest of the curve. In this case, a standard least square regression would favor the interpolation of the high values, to the detriment of the other points of the curve. To avoid this problem, we adopted a weighted least squares approach, choosing the weights of the residuals to be inversely proportional to the estimated values.

Another point to be carefully considered is the presence of outliers, which are inevitably introduced in operator-dependent experimental measurements of the type taken in a clinical scenario. In order to cope with this issue, a robust identification procedure has been implemented in combination with the weighted least squares. The robust implemented procedure belongs to the class of iteratively reweighted robust (IRR) algorithms [24]. The details are reported in the form of a pseudo-code in Algorithm 1.

Algorithm 1 Pseudo-code of the implemented variant of Iteratively Reweighted Robust algorithm

Init: $w \leftarrow w_i \leftarrow 1$
for iter = 1 to $\mathrm{N_{iter}}$ **do**
  $\theta_{\mathrm{opt}} = \arg\min_\theta \sum_k w_k \left(y_k - \hat{y}_k\right)^2$
  $r \leftarrow |y - \hat{y}|$
  if $r/N < B$ **then**
    $w_e \leftarrow (1 - (r/N)/B^2)^2$
  else
    $w_e \leftarrow 0$
  end if
  $w_i = \hat{y}^{-2}$
  $w \leftarrow w_e \cdot w_i$
end for

θ: model parameters; r: residuals; w_e, w_i, w: external, internal and total weights; y: experimental data; $\hat{y}$: estimated response; B: tuning parameter; N: normalization factor.

The IRR regression algorithm typically converges in a few iterations ($N_{\mathrm{iter}} = 5$ in our case). The implemented variant blends two weights: the internal weight compensates the magnitude differences of the data points, whereas the external one confers robustness to outliers. To this end, the latter weight is set to lower values for data points farther away from the current estimated curve and takes zero value when such distance exceeds a certain threshold, which can be tuned. The tuning parameter B was set to 4.68 and the normalization factor N to 1.48×median(r) as suggested in [24].

Finally, we had to define the admissible bounds for each parameter. In this case, the bounds cannot be derived from physical considerations, therefore the upper and lower bounds were selected heuristically, according to the following procedure: an initial guess has been evaluated on the basis of the typical rates of change and concentration values attained by the release curves. Subsequently, starting from this guess, we have gradually enlarged the admissible range of the parameter values, iteratively repeating the identification, and evaluated the distribution of the parameters over the population: the optimization ranges were fixed once most of the identified values were lying strictly between the bounds.

The calculations were carried out on a PC with Intel Core i7 processor, 16 Gb RAM and 2.5 GHz, resulting in a mean simulation time of 15 minutes for each patient.

The distribution of the estimated parameter values over our training set is reported in Fig. 4: the parameters T_d and C_{s0} exhibit a Gaussian-like trend, centered around a modal value, whereas a, b and C_{c0} appear to belong to long-tailed distributions. Therefore, we can conclude that the latter parameters are those that are mostly affected by inter-individual variability. This consideration can be exploited, as illustated in the following, in order to reduce the number of optimization parameters.

3.5 Confidence Intervals

The confidence intervals (CIs) of the estimated parameters allow us to evaluate the reliability of the fitting results: the tighter the CIs, the smaller the uncertainty on the estimated values.

For the computation of the CIs, first of all the original data have been used to get estimated release curves for each patient; then, the estimated curves have been corrupted with multiplicative noise, according to the error model $y_{\text{meas}} = y_{\text{true}}(1+\delta)$, using the estimated value, y_{est}, in place of the unknown true value y_{true}, and assuming $\delta \in \mathcal{N}(0, \bar{\sigma})$. The value of the standard deviation $\bar{\sigma} = 0.225$ has been derived by the analysis of the residuals.

Thus, a sample of 50 noise-corrupted time-series data has been generated and, by applying the identification procedure on each time-series, we have computed the corresponding sample of parameter vectors, which is used to compute the relative confidence intervals (i.e., each CI is normalized with respect to the average estimated value of the parameter) shown in Fig. 5.

This information, along with the results of the sensitivity analysis, is exploited to identify the most critical parameters for the identification procedure.

3.6 Optimal Experimental Design

The objective of the Optimal Experimental design (OED) is to maximize the confidence of the estimated parameters by designing the most informative experiments. The problem may be mathematically formulated as a general dynamic optimization problem searching for those manipulable variables (time-dependent stimuli, initial experimental conditions, experiment durations, sensor locations, sampling times or type of observables) that maximize the information.

In this work, the objective is to minimize the uncertainty of the unknown parameters ($\Theta = [a, b, T_d, C_{s0}, C_{c0}]$) by selecting the best sampling times (t_s). The method assumes that each measurement C_p is independent and identical distributed (i.i.d) normal random variable. The mean ($\bar{C_p}$) is obtained from the model in (6).

The Fisher Information Matrix (FIM) is the standard measure for the amount of information that measurements carry about an unknown parameter. The FIM is defined as the variance of the score:

$$\mathscr{F}(\Theta, t_s) = E\left\{\left(\frac{\partial J_{ml}}{\partial \Theta}\right)\left(\frac{\partial J_{ml}}{\partial \Theta}\right)^{\mathrm{T}}\right\} \tag{7}$$

where J_{ml} is the negative log-likelihood function [25]. For independent and identical distributed (i.i.d.) normal error the Fisher information is equivalent to:

$$\mathcal{F}(\Theta, t_s) = \frac{1}{\sigma^2}\mathbf{E}\left\{\left(\frac{\partial J_m(\Theta, t_s)}{\partial \Theta}\right)^T\left(\frac{\partial J_m(\Theta, t_s)}{\partial \Theta}\right)\right\} \tag{8}$$

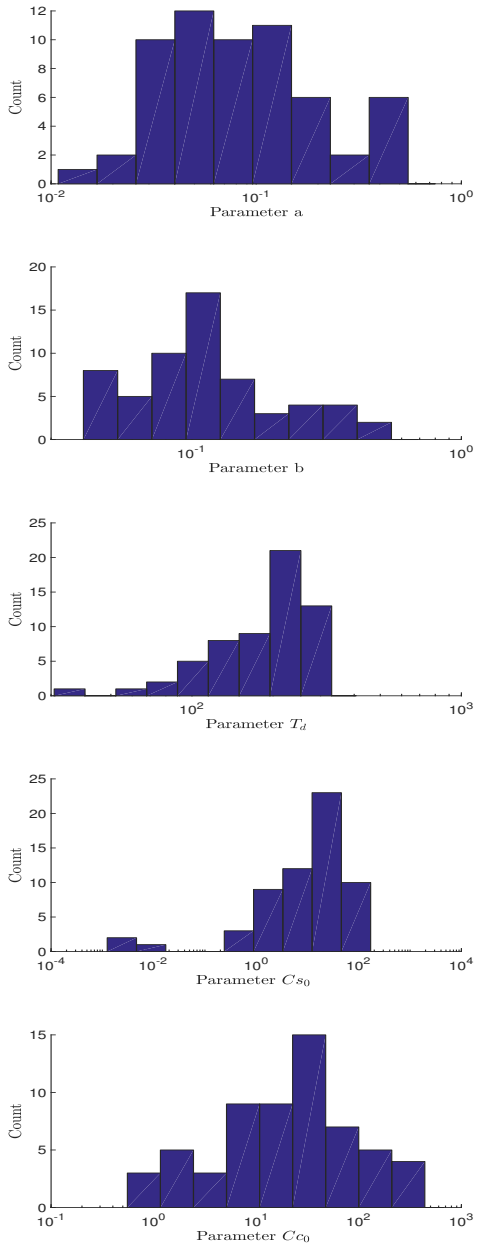


Figure 4: Distribution of estimated parameter values over the patients population.

where $J_m(\Theta, t_s)$ is the sum of the square errors and σ its standard deviation.

Therefore the problem is formulated as the following dynamic optimization problem: *Calculate best sampling times so as to optimize a scalar measure of the FIM (J_{OED}).*

Different scalar functions of the FIM are formulated (J_{OED}). The FIM determines typically a hyperellipsoid in the parameter space. This hyperellipsoid represents the quantity and quality of information of the selected experiments. The largest and the more spherical the hyperellipsoid defined by the FIM, the better the experimental design. Common criteria are:

- D criterion that maximizes the volume of the information hyperellipsoid but not its shape. The higher

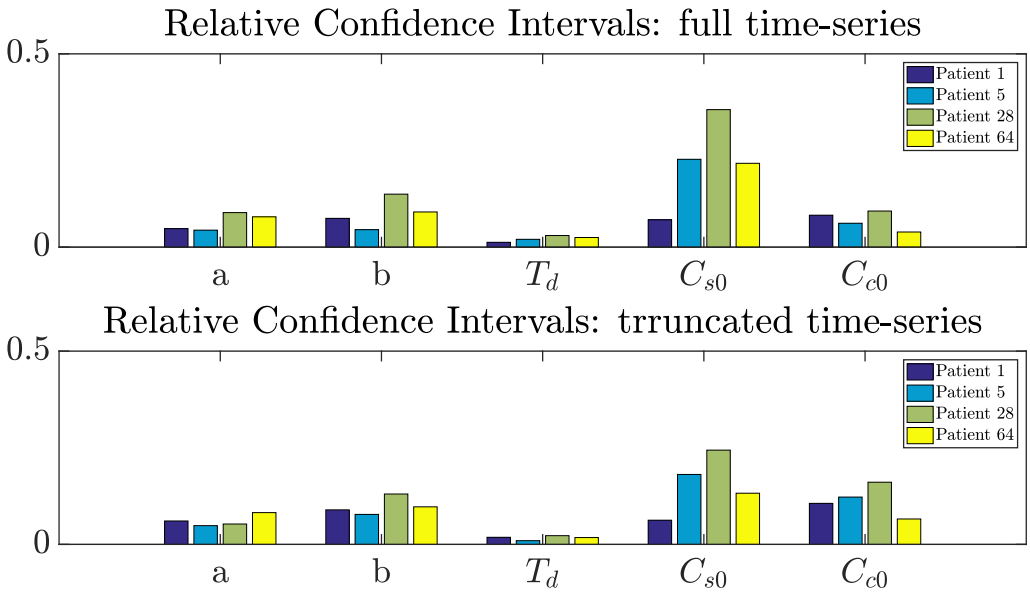


Figure 5: *Relative confidence intervals calculated on four patients, using the full time-series (top panel) and the truncated time-series (low panel, see Section 4.1 for details on the truncated time-series).*

its value the biggest the expected parameter confiance for the parameter estimates.

$$J_D = \max \det(\mathscr{F})$$

- A criterion that maximizes the arithmetic mean of the hyperellipsoid semi-axes.

$$J_A = \max trace(\mathscr{F})$$

- E criterion that maximizes the minimum semi-axis of the hyperellipsoid and offers a compromise between D and $Emod$

$$J_E = \max \lambda_{min}(\mathscr{F})$$

- Modified E criterion ($Emod$) that minimizes the relationship between the longest and shortest semi-axes of the hyperellipsoid, i.e., improves the eccentricity of the information hyperellipsoid. The global optimal solution is $J_{Emod} = 1$, when the parameter confidence is equally distributed

$$J_{Emod} = \min \frac{\lambda_{max}(\mathscr{F})}{\lambda_{min}(\mathscr{F})}$$

Generally speaking, D and A are criteria to improve information and $Emod$ to decorrelate parameters. E criterion is the option to optimize a compromise between improving information and parameter decorrelation.

Numerics of the optimization depends on the criterion selected. A criterion is insensitive to singular FIM matrices. Thus A may lead to undesirable results as was already pointed out by [26] and has not been used in this work. E criterion is non-differentiable and requires the use of appropriate global optimizers [27]. In this work we selected the global optimizer based on scatter search (eSS, Enhanced Scatter Search) method [28]. This algorithm can optimize non-differentiable functions and it is very efficient in finding the best experimental designs. Last, when parameters are of different orders of magnitude, FIM is usually badly

conditioned and criteria such as $E_{mod} = 1$ are not necessarily decorrelating the parameters (see [29] for details). To avoid these problems we used the normalization proposed in [30].

Computations of the different optimal experimental designs require advanced numerical techniques. In this work we used AMIGO2 (Advanced Model Identification using Global Optimization), a multi-platform toolbox implemented in Matlab [31]. The optimization of the FIM was computed using the common control vector parametrization technique, which transforms the original infinite dimension optimization problem into a non-linear programming problem [32].

4 Results

4.1 Estimation of Patient-Specific Responses

The system identification approach illustrated in Section 3.4 was first applied to the training set, to evaluate the distribution of the parameters (reported in Fig. 4): to give a feeling of how the model fits the experimental data, in Fig. 6 we report the estimated response for 4 patients. Since the use of the standard R^2 index is debatable in the case of nonlinear regression, in order to assess the quality of the interpolation results on each curve, we evaluate the Average Normalized Residual (ANR), defined as

$$ANR = \frac{1}{N_p} \sum \sqrt{\frac{(y_{\text{meas}} - y_{\text{est}})^2}{y_{\text{est}}^2}}\,, \tag{9}$$

where y_{meas} represents the experimental acquisition, y_{est} the vector of the estimated values, and N_p the number of the experimental acquisitions, for each curve.

The distribution of the ANR metrics for the training and validation datasets, along with the results obtained optimizing all the 5 parameters (hereafter we refer to this case as 5-parameters model), are shown in Fig. 7. Note that smaller ANR values are associated to better interpolation.

The fitting results are generally good, which suggests that the devised model provides a satisfactory description of the cTnT release dynamics. The usefulness of the model, so far, consists in the possibility of recovering a detailed picture of the evolution of cTnT concentration after AMI, starting from a few experimental samples.

To further evaluate the applicability and usefulness of the model, we decided to investigate how the estimated response changes when some experimental data points are removed (the identification algorithm is run on a truncated time-series). This resembles the most very common situation that occurs when the patient are presented to the hospital after several hours, or even days, after the onset of AMI. Under this circumstance, the initial data points of the curve would not be available and the clinician would not be able to estimate the time and amplitude of the first cTnT peak. Thus, we challenged our model to recover this information. As shown in Fig. 6, on the same 4 patients we simulated the situation above, by removing all of the initial data points until the first peak (marked with green crosses), before performing the model identification step. The resulting estimated responses (solid red line) show that we are still able to identify with good approximation the first peak time point, although it is not possible to precisely estimate the amplitude of the peak. A look at Fig. 5 confirms that, as expected, the CIs obtained with the truncated time-series data are higher than with the full time-series data. Thus, the proposed model confirms as an effective tools not only for the interpolation of the experimental data, but also for extrapolating values of the release curve outside the experimental observation window.

Particular attention must be paid when the number of experimental data points is too low with respect to the number of optimization parameters. In these situations, the model is expected to overfit the data and the estimated response is, therefore, not reliable. An example of this is represented by patient 28 estimated response on the truncated time-series: having only 4 points left (and 5 optimization parameters) the model overfits the data (though, in this case, the time of the estimated peak is still quite close to the one estimated on the full time-series). A possible way to overcome this issue, when it is not possible to increase the number of acquisitions, consists in reducing the number of optimization parameters, by setting one or more parameters to a fixed value.

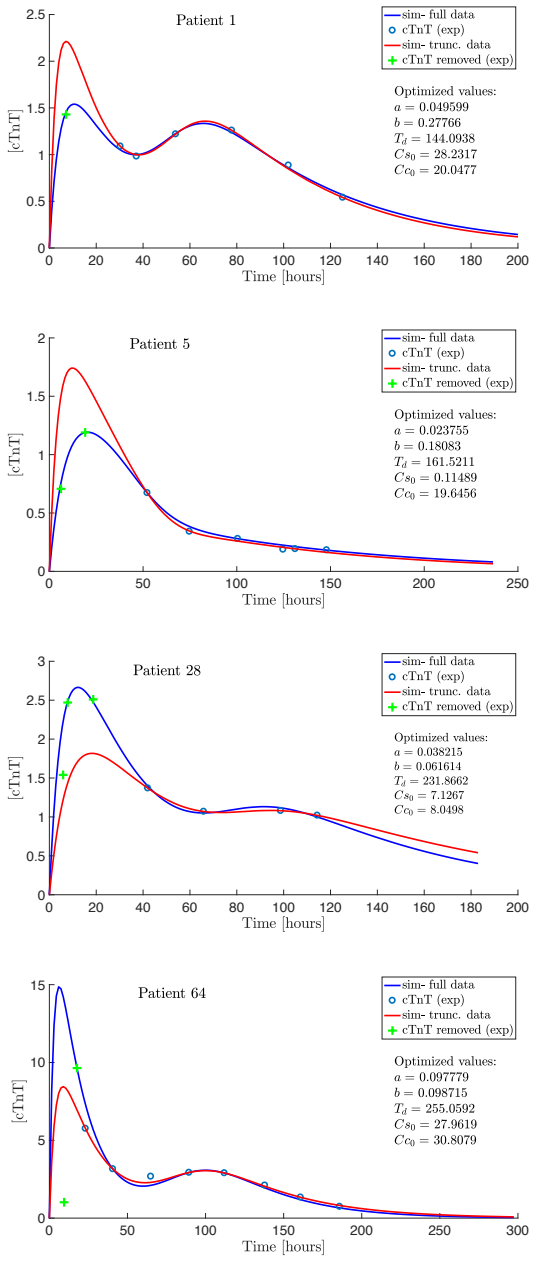


Figure 6: Model fitting using the full and truncated dataset. The blue curves (–) are obtained by fitting the model to all the experimental time-points. The red curves (–), instead, are obtained by removing the time-points marked with a green cross(+), to simulate analysis of patients admitted at the hospital after the concentration peak has been attained.

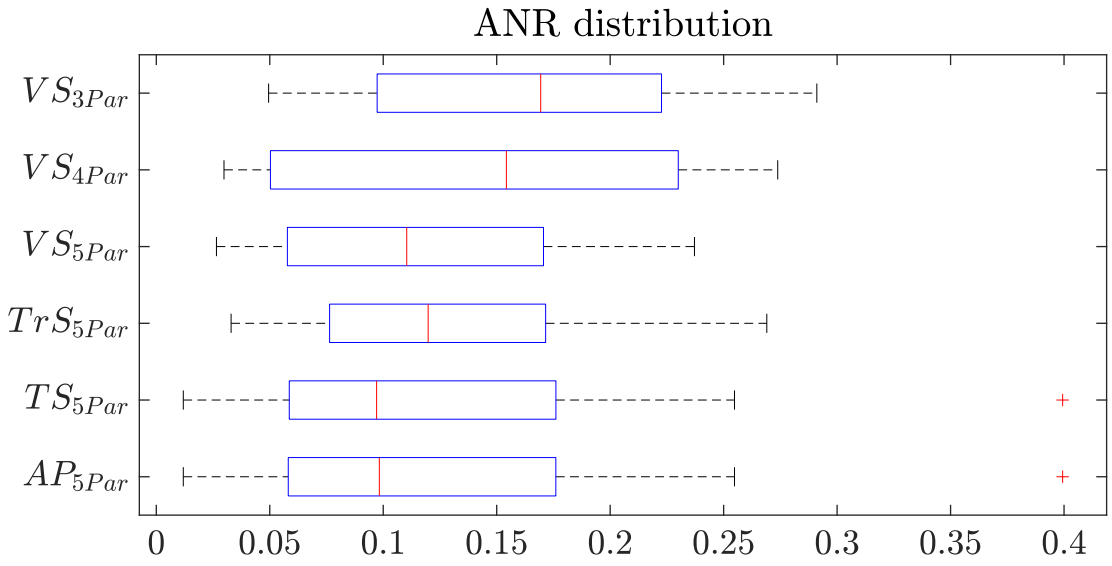


Figure 7: Distribution of the ANR index over the dataset, using the 5 parameters model on the whole (AP), training (TS), truncated (TrS) and validation (VS) dataset, and the 4- and 3- parameters models on the validation (VS) dataset.

To evaluate the capability of the model to estimate the first missing acquisitions, we computed the ANR index on the results obtained from the experiment described above. In this case, to avoid overfitting problems when reducing the dataset, we have selected 46 patients such that the truncated dataset would still contain at least 6 acquisitions. The distribution of the ANR metrics for these 46 patients are shown in Fig.7 (label $TrS_{5\mathrm{Par}}$).

4.2 Sensitivity Analysis

To gain a deeper insight on how the various parameters influence the response of model (6), we performed sensitivity analysis (SA) [33]. More specifically, in our case SA has been applied to identify the parameters that have major influence on the two peaks of the curve, corresponding to the two release phases. Properly identifying these values is of great importance to the clinician for diagnostic purposes.

We have performed a normalized local SA [34], numerically computing the first-order partial derivatives of the quantities of interest, with respect to each parameter, that is

$$S_i = \frac{\partial y_i / y_i}{\partial p / p} = \frac{\partial ln(y_i)}{\partial ln(p)}$$

Numerically, the sensitivity analysis was computed by varying the parameters by a small amount (around 1%), computing the corresponding changes in the peaks and then the ratio.
The results of the sensitivity analysis are reported in Fig. 8: the first peak on the curve is mainly affected by the initial concentration of cTnT into the cytosol, C_{c0}, and by the normalized diffusion coefficients a and b. The second peak, instead, is influenced by the normalized diffusion coefficient b, by the initial concentration of cTnT into the sarcomere, C_{s0}, and by the threshold of the modulator function, T_d.

4.3 Estimated response with some fixed parameters

As discussed above, a critical requisite, to get useful response estimates from the proposed cTnT model, is that the number of experimental acquisitions is not too low. This requirement would hinder the application

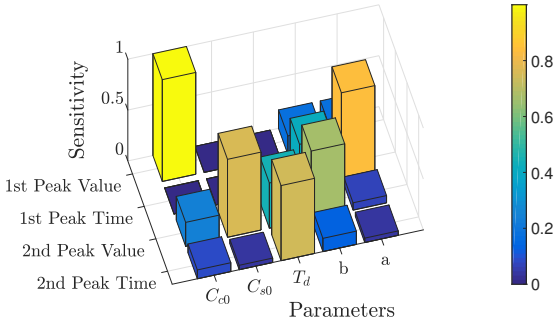


Figure 8: *Sensitivity analysis for the peak values and their time, by increasing the parameter values by* 1% *with respect to the optimal starting value.*

of the devised estimation approach in the first hours after the onset of AMI, when only 2-3 points are typically available.

A possible strategy, to enable response estimation also during this first phase, may consist in reducing the number of optimization parameters, so as to avoid the problems of inconsistency and reliability of the results associated to overfitting. On the other hand, to reduce the number of optimization parameters, we need to fix one or more parameters to a preassigned value. A sensible choice would be to fix the parameters that have the lowest influence (i.e., low sensitivity) on the response shape and that exhibit the lowest variability among our patients sample. The selection of the fixed parameters and of the preassigned values was done on the basis of the empirical distributions of the parameters and of the sensitivity analysis computed in the previous sections.

In particular, we have chosen to fix T_d and C_{s0}, because these parameters have a negligible influence on the first peak, which enables us to provide good estimates of the first peak even with fewer optimization parameters. We computed the estimated response i) with only one fixed parameter, namely $C_{s0} = 10$ (this case is denoted as (4- parameters model), and ii) with two fixed parameters, namely $C_{s0} = 10$, $T_d = 200$ (3-parameters model).

Some estimation results obtained with the 5-, 4-, and 3-parameters models are compared in Fig. 9. Note that the tests were run on patients of the validation set, whereas the preassigned values of the fixed parameters have been chosen (approximately) as the median values of the empirical distributions obtained on the training dataset. Interestingly, the degradation of the estimated responses when fixing one or two parameters does not appear to be dramatic (see Fig.7, labels $VS_{4\text{Par}}$ and $VS_{3\text{Par}}$), although a deeper investigation on a larger dataset would be required to definitely state the effectiveness of this strategy.

4.4 Optimal Experimental Design Analysis

The OED problem for model (6) has been solved using various objective functions, to test the best solution for our case. More specifically, we run OED with the criteria Dopt, Eopt and Emod; for each criterion, we computed the optimal sampling times, assuming that it is possible to take 5, 10, and 15 measurements. We consider the following optimization constrations: intervals between successive measurements should be of at least half an hour, and measurements cannot be taken during the first hour. Once the optimal sampling times were calculated, we compute the CIs corresponding to the optimized experimental data points. Since the real concentration values are not available at the optimal time-points, they have been replaced by the values estimated using the originally available experimental time-series. These estimated optimal data points are corrupted with a multiplicative noise vector before applying the identification algorithm.

The Eopt criterion gave the best results, probably because it provides a balance between information maximization and parameters decorrelation. The other two criteria only focus on one objective, either information maximization (Dopt) or parameters decorrelation (Emod). Interestingly, Eopt performed better

than Emod, which suggests that the parameters of our model are not highly correlated

In Fig. 10 we report the OED time points for two cases with only 5 acquisitions. The results of the OED analysis suggest that most of the measurements should ideally be located just before the first peak is attained, possibly repeating the acquisition 3-4 times in the first 4 hours when the cTnT concentration rapidly increases and leaving only one acquisition in the intermediate plateau phase. This choice is quite different with respect to the classical experimental routine, which generically aims to cover the whole period of cTnT release with a much more regular sampling (typically on a daily basis).

To further confirm the indications of OED, we repeated the OED optimization using a generic patient's response (i.e., setting each parameter equal to the median value of the corresponding empirical distribution). For this generic patient, we computed the optimal sampling times with 10 and 15 measurements. The results, reported in Fig. 11 are consistent with those found with 5 points: the optimal allocation of the measurements is not obtained with a regular spanning of the interval, but concentrating the points mostly around the first peak and then a few points allocated in the later days.

Clearly, an immediate consequence of these OED results, should they be confirmed by experimental tests, would be the possibility to significantly reduce the discomfort of the patient, who is currently compelled to undergo daily blood tests for a period of one/two weeks after the ischemic event. If the model was proven to provide reliable patient-specific estimates of the cTnT evolution, it would imply the necessity of a smaller number of experimental measurements to be performed, with consequences also in terms of diagnostic and follow-up costs.

5 Discussion and Conclusions

The main contribution of the present work consists in the development of a novel mathematical model of cTnT release into the plasma after the onset of AMI. The model was built using a bottom-up mechanistic approach, starting from the dynamics underlying the diffusion of species of interest between the involved compartments. This is, to our knowledge, the first model accounting for the dynamics of cardiac biomarker release. Once the model structure was devised, we performed a structural identification analysis, a step that is often neglected in system identification, but that is crucial to obtain consistent findings, especially when dealing with biological systems, which are intrinsically subject to high inter-individual variability. This analysis enabled us to derive a structurally identifiable version of the cTnT release model, which was subsequently used for the identification of the parameters and estimation of the response on a sample of patients, whose data were collected by the interventional cardiology unit at our University Hospital in Catanzaro.

The model interpolation results show a good agreement with the experimental data and the capability of reproducing the characteristics biphasic kinetics of cTnT release. In particular, we have shown that the model, along with the proposed identification approach, can be effectively exploited to estimate the patient-specific cTnT concentration time-course, both between experimental samples (interpolation) and outside the observational window (extrapolation). identification of clinically relevant patient-specific features. Although model simulations cannot replace clinical acquisitions, fitting the model to the available acquisitions enables the estimation of otherwise unavailable clinical data, like cTnT concentration evolution before the patient hospitalization and, in particular, the time of occurrence of the peak concentration, as shown by the tests illustrated above, which is a valuable information for a correct diagnosis and prognosis. Furthermore, the model is expected to prove itself a useful tool also for other purposes, e.g., in retrospective clinical studies.

We have pointed out that the parameter identification procedure requires a minimum number of experimental data points, which poses the problem of how to exploit the model in the early phase after the ischemic event, when few measurements are available. A possible strategy to cope with this issue has been proposed, which seems to yield promising results.

Overall, the results of our work suggest that the devised model can be leveraged as a quantitative diagnostic tool, e.g., to give estimates of the the time of occurrence and the extent of the AMI, to assess the current stage of the pathology and to predict its evolution in the short term.

Furthermore, the optimal choice of the sampling times was also investigated, using the methods of Optimal

Experimental Design. The results of this optimization are clinically interesting, indeed they suggest that, in principle, to obtain effective information on the cTnT response, it is not necessary to take measurements on a daily basis, but it could be sufficient to have a tight sampling in the first hours and then an acquisition in the plateau phase.

Future work will focus on the extension and refinement of the model along several directions: firstly, an important study will regard the role played by additional risk factors like, for instance, diabetes, renal insufficiency or previous myocardial damage, on the release kinetics of cTnT; secondly, we will explore the advantages of linking the troponin release model with other cardiac disesases biomarkers, like creating kinase (CK-MB). Furthermore, a particular attention will be devoted to further validation of the model, both expanding the patients database and by using robustness analysis methods, which can be effectively used as an alternative (with respect to data fitting) means of model validation as discussed in [35], [36]. Finally, a crucial point for the applicability of the methods described in this work will be the development of a simple and efficient software platform for the real-time automated analysis of the experimental cTnT release data, to be used by clinicians during the treatment of AMI patients.

Acknowledgment

M.R.G. acknowledges financial support from the Spanish Government and the European Regional Development Fund through the project DPI2014-54085-JIN.

References

[1] B. Ibanez *et al.*, "2017 esc guidelines for the management of acute myocardial infarction in patients presenting with st-segment elevation: The task force for the management of acute myocardial infarction in patients presenting with st-segment elevation of the european society of cardiology (esc).," *European Heart Journal*, vol. 39, no. 2, pp. 119–177, 2018.

[2] M. Roffi *et al.*, "2015 esc guidelines for the management of acute coronary syndromes in patients presenting without persistent st-segment elevation: Task force for the management of acute coronary syndromes in patients presenting without persistent st-segment elevation of the european society of cardiology (esc).," *European Heart Journal*, vol. 37, no. 3, pp. 267–315, 2016.

[3] F. Rodriguez and K. W. Mahaffey, "Updates to the accf/aha and esc stemi and nstemi guidelines: putting guidelines into clinical practice," *J. Am. Coll. Cardiol.*, vol. 68, no. 3, pp. 313–321, 2016.

[4] C. Cosentino, W. Curatola, F. Montefusco, M. Bansal, D. Di Bernardo, and F. Amato, "Linear matrix inequalities approach to reconstruction of biological networks," *IET Systems Biology*, vol. 1, no. 3, pp. 164–173, 2007.

[5] F. Montefusco, C. Cosentino, and F. Amato, "Core-net: exploiting prior knowledge and preferential attachment to infer biological interaction networks," *IET Systems Biology*, vol. 4, no. 5, pp. 296–310, 2010.

[6] L. Salerno, C. Cosentino, G. Morrone, and F. Amato, "Computational modeling of a transcriptional switch underlying b-lymphocyte lineage commitment of hematopoietic multipotent cells," *PloS one*, vol. 10, no. 7, p. e0132208, 2015.

[7] H. M. Byrne, "Dissecting cancer through mathematics: from the cell to the animal model," *Nature Reviews Cancer*, vol. 10, no. 3, pp. 221–230, 2010.

[8] D. Pe'er and N. Hacohen, "Principles and strategies for developing network models in cancer," *Cell*, vol. 144, no. 6, pp. 864–873, 2011.

[9] S. S. Hori and S. S. Gambhir, "Mathematical model identifies blood biomarker–based early cancer detection strategies and limitations," *Science Translational Medicine*, vol. 3, no. 109, pp. 109ra116–109ra116, 2011.

[10] R. Roberts, P. D. Henry, and B. E. Sobel, "An improved basis for enzymatic estimation of infarct size.," *Circulation*, vol. 52, no. 5, pp. 743–754, 1975.

[11] A. "Procopio and et al", "Predictive mathematical model of cardiac troponin release following acute myocardial infarction," in *Proceedings of the 14th IEEE International Conference on Networking, Sensing and Control*, (Calabria, Italy), May 16-18, 2017.

[12] A. "Procopio and et al", "A model of cardiac troponin T release in patient with acute myocardial infarction," in *Proceedings of the 56th IEEE Conference on Decision and Control*, (Melbourne, Australia), December 12-15, 2017.

[13] L. H. J. Jacobs, *The release of cardiac troponin: when, where and how.* PhD thesis, Maastricht University, 2012.

[14] J. Crank, *The mathematics of diffusion.* Oxford university press, 1979.

[15] M. J. Chapman, K. R. Godfrey, M. J. Chappell, and N. D. Evans, "Structural identifiability for a class of non-linear compartmental systems using linear/non-linear splitting and symbolic computation," *Mathematical biosciences*, vol. 183, no. 1, pp. 1–14, 2003.

[16] O. Chis, J. R. Banga, and E. Balsa-Canto, "Structural identifiability of systems biology models: a critical comparison of methods," *PloS one*, vol. 6, no. 11, p. e27755, 2011.

[17] G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò, "Daisy: A new software tool to test global identifiability of biological and physiological systems," *Computer methods and programs in biomedicine*, vol. 88, no. 1, pp. 52–61, 2007.

[18] O. Chiş, J. R. Banga, and E. Balsa-Canto, "Genssi: a software toolbox for structural identifiability analysis of biological models," *Bioinformatics*, vol. 27, no. 18, pp. 2610–2611, 2011.

[19] N. Meshkat, C. E.-z. Kuo, and J. DiStefano III, "On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and combos: a novel web implementation," *PloS one*, vol. 9, no. 10, p. e110261, 2014.

[20] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou, "Structural identifiability of dynamic systems biology models," *PLOS Computational Biology*, vol. 12, no. 10, p. e1005153, 2016.

[21] M. Lindstrom and D. Bates, "Nonlinear mixed effects models for repeated measures data," *Biometrics*, vol. 46, no. 3, pp. 673–687, 1990.

[22] R. Waltz, J. Morales, J. Nocedal, and D. Orban, "An interior algorithm for nonlinear optimization that combines line search and trust region steps," *Mathematical programming*, vol. 107, no. 3, pp. 391–408, 2006.

[23] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, (Perth, Australia), November 27 - December 1, 1995.

[24] L. Tóthfalusi and L. Endrényi, "Algorithms for robust nonlinear regression with heteroscedastic errors," *International journal of bio-medical computing*, vol. 42, no. 3, pp. 181–190, 1996.

[25] S. M. Kay, *Fundamentals of statistical signal processing, volume I: estimation theory.* Prentice Hall, 1993.

[26] R. Mehra, "Optimal input signals for parameter estimation in dynamic systems–survey and new results," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 753–768, 1974.

[27] D. Telen, N. Van Riet, F. Logist, and J. Van Impe, "A differentiable reformulation for e-optimal design of experiments in nonlinear dynamic biosystems," *Mathematical Biosciences*, vol. 264, pp. 1–7, 2015.

[28] J. A. Egea, R. Martí, and J. R. Banga, "An evolutionary method for complex-process optimization," *Computers & Operations Research*, vol. 37, no. 2, pp. 315–324, 2010.

[29] M. R. García, *Identification and real time optimisation in the food processing and biotechnology industries*. PhD thesis, Universidade de Vigo, 2008.

[30] M. García, A. Alonso, and E. Balsa-Canto, "A normalisation strategy to optimally design experiments in computational biology," *Advances in Intelligent Systems and Computing*, vol. 616, pp. 126–136, 2017.

[31] E. Balsa-Canto, D. Henriques, A. Gabor, and J. R. Banga, "Amigo2, a toolbox for dynamic modeling, optimization and control in systems biology," *Bioinformatics*, vol. 32, no. 21, p. 3357, 2016.

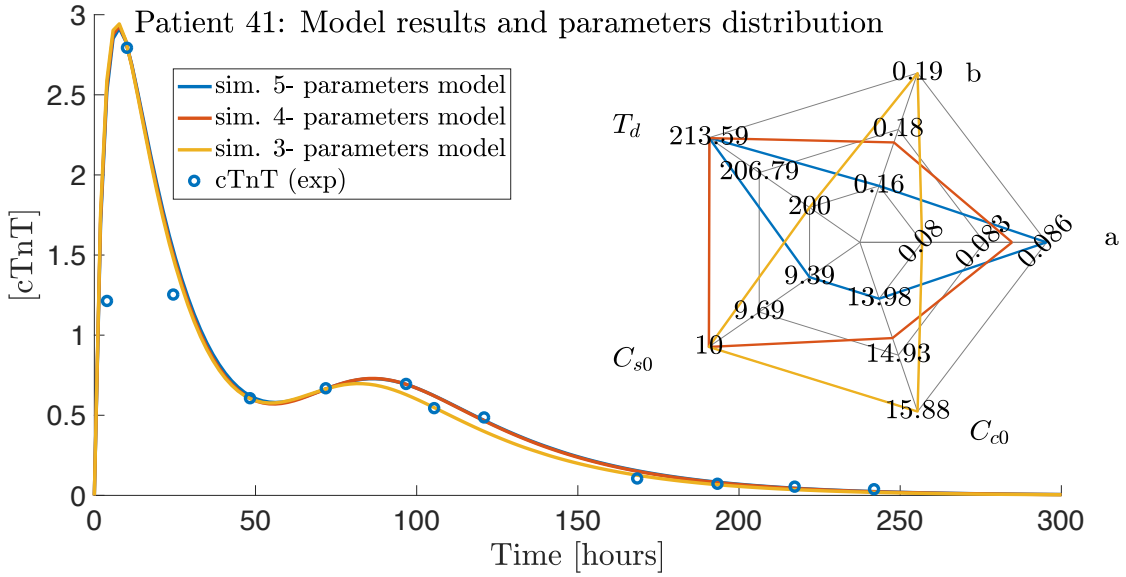
[32] E. Balsa-Canto, A. A. Alonso, A. Arias-Méndez, M. R. García, A. López-Núñez, M. Mosquera-Fernández, C. Vázquez, and C. Vilas, "Modeling and optimization techniques with applications in food processes, bio-processes and bio-systems," in *Numerical Simulation in Physics and Engineering*, pp. 187–216, Springer, 2016.

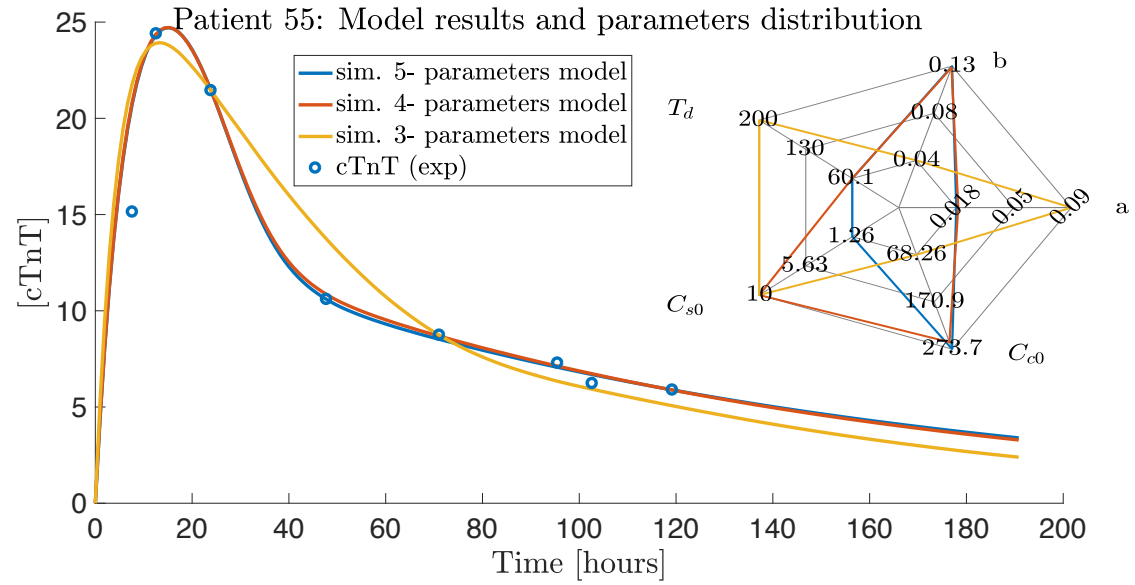
[33] M. Eslami, *Theory of sensitivity in dynamic systems: an introduction*. Springer Science & Business Media, 2013.

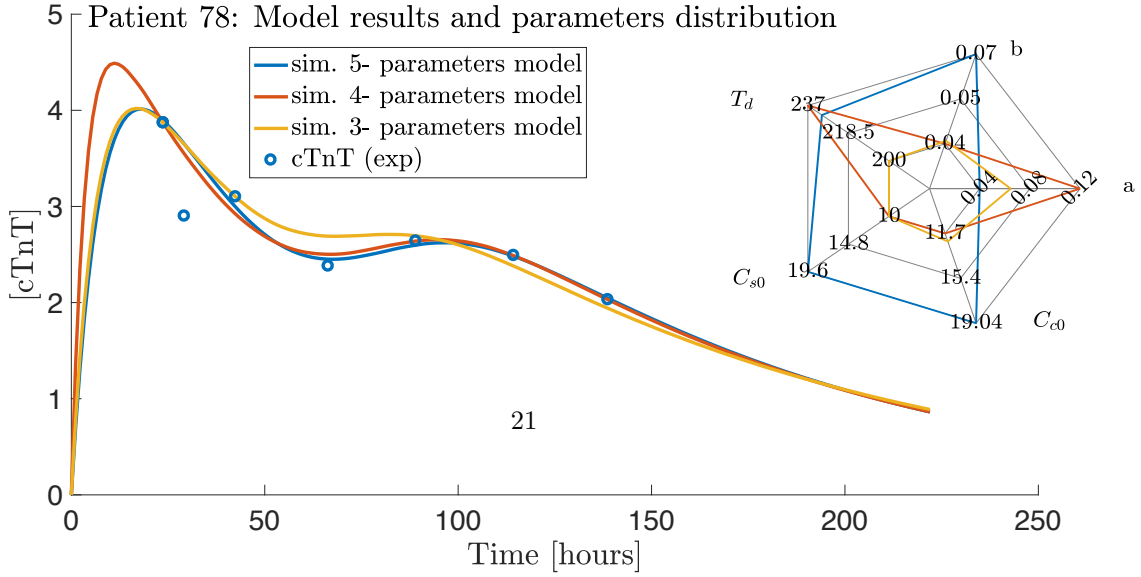
[34] Z. Zi, "Sensitivity analysis approaches applied to systems biology models," *IET systems biology*, vol. 5, no. 6, pp. 336–346, 2011.

[35] D. G. Bates and C. Cosentino, "Validation and invalidation of systems biology models using robustness analysis," *IET Systems Biology*, vol. 5, no. 4, pp. 229–244, 2011.

[36] L. Salerno, C. Cosentino, A. Merola, D. Bates, and A. Francesco, "Validation of a model of the GAL regulatory system via robustness analysis of its bistability characteristics," *BMC Systems Biology*, vol. 7, no. 39, 2013.


Patient 41: Model results and parameters distribution
sim. 5- parameters model
sim. 4- parameters model
sim. 3- parameters model
cTnT (exp)
[cTnT]
Time [hours]
T_d
C_{s0}
C_{c0}
a
b
0.19
0.18
0.16
0.08
0.083
0.086
213.59
206.79
200
9.39
13.98
9.69
10
14.93
15.88


Patient 55: Model results and parameters distribution
sim. 5- parameters model
sim. 4- parameters model
sim. 3- parameters model
cTnT (exp)
[cTnT]
Time [hours]
T_d
C_{s0}
C_{c0}
a
b
0.13
0.08
0.04
0.018
0.05
0.09
200
130
60.1
1.26
68.26
5.63
10
170.9
273.7


Patient 78: Model results and parameters distribution
sim. 5- parameters model
sim. 4- parameters model
sim. 3- parameters model
cTnT (exp)
[cTnT]
Time [hours]
T_d
C_{s0}
C_{c0}
a
b
0.07
0.05
0.04
0.04
0.08
0.12
237
218.5
200
10
11.7
14.8
19.6
15.4
19.04

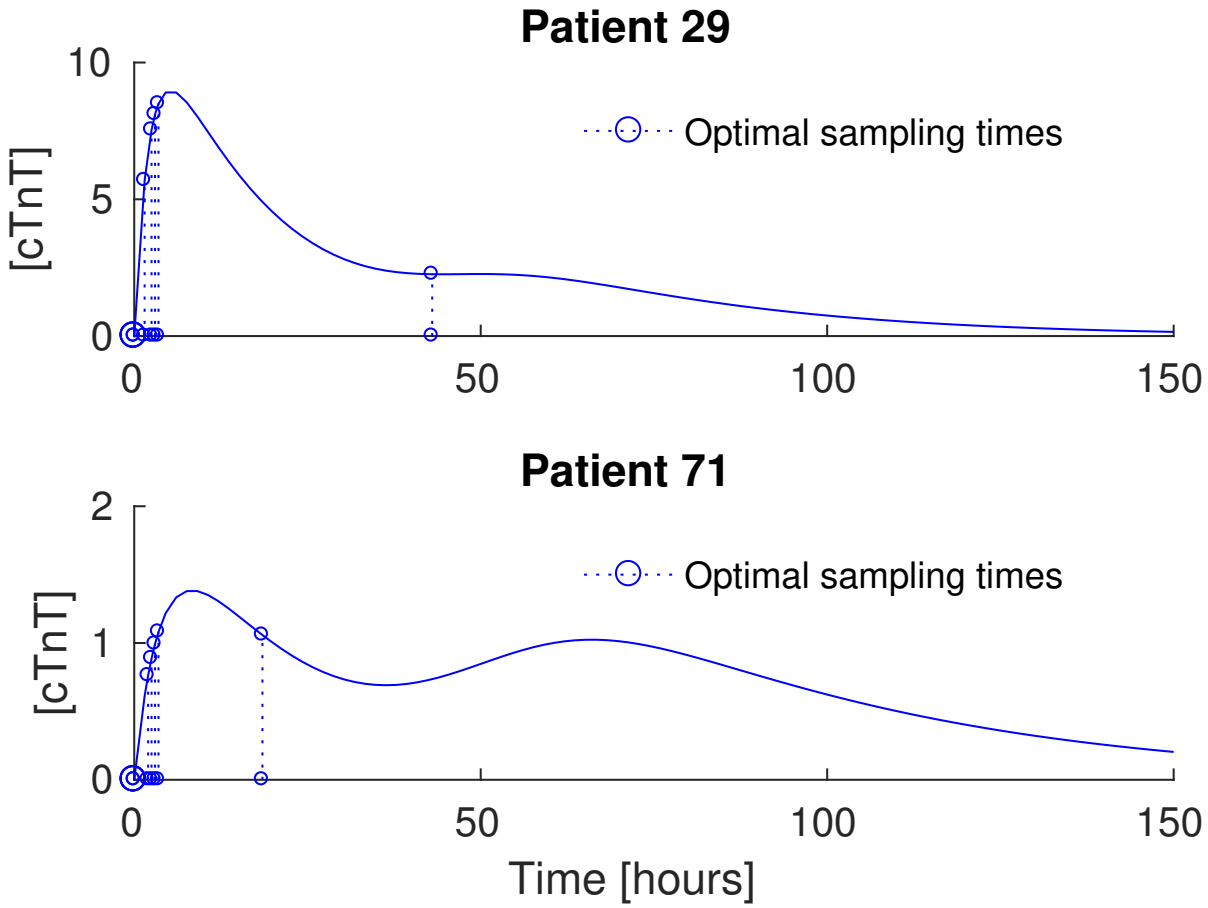


Figure 10: *Optimal sampling times with Eopt for patient 29 and 71.*

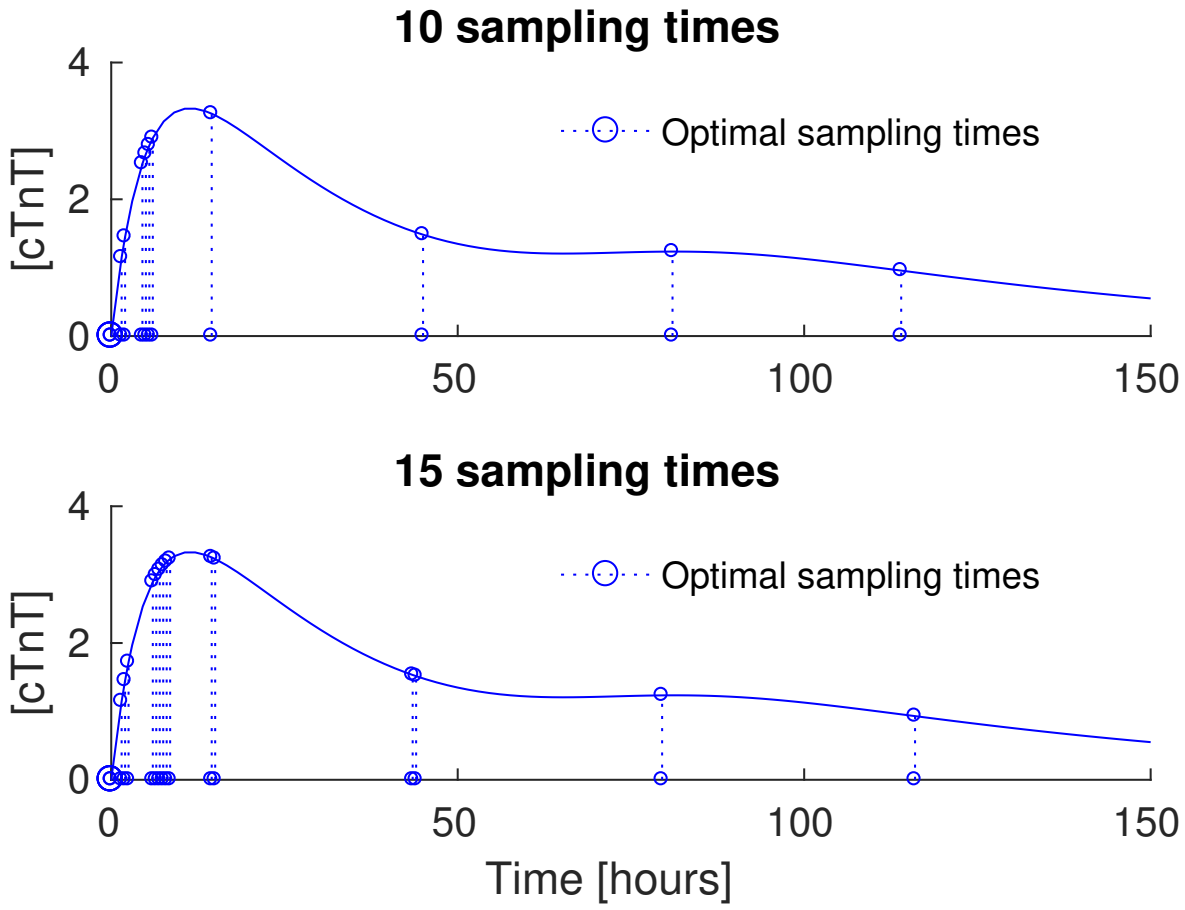


Figure 11: *Optimal sampling times with Eopt for a generic patient, using 10 and 15 measurements.*