Formal Verification of Biomedical Devices via In Silico Clinical Trials on Adversarial Scenarios

A Case Study on an Artificial Pancreas for T1DM Patients

Agostina Calabrese*

Computer Science Department, Sapienza University of Rome
calabrese.1657689@studenti.uniroma1.it

Abstract. Biomedical Devices improve the quality of life in patients by making the treatments they follow completely, or partially, automated. However, when the effects of a biomedical device are relevant on health, the consequences due to a possible malfunction might be critical. As a consequence, design of biomedical devices is often a long and expensive process and requires a verification of the device in each of the possible relevant scenarios. When performing an in vivo clinical trial, the set of involved patients is often small and the devices can be tested only in the scenarios that actually occur. As a consequence, performing the verification of a biomedical device by means of an in vivo clinical trial is not feasible.
In this paper we show a technology for performing In Silico Clinical Trials (ISCTs) of biomedical devices. As a case study, we describe the results concerning the preliminary phase of an ISCT of the Medtronic MiniMed ePID System [29], an artificial pancreas for Type 1 Diabetes Mellitus (T1DM) patients.

Keywords: In Silico Clinical Trials · Simulation-based Verification · Cyber-Physical Systems · VPH models · Model Checking · Simulation.

1 Introduction

The design of new biomedical devices is registering a positive trend due to the advance of biomedical engineering. Such devices are meant to improve the quality of life in different kinds of patients by making the treatments they follow completely, or partially, automated.

As one would expect, the more the effects of a biomedical device are relevant on health, the more the consequences due to a possible malfunction are critical. For instance, the artificial pancreas (see, *e.g.*, [31]) is a safety-critical device for blood glucose levels monitoring and regulation in patients with Type 1 Diabetes Mellitus (T1DM). If not correctly designed, the artificial pancreas has the capability to lead a patient to coma or worst, to death. As a consequence, design of biomedical devices is often a long and expensive process.

* Alternative email address: agostina.calabrese3@gmail.com

1.1 Motivation

Biomedical devices are often composed of one or more physical components controlled by software. This class of systems is known as Cyber Physical Systems (CPSs). For instance, an artificial pancreas includes glucose sensors and insulin and/or glucagon pumps (the actuators), both interacting with a control algorithm. In order to verify that the behaviour of a CPS meets the specification, we would need to verify it in each of the possible relevant scenarios. A scenario can be defined as a finite sequence of either ordinary or anomalous events. For instance, in the field of artificial pancreases validation, an ordinary event could be the occurrence of a meal, while an anomalous event could be a sudden obstruction of the insulin pump.

Despite the improvements in sensor and pump design and realisation, the artificial pancreas must counter delays and inaccuracies in both glucose measurement and insulin delivery [4]. For instance, Continuous Glucose Monitoring (CGM) devices measure glucose levels in the interstitial fluid, but there is a physiological (and sensor-indipendent) delay representing the transport of glucose from blood to interstitial fluid that must be taken into account [13].

Even more important than delays are the potential deviation between the sensed and the actual glucose levels and the possible difference between the amount of administered insulin and the computed dose. The occurrence of these events can be modelled by variations in the parameter values of the model defining the System Under Verification (SUV).

In the case of a biomedical device, the verification activity should be repeated for each patient taken from a possibly complete population. When performing an in vivo clinical trial, the set of involved patients is often small and the devices can be tested only in the scenarios that actually occur. This means that, if no obstructions occurs during the in vivo clinical trial, nothing can be argued about what the behaviour of the artificial pancreas would be in the case of such realistic anomalies. As a consequence, performing the verification of a biomedical device by means of an in vivo clinical trial is a very time consuming and expensive process which requires the recruitment of many volunteers for a long period of time.

These objections do not apply in the case of in In Silico Clinical Trials (ISCTs). An ISCT is a clinical trial performed by means of computer simulations over a population of Virtual Patients (VPs) (see, *e.g.*, http://paeon.di.uniroma1.it), and can greatly help in the early phases of the design of a new biomedical device in order to spot design errors or fragile design choices.

Being entirely *model-based*, performing an ISCT is much cheaper and faster than an in vivo trial, requiring only a mathematical model of both the physical and the cyber components of the device to be used, in synergy with a model of the patient (Virtual Physiological Human, VPH, model) and a model of the Pharmaco-Kinetics/-Dynamics (PKPD) of the relevant medicinal drugs (see, *e.g.*, [7]). Such heterogeneous models need to be integrated in order to be simulated as a *closed loop*.

1.2 Contribution

In this work we show a technology to perform ISCTs of biomedical devices, by focussing, as a case study, on the preliminary phase of an ISCT of an artificial pancreas for patients affected by T1DM.

Our technology is based on Modelica, one of the major open-standard general-purpose languages for modelling dynamical systems, widely used in application domains as diverse as mechanical, electrical, electronic, hydraulic, thermal, control, electric engineering, but also physiology and pharmacology (see, *e.g.*, [26]). Translators are also available to integrate biochemical models in Systems Biology Markup Language (SBML) into Modelica (see, *e.g.*, [17]). Several efficient and highly-configurable Modelica-based simulators are currently available, both open-source (*e.g.*, OpenModelica and JModelica) and proprietary (*e.g.*, Dymola).

In our case study, we defined in Modelica the Medtronic MiniMed ePID System described in [29]. The ePID system uses a Proportional–Integral–Derivative (PID) controller, and hence is purely reactive and respond to alterations in blood glucose levels only after they have occurred. Because of this, PID algorithms must cope with the time lags in both glucose sensing and insulin action and delivery [4].

Therefore, defining an *adversarial model* of the uncontrollable events that may occur and impact the correct functioning of the ePID system (*disturbance model*) is of fundamental importance in order to perform a reliable System Level Formal Verification (SLFV) of the biomedical device.

We defined such a disturbance model (again in Modelica) in terms of possible temporary faults in the sensors and actuators of the device (a time series of such events defines an *operational scenario*), and we used the System Level Formal Verifier (SyLVer) tool [18,22] developed by the Model Checking Laboratory (MCLab) (http://mclab.di.uniroma1.it/site) to generate an optimized simulation campaign that verifies the closed-loop artificial pancreas–virtual patient system on *all* such scenarios.

In this work we show an extension of the SyLVer approach where the monitor functionalities are no more limited to a PASS/FAIL decision. In our extension, the monitor is used to compute the values of application-dependent Key Performance Indicators (KPIs), allowing statistical analysis of results and thus giving back to the designers both counter-examples (*i.e.*, scenarios where the device performance are unsatisfactory and might pose the patient safety at risk) as well as *aggregate/statistical information* on the overall device performance and robustness.

1.3 Paper Outline

The paper is organised as follows. Section 2 describes the T1DM VP population involved in the ISCT, while Section 3 is dedicated to the description of the model of the biomedical device. The disturbance model and the generation of the simulation campaigns are shown in Section 4. Finally, the results of the preliminary phase of our ISCT are discussed in Section 6, while conclusions are drawn in Section 7.

2 T1DM Virtual Patient Population

The starting point to carry out an ISCT for the verification of a biomedical device is the availability of a suitable population of VPs. Such a population must be *complete*, *i.e.*, large enough to represent all relevant human patient phenotypes, whose spectrum can be quite large in hormonal regulatory systems, since they typically occur within a complex network of endocrinological, neurological, and psychological factors (see, *e.g.*, [10,16,15]).

To compute such a population of virtual patients, we need a VPH model, *i.e.*, a mathematical model of the (patho-)physiology of interest and the kinetics-dynamics of relevant drugs. Often, VPH models are in the form of parametric systems of Ordinary Differential Equations (ODEs), where parameters are used to model inter-subject variability, meaning that different assignments determine the behaviour of different patients.

Unfortunately, as argued in, *e.g.*, [30,25,19], many of the possible assignments to the parameters of a VPH model lead to time evolutions that are not biologically admissible (*i.e.*, coherent with the laws of biology).

As a consequence, a representative population of virtual patients cannot be built by arbitrarily picking assignments to the parameters of a VPH model, but an intelligent search in the parameter space is needed.

The work in [3] describes the computation of a representative population of T1DM VPs, obtained by exploiting the Medtronic VPH model of the human glucose regulation system [14]. This model is simpler (hence, faster to simulate) than other models, *e.g.*, those in [8,5], but is similarly effective in predicting the evolution of blood glucose and plasma insulin concentrations.

The population of VPs has been generated by using the VP generator originally presented in [30,25], which performs an AI-guided randomised search in the space of the model parameters. It is important to note that, in most cases, the size of the parameter space is such that, even after proper discretisation, an exhaustive search would be infeasible. To counteract this issue, our VP generator exploits statistical hypothesis rejection methods (see, *e.g.*, [9,24,23]).

3 The Medtronic MiniMed ePID System

An artificial pancreas is a CPS consisting of sensors, a control algorithm, and actuators. Typically, a CGM sensor gains information about current glucose blood level. The collected information feeds the control algorithm which computes the amount of drug to be injected into the patient. The actuators of artificial pancreases are hormonal pumps. The most common devices include only an insulin pump, but recent research is working forward bi-hormonal controllers for blood glucose regulation [8,11] having an additional pump for glucagon administration.

In our case study we verified the Medtronic MiniMed external physiological insulin delivery (ePID) system [29]. This closed-loop controller for glucose regulation is composed of a CGM sensor, a PID controller and an insulin pump. The

control algorithm is described by the following equations:

$$P(t) = K_p[SG(t) - \text{target}] \tag{1}$$

$$I(t) = I(t-1) + \frac{K_p}{T_I}[SG(t) - \text{target}] \tag{2}$$

$$D(t) = K_p \cdot T_D \cdot \dot{SG}(t) \tag{3}$$

where $SG(t)$ is the measured blood glucose concentration at time t, and *target* is the target glucose level. The insulin dose that the pump has to administrate at time t is given by the equation:

$$PID(t) = P(t) + I(t) + D(t) \tag{4}$$

This model includes the following 3 parameters:

- K_p ($\mu U/min^2$) is a factor depending on the subject's daily dose of insulin,
- T_I (min) is a parameter used to allow small changes in the integral compartment during the day and rapid changes during the night,
- T_D (min) is a factor used to regulate the insulin dose according to glucose rising and falling.

The first parameter is patient-specific, but its value is uniquely determined by the daily insulin dose. The remaining two parameters are set to the same value for all patients. Since CGM devices measure glucose levels in the interstitial fluid, the model of the sensor "reads" the blood glucose concentration value from the VPH model and adds the time lag as in the following equation:

$$\dot{G_{\text{ISF}}}(t) = -\frac{1}{\tau_{\text{SEN}}} \cdot G_{\text{ISF}}(t) + \frac{1}{\tau_{\text{SEN}}} \cdot (G(t) + \text{error}(t)) \tag{5}$$

where G is the blood glucose concentration and τ_{SEN} is the interstitial fluid delay (min).

4 Adversarial Operational Scenarios

While falsification approaches (see, *e.g.*, [1,6,2,28]) are incomplete approaches aiming at finding errors in the SUV, SLFV aims at certifying the absence of errors by verifying the SUV on *all* the simulation scenarios that are considered relevant.

Indeed, in our SLFV activity of the artificial pancreas described in Section 3, we need to generate an *exhaustive* simulation campaign *i.e.*, a simulation campaign that includes all simulation scenarios deemed relevant. Usually it requires weeks or even months of simulation activity to perform an exhaustive campaign, and the prospect is even worse if considering that when the SUV is a biomedical device, the simulation campaign should be repeated for more than one patient (the complete population of patients, if possible).

SyLVer [18,22] is a tool for the generation of optimized simulation campaigns starting from a model of the operational environment of the SUV (namely, a disturbance model). Such simulation campaigns are *exhaustive*, in that they exercise the SUV on *all* scenarios entailed by the operational environment model. However, by suitable randomising the verification order of the operational scenarios to be simulated [20], the simulation campaign computed by SyLVer is also *anytime*, in that during the verification process the system outputs an upper bound to the Omission Probability [21], *i.e.*, the probability that an error will be found during the simulation of a yet-to-be-simulated scenario. This feature allows the verification engineer to stop the (otherwise exhaustive) verification process when the omission probability goes below a given threshold. The simulation campaign computed by SyLVer is also *parallel*, in that it is designed to be executed on possibly large high-performance computing infrastructures [22].

Our Modelica definition of the PID controller allows the injection of temporary faults into the glucose-sensing mechanism and the insulin-delivery mechanism. What we need to do is to formally describe how often an event able to affect at least one of these two errors can happen, and in which measure it can contribute to the errors. By doing this we defined all the operational scenarios that are relevant to the SLFV of the artificial pancreas. The goal is to describe all the admissible sequences of events (*i.e.*, a scenario) by means of a Finite State Automaton (FSA), in order to give it as input to the SyLVer tool. To this end, we modelled the disturbance sequences characterizing the operational environment of the artificial pancreas, again using Modelica. The FSA is then automatically generated starting from this high-level description.

We equipped our Modelica model of the biomedical device with a general module for the application of disturbances on a signal. This module can be seen as a function that takes as input the original signal and returns the disturbed signal according to the following equation:

$$f(S(t)) = \alpha \cdot S(t - \tau) + \beta \tag{6}$$

where $S(t)$ is the signal and α, β and τ are three parameters. In this way, we can instantiate the equation above by assigning different parameter values for, respectively, glucose-sensing error and insulin-delivery error. Since a calibration error in glucose-sensing can be modelled as an additive error, we fixed α to 1 and τ to 0. In order to define only realistic scenarios, the calibration error in glucose-sensing should be constant through the time and restricted to a small domain centered in 0. We discretised this range and defined the domain of β as the set $\{-5, 0, 5\}$. The values in the set are expressed in mg/dl. The value of β, initialized at 0, is chosen one hour after the start of the simulation and never changed through the scenario. We chose to not inject disturbances during the first hour of the in-silico clinical trial (*i.e.*, the SLFV of the artificial pancreas on the virtual T1DM patients) in order to let the system reach stability. We modelled the possible errors in the insulin delivery mechanism by defining the domain of α as the set $\{0.8, 1, 1.2\}$. In order to simulate the occurrence of sudden failures, *e.g.*, a partial obstruction of the pump, the value of α, initialized at 1,

can be modified every 6 hours starting from the end of the first hour of the trial. Since it is more natural to model the effect of an obstruction event as a proportional error, we fixed both β and τ to 0.

5 Monitor

The last ingredient to perform the SLFV activity is a criterion to evaluate the behaviour of the artificial pancreas. In order to best fit the requirements of in-silico clinical trials, we extended the SyLVer approach by defining a monitor for the SUV that returns the values of the KPIs of interest instead of the boolean PASS/FAIL. This is done in order to allow statistical analysis of results. To this purpose, we defined the following KPIs:

– the average of function GRADE during time. GRADE is a function introduced in [12] in order to provide a method to evaluate the degree of dangerousness of blood glucose levels. The GRADE function assigns to glucose concentrations (expressed in mg/dl) a score from the interval $[0, 50]$ (see Figure 1) and it is defined as:

$$\text{GRADE}(g) = \begin{cases} 425 \cdot \{log_{10}[log_{10}(\frac{x}{18})] + 0.16\}^2 & \text{if } g \in [37, 630] \\ 50 & \text{otherwise} \end{cases}$$

Accordingly to its definition, the GRADE function assigns a score ≤ 5 if and only if the corresponding blood glucose level is within the euglycemic range (*i.e.*, 70-140 mg/dl), while high scores are assigned in case of both hypoglycemia and hyperglycemia. This KPI can thus be calculated as:

$$\overline{\text{GRADE}} = \frac{\int_0^h \text{GRADE}(t)dt}{h} \tag{7}$$

where h is the horizon of the simulation.

– the mean deviation from target (see (1) and (2)). In [29] the target glucose concentration was 120 mg/dl for safety reasons, but the authors themselves argued that better results could be probably achieved by setting the target to a lower value. In the case of in-silico clinical trials the limitations due to safety reasons do not apply, so we fixed the target at 105 mg/dl (*i.e.*, the center of the euglycemic range). This KPI is calculated as:

$$\overline{\text{targetDev}} = \frac{\int_0^h \frac{|G(t) - \text{target}|}{\text{target}} dt}{h} \tag{8}$$

– the highest glucose level registered during the trial

$$\max_G = \max_{0 \leq t \leq h} G(t) \tag{9}$$

– the lowest glucose level registered during the trial

$$\min_G = \min_{0 \leq t \leq h} G(t) \tag{10}$$

– the fraction of time that the patient spent in the euglycemin range, calculated as:

$$\text{isEuglycemic}(t) = \begin{cases} 1 & \text{if GRADE}(t) < 5 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{safeTime} = \frac{\int_0^h \text{isEuglycemic}(t)dt}{h} \tag{11}$$

We defined the PASS/FAIL test as in the following:

$$(\overline{\text{GRADE}} \leq 20 \wedge \max_G \leq 300 \wedge \min_G \geq 50 \wedge \text{safeTime} \geq 55.00\%) \iff \text{PASS}$$

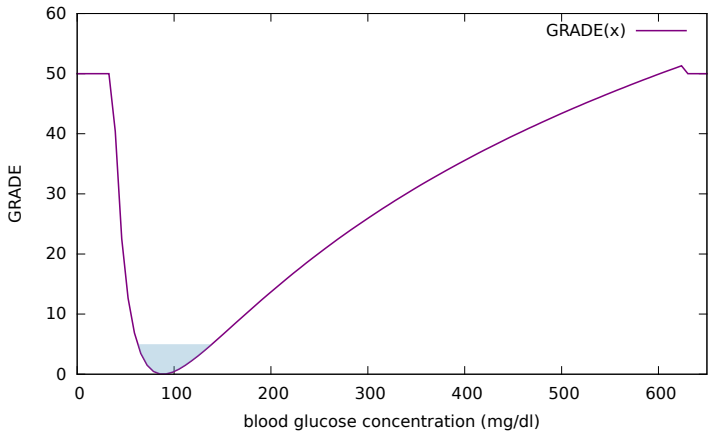


Fig. 1: The GRADE function (the highlighted area shows the euglycemic range).

6 Preliminary Experimental Results

In this section we discuss the preliminary experimental results obtained from the ISCT on 40 representative VPs (*i.e.*, 4 times the number of patients included in the corresponding in vivo clinical trial [29]). In order to let the SyLVer tool generate the optimised simulation campaigns we need to fix the horizon of the simulation scenarios. We decided to extend the verification period chosen in the in vivo clinical trial concerning the PID controller [29] (32 h) to 48 h, thus considering all the simulation scenarios having 48 h as horizon defined by our disturbance model (see Section 4). The optimized simulation campaign generated by SyLVer counts 65 769 different scenarios. For each scenario, we verified the device under three different inputs (*i.e.*, normal, hyperglycemic and

hypoglycemic condition). More specifically, we adopted the portfolio of inputs described in [3], hence altering the amount of ingested carbohydrates and the daily dose of injected insulin by the specified multiplication factors. As a result, the PID controller was validated in 197307 different scenarios for each of the 40 patients.

Despite the number of involved patients in our preliminary experiments is too small to compute meaningful statistics on the robustness of the Medtronic MiniMed external physiological insulin delivery (ePID) system, it was enough to detect *faulty behaviours* of the system. Figure 2 shows the percentage of successful scenarios registered by each VP. As shown in the diagram, almost half of the VPs registered a PASS in more than 90% of the scenarios. However, it is not negligible that the artificial pancreas failed in almost all the scenarios for 10% of the VPs.

In the following analysis we will focus on the VPs with, respectively, the lowest (p_l) and the highest (p_h) percentages of successful scenarios (besides VPs with a success rate of 100%). As shown in Figure 3a, p_l registered a FAIL in all the scenarios because of a too low minimum blood glucose concentration. This faulty behaviour is probably due to the high insulin sensitivity of this VP (*i.e.*, 0.002 ml/μU), showing that ISCTs can be of fundamental importance in determining personalised settings of the device. p_h has a failure rate of 2.70%, and the main cause is the constraint on the minimum blood glucose concentration (see Figure 3b). The failure of the system is due to scenarios involving an error in the insulin delivery component, showing that the artificial pancreas cannot cope with all the relevant adversarial scenarios.

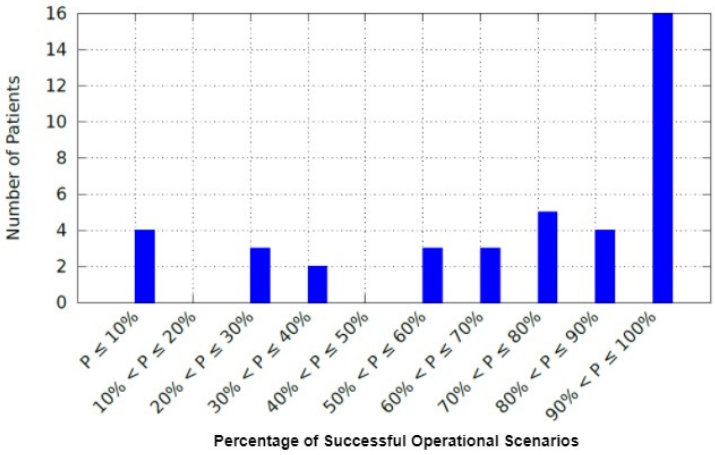
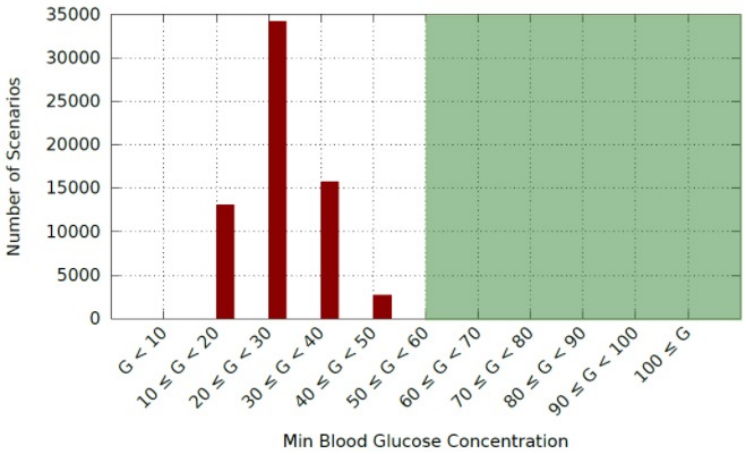
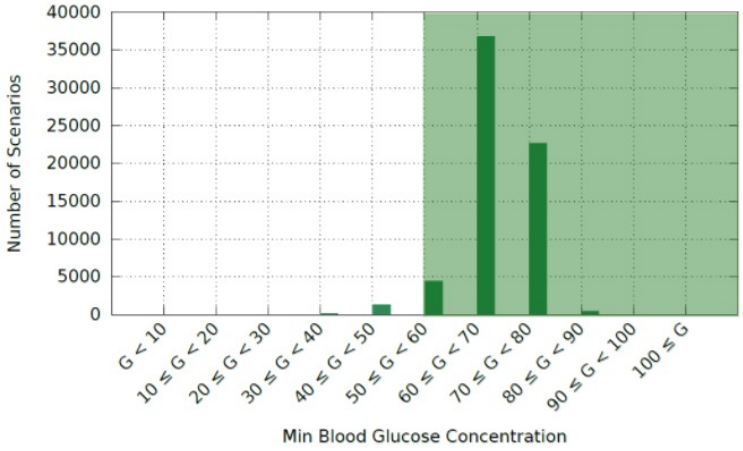


Fig. 2: Percentages of successful scenarios for each VP.



(a) VP with the lowest percentage of successful scenarios (p_l).



(b) VP with the highest percentage of successful scenarios (p_h) besides VPs with a success rate of 100%.

Fig. 3: Minimum blood glucose concentrations for each scenario.

7 Conclusions

In this work, we defined an *in silico* clinical trial for the system-level verification of the Medtronic MiniMed external physiological insulin delivery (ePID) system [29] on a representative population of VPs and on adversarial scenarios encompassing temporary faults in the glucose sensor and the insulin delivery mechanism of the biomedical device.

The entire workflow of our ISCT has been performed in Modelica. The population of VPs on which we carried out our verification activity has been computed in [3], while the generation of the adversarial operational scenarios has been performed starting from a high-level model given as input to SyLVer [18,22].

In order to provide the user with statistical information about the robustness of the biomedical device under verification, we extended the SyLVer approach by defining a monitor for the SUV computing the values of suitable KPIs. Our preliminary experiments highlighted a few scenarios resulting in faulty behaviours of the artificial pancreas. These failures were caused by the lack of personalised settings of the device and by errors in the insulin delivery mechanism.

In future work we plan to extend our verification activity by including in our adversarial operational scenario model unexpected patient behaviours in terms of carbohydrates intake and meal profiles, along the lines of, *e.g.*, [27].

Acknowledgements

This work was partially supported by the following research projects/grants: Italian Ministry of University & Research (MIUR) grant "Dipartimenti di Eccellenza 2018–2022" (Dept. Computer Science, Sapienza Univ. of Rome); EC FP7 project PAEON (Model Driven Computation of Treatments for Infertility Related Endocrinological Diseases, 600773); INdAM "GNCS Project 2019". The experimental part has been run on the Marconi CINECA cluster, thanks to Class C ISCRA Project n. HP10COBFWG.

The author is grateful to Stefano Sinisi (Dept. Computer Science, Sapienza Univ. of Rome) for having carefully supervised this work.

References

1. H. Abbas, G. Fainekos, S. Sankaranarayanan, F. Ivančić, and A. Gupta. Probabilistic temporal logic falsification of cyber-physical systems. *ACM TECS*, 12(2s), 2013.
2. A. Adimoolam, T. Dang, A. Donzé, J. Kapinski, and X. Jin. Classification and coverage-based falsification for embedded control systems. In *CAV 2017*, volume 10426 of *LNCS*. Springer, 2017.
3. A. Calabrese, T. Mancini, A. Massini, S. Sinisi, and E. Tronci. Generating T1DM virtual patients for in silico clinical trials via AI-guided statistical model checking. In *RCRA 2019*, 2019.
4. C. Cobelli, E. Renard, and B. Kovatchev. Artificial pancreas: Past, present, future. *Diabetes*, 60, 2011.

5. C. Dalla Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli. The UVA/Padova type 1 diabetes simulator: New features. *JDST*, 8, 2014.
6. A. Dokhanchi, A. Zutshi, R. Sriniva, S. Sankaranarayanan, and G. Fainekos. Requirements driven falsification with coverage metrics. In *EMSOFT 2015*. IEEE, 2015.
7. T. Eissing, L. Kuepfer, C. Becker, M. Block, K. Coboeken, T. Gaub, L. Goerlitz, J. Jaeger, R. Loosen, B. Ludewig, M. Meyer, C. Niederalt, M. Sevestre, H. Siegmund, J. Solodenko, K. Thelen, U. Telle, W. Weiss, T. Wendl, S. Willmann, and J. Lippert. A computational systems biology software platform for multiscale modeling and simulation: Integrating whole-body physiology, disease biology, and molecular reaction networks. *Front. Physiol*, 2, 2011.
8. F. El-Khatib, S. Russell, D. Nathan, R. Sutherlin, and E. Damiano. Bi-hormonal closed-loop blood glucose control for type 1 diabetes. *Sc. Transl. Med.*, 2, 2010.
9. R. Grosu and S. Smolka. Monte Carlo model checking. In *TACAS 2005*, volume 3440 of *LNCS*. Springer, 2005.
10. M. Hengartner, T. Kruger, K. Geraedts, E. Tronci, T. Mancini, F. Ille, M. Egli, S. Roeblitz, R. Ehrig, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, and B. Leeners. Negative affect is unrelated to fluctuations in hormone levels across the menstrual cycle: Evidence from a multisite observational study across two successive cycles. *J. Psycho. Res.*, 99, 2017.
11. P. Herrero, P. Georgiou, N. Oliver, M. Reddy, D. Johnston, and C. Toumazou. A composite model of glucagon–glucose dynamics for in silico testing of bihormonal glucose controllers. *JDST*, 7, 2013.
12. N. Hill, P. Hindmarsh, R. Stevens, I. Stratton, J. Levy, and D. Matthews. A method for assessing quality of control from glucose profiles. *Diabetic Medicine*, 24, 2007.
13. R. Hovorka. Closed-loop insulin delivery: from bench to clinical practice. *Nat. Rev. End.*, 7, 2011.
14. S. Kanderian, S. Weinzimer, G. Voskanyan, and G. Steil. Identification of intraday metabolic profiles during closed-loop glucose control in individuals with type 1 diabetes. *JDST*, 3, 2009.
15. B. Leeners, T. Krüger, K. Geraedts, E. Tronci, T. Mancini, M. Egli, S. Röblitz, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, and F. Ille. Associations between natural physiological and supraphysiological estradiol levels and stress perception. *Front. Psycol.*, 10, 2019.
16. B. Leeners, T. Kruger, K. Geraedts, E. Tronci, T. Mancini, F. Ille, M. Egli, S. Roeblitz, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, and M. Hengartner. Lack of associations between female hormone levels and visuospatial working memory, divided attention and cognitive bias across two consecutive menstrual cycles. *Front. Behav. Neuro.*, 11, 2017.
17. F. Maggioli, T. Mancini, and E. Tronci. SBML2Modelica: Integrating biochemical models within open-standard simulation ecosystems. *Bioinformatics*, 2019.
18. T. Mancini, F. Mari, A. Massini, I. Melatti, F. Merli, and E. Tronci. System level formal verification via model checking driven simulation. In *CAV 2013*, volume 8044 of *LNCS*. Springer, 2013.
19. T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, S. Sinisi, E. Tronci, R. Ehrig, S. Röblitz, and B. Leeners. Computing personalised treatments through in silico clinical trials. A case study on downregulation in assisted reproduction. In *RCRA 2018*, 2018.
20. T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, and E. Tronci. On minimising the maximum expected verification time. *Inf. Proc. Lett.*, 122, 2017.

21. T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. Anytime system level verification via parallel random exhaustive hardware in the loop simulation. *Microprocessors and Microsystems*, 41, 2016.
22. T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. SyLVaaS: System level formal verification as a service. *Fundam. Inform.*, 1–2, 2016.
23. T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, and L. Elmegaard. Parallel statistical model checking for safety verification in smart grids. In *SmartGridComm 2018*. IEEE, 2018.
24. T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, M. Prodanovic, and L. Elmegaard. User flexibility aware price policy synthesis for smart grids. In *DSD 2015*. IEEE, 2015.
25. T. Mancini, E. Tronci, I. Salvo, F. Mari, A. Massini, and I. Melatti. Computing biological model parameters by parallel statistical model checking. In *IWBBIO 2015*, volume 9044 of *LNCS*. Springer, 2015.
26. M. Mateják and J. Kofránek. Physiomodel – An integrative physiology in Modelica. In *EMBC 2015*. IEEE, 2015.
27. N. Paoletti, K. Liu, H. Chen, S. Smolka, and S. Lin. Data-driven robust control for a closed-loop artificial pancreas. *IEEE/ACM Trans. Comp. Biol. Bioinf.*, 2019. To appear.
28. S. Sankaranarayanan, S. Kumar, F. Cameron, B. Bequette, G. Fainekos, and D. Maahs. Model-based falsification of an artificial pancreas control system. *ACM SIGBED Review*, 14(2), 2017.
29. G. Steil, K. Rebrin, C. Darwin, F. Hariri, and M. Saad. Feasibility of automating insulin delivery for the treatment of type 1 diabetes. *Diabetes*, 55, 2006.
30. E. Tronci, T. Mancini, I. Salvo, S. Sinisi, F. Mari, I. Melatti, A. Massini, F. Davi', T. Dierkes, R. Ehrig, S. Röblitz, B. Leeners, T. H. C. Krüger, M. Egli, and F. Ille. Patient-specific models from inter-patient biological models and clinical records. In *FMCAD 2014*. IEEE, 2014.
31. S. Weinzimer, G. Steil, K. Swan, J. Dziura, N. Kurtz, and W. Tamborlane. Fully automated closed-loop insulin delivery versus semiautomated hybrid control in pediatric patients with type 1 diabetes using an artificial pancreas. *Diab Care*, 31, 2008.