# In Silico Clinical Trials through AI and Statistical Model Checking

Vadim Alimguzhin<sup>1</sup>, Toni Mancini<sup>1</sup>, Annalisa Massini<sup>1</sup>, Stefano Sinisi<sup>1</sup>, and Enrico Tronci<sup>1</sup>

<sup>1</sup>Computer Science Department, Sapienza University of Rome {alimguzhin, tmancini, massini, sinisi, tronci}@di.uniroma1.it

#### Abstract

A Virtual Patient (VP) is a computational model accounting for individualised (patho-) physiology and Pharmaco-Kinetics/Dynamics of relevant drugs. Availability of VPs is among the enabling technology for In Silico Clinical Trials. Here we shortly outline the state of the art as for VP generation and summarise our recent work on Artificial Intelligence (AI) and Statistical Model Checking based generation of VPs.

#### 1 Introduction

Assessing safety and efficacy of pharmaceutical drugs and treatments is typically done through Randomised In Vivo Clinical Trials (RVCTs) which tests the candidate drug/treatment on suitably chosen patients. Unfortunately this approach has many drawbacks. First, it is very time consuming: in many cases, even more than 5 years are needed to have a new drug approved by the regulatory bodies. Second, it is very expensive: hundreds of million Euros are typically spent to complete a RVCT. Third, it is hardly usable for rare diseases, because of costs and lack of patients.

The above state of affairs motivates investigation of methods that can save on time and costs for safety and efficacy assessment of drugs and treatments. In such a scenario, the In Silico Clinical Trials (ISCTs) approach is one of the most promising ones. Shortly, much as done in simulation-based verification of Cyber-Physical Systems (CPSs), ISCTs replace the physical system (a patient in this case) with a computational model, a Virtual Patient (VP), accounting for the relevant features of the physical system. Basically, a VP is a computational model encompassing patient (patho-) physiology along with Pharmaco-Kinetics/Dynamics (PKPD) of relevant drugs. In such a context, safety and efficacy assessment can be done by simulating the effect of the drug on, ideally, all VPs, exactly as, ideally, CPS verification aims at evaluating requirements under all possible operational scenarios.

#### 2 State of the art

ISCTs basically entail the following steps. First, generate a set of VPs whose predictions are in agreement with (patho-) physiology, PKPD, and *in vivo* clinical data. This *model validation* activity is done once and for all. Second, use the above *in vivo* validated VPs to assess, through simulation, safety and efficacy of a candidate drug.



The main obstacle to overcome to carry out an ISCT is the generation of a cohort of VPs that is complete, i.e., is large enough to represent all relevant human patient phenotypes, and is minimal, i.e., does not contain behaviours that are not compatible with the human (patho-) physiology of interest. In the following we briefly review the state of the art on VPs generation methods.

First, we note that achieving simultaneously completeness and minimality is presently out of reach. Thus, usually, completeness is privileged. Basically, this aims at guaranteeing that all human patient phenotypes are adequately represented *in silico*.

Computational models for VPs are usually developed using medical knowledge from the literature and from pathway databases such as KEGG (Kyoto Encyclopedia of Genes and Genomes) or Reactome. Open formats, such as the Systems Biology Markup Language (SBML), allow exchange of computational models across platforms.

Unfortunately, physiology knowledge is often qualitative. To overcome such an obstacle, many qualitative as well as quantitative approaches have been devised. An overview is in [44].

Qualitative approaches, often referred to as *logic based*, discretise the values of interest. In this context, Boolean models have been widely studied (see, e.g., [18] and citations thereof). Logical models are very useful for a qualitative analysis, but are quite difficult to use within a compositional framework where quantitative models from physiology or pharmacology are present. Since in an ISCT quantitative aspects are extremely relevant (e.g., drug dosage) we will focus on quantitative models.

Quantitative models typically focus on representing dynamics of chemical reactions through Ordinary Differential Equations (ODEs) with stoichiometric parameters and reaction rates estimated from clinical data using model identification techniques, as, e.g., in [55]. Such models are typically used to support ISCT, as, e.g., in [26]. The main obstacle to overcome in using quantitative models is the lack of knowledge about the values for their parameters. In fact, very few model parameter values can be estimated using clinical data (see, e.g., [51] for a survey). As a result, most parameter values have to be estimated through computational methods, typically using model identification techniques (see, e.g., [57, 54]).

In such a setting two approaches are typically used: those based on *optimisation* techniques and those based on *statistical* techniques.

Both global and local optimisation-based approaches, relying on Artificial Intelligence (AI) as well as on numerical methods, have been widely studied. Global optimisation—based approaches are computationally heavy, however convergence is guaranteed to a global optimum (see, e.g., [27, 56, 42]). On the other hand, local optimisation approaches (e.g., [28, 34, 38]) are in general computationally lighter than global ones, but convergence is only guaranteed towards a local optimum (e.g., [20, 21, 40]). The reader is referred to, e.g., [43] for an in-depth comparison among the two approaches and [30] for an example of use of AI-based optimisation techniques for computing personalised treatments.

Typically, both global as well as local optimisation approaches rely on model identifiability, *i.e.*, different values for the model parameters lead to different model behaviours (see, *e.g.*, [25]). In such a case, different model identification techniques can be used. Examples are in [9, 45, 10, 53, 1, 49].

When clinical data are scarce, identification approaches can be applied by averaging data coming from different patients. However, the resulting parameter value yields an *inter-patient* model behaviour (see, e.g., [46]).

Since the goal of the above mentioned approaches is to generate a model parameter value that fits available experimental data, a huge amount of data per patient is needed in order to generate a virtual population that is representative of all human phenotypes. As a result, generating a complete set of VPs using model identification techniques would require considering a large amount of patients (ideally at least one for each relevant phenotype) along with a huge amount of clinical data for each of such patients. Unfortunately, this is exactly one of the obstacles ISCTs aims at overcoming. Thus, while model identification techniques are essential to validate models against clinical data and to develop *patient-specific* models, they can be hardly used to generate a complete set of VPs.

Moreover, VP models are often globally or partially unidentifiable, *i.e.*, wide ranges of parameter values lead to very similar model behaviours (see, *e.g.*, [8]). However, being able to generate VPs (*i.e.*, parameter values) which yield clearly distinguishable model behaviours is of crucial importance for ISCTs.

Statistical approaches are also widely used. In such approaches a parameter probability distribution, rather than a single value, is inferred (see, e.g., [17, 22, 48]). They are typically used for physiology-based PKPD models (see, e.g., [47]), i.e., quantitative VP models where parameter values are measurable physiological quantities (such as blood flow, organ volumes, etc). Unfortunately, most VP models have

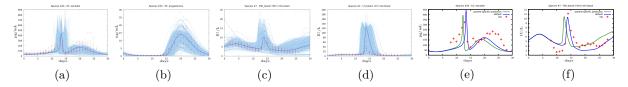


Figure 1: (a)–(d): system trajectories under *all* Biologically Admissible (BA) parameters (dark blue curves denote trajectories under default parameter); (e)–(f): patient-specific (green curves) vs. default predictions (blue curves) for, respectively, E2 and FSH.

hard-to-measure parameters such as stoichiometric constants and reaction rates. In fact, probability distribution functions for them are typically unknown.

Limited knowledge about VP model parameters calls for methods that handle models whose parameters are partially unknown. This guarantees completeness, possibly at the price of minimality. In other words, we use an over-approximation of the set of all *physiologically admissible* (*i.e.*, whose behaviour is plausible with respect to physiology) VPs.

Since model-checking techniques enable exploration of all possible behaviours of a model, it is not surprising that such approaches have been investigated in order to generate VPs for both qualitative as well as quantitative VP models. In particular, for qualitative models (e.g., Boolean models) the problem of finding model parameter values yielding a biological meaningful behaviour is reduced to the problem of finding stable states of the model, i.e., attractors. In such a setting, symbolic model checking is widely used (see, e.g., [41, 13, 2, 19]). There, physiological admissibility can be defined using a temporal logic, such as CTL or LTL. Some examples are in [7, 6, 5, 11, 4] or in [15], where a probabilistic model checking approach is used.

Unfortunately, even using model approximation techniques like those in [3, 39], the above mentioned model checking approaches cannot be used to generate VPs from quantitative ODE-based models, *i.e.*, those sought to support ISCTs. Indeed, in such a setting only simulation-based approaches are viable. This suggests to investigate use of Statistical Model Checking (SMC) techniques to support VP generation.

## 3 Generating a complete set of Virtual Patients

Here we outline our AI and SMC based VP generation algorithms described in [52, 37]. Our algorithms take as input an ODE model for a biological system along with *default values* for all model parameters. Such default values are typically obtained by *averaging* among the behaviour of many patients (*inter-patient* model), as, *e.g.*, in [46]. Our approach can be summarised as follows.

Formalisation of biological admissibility In order to build a general-purpose tool that can automatically search through millions of model parameter values we need a criterion to automatically filter out (most of) the parameter values leading to time evolutions that are not biologically meaningful. We provide such a criterion by defining, as BA parameter values, those entailing time evolutions having a second-order statistics  $close\ enough$  to that of the model default parameter values.

Computation of a complete set of Virtual Patients Our goal is to compute a set of BA values for the model parameters (VPs) that encompasses as many biologically meaningful behaviours as possible, but at the same time is not too large, in order to speed up our on-line computation. Thus, taking into account that differences in values below a certain threshold are meaningless from a biological point of view, we discretise the range of values for each model parameter. In such a framework, we present a SMC based algorithm that computes a set S containing only and (with arbitrarily high confidence) all BA values for our model parameters. Note that such an algorithm does not depend on patient-specific data. Thus it can be run once and for all off-line, and its output can be stored for further processing.

Basically, our algorithm makes an AI-guided sampling of the space of the model parameters (building on our Model Checking (MC) tools in, e.g., [31, 14, 32, 33, 36, 29]) and then uses hypothesis testing—based

SMC (as in [12, 35]) to guarantee, with any user-provided statistical confidence, completeness of the set of VPs generated.

Effectiveness of our approach has been evaluated on GynCycle, a model of the hormones regulating the human female menstrual cycle introduced in [46]. As hormonal regulatory systems occur within a complex network of endocrinological, neurological, and psychological factors [16, 24, 23], they are difficult to capture within clinical studies, and model-based approaches might be of great aid in taking these many factors under better control. GynCycle has 114 parameters, 75 of which are patient-specific (at least for our purposes), and consists of 41 algebraic-differential equations defining the time evolution of 33 species.

We experimentally evaluate soundness and completeness of our notion of biological admissibility, using reference values from the literature (e.g., [50]) and a dataset (courtesy of Pfizer), encompassing daily measurements of the blood level of E2, P4, FSH, and LH on 12 women during an entire menstrual cycle (totalling more than 1000 data points). Figures 1a to 1d show the trajectories for those 4 hormones obtained by running the GynCycle model on all parameter values computed by our algorithm. We see that most of such trajectories are biologically meaningful, being in agreement with the trajectories in [50]. This shows (experimentally) soundness of our biological admissibility notion. Furthermore (experimental completeness), most of the measurements in our dataset (red crosses in Figures 1a to 1d) lie within the region covered by our trajectories.

Figures 1e and 1f give an example of patient-specific (i.e., using our computed model parameters) predictions for, respectively, E2 and FSH and compare them with default predictions (i.e., using the default model parameters) and in vivo clinical measurements. The achieved error reduction is of about 10%. This value has a relevant impact from a clinical standpoint, as it can move hormone peaks (which are among the main fertility/infertility indicators) by several days (see Figures 1a to 1d).

### 4 Conclusions

Computation of a complete set of VPs is among the enabling technology for ISCTs. We have shown how using synergies between AI and SMC techniques it is possible to effectively generate a complete set of VPs starting from an ODE model and default values to the model parameters.

**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".

### References

- [1] U. Abdulla and R. Poteau. Identification of parameters in systems biology. Math. Biosci., 305:133-145, 2018.
- [2] W. Abou-Jaoudé, P. Monteiro, A. Naldi, M. Grandclaudon, V. Soumelis, C. Chaouiya, and D. Thieffry. Model checking to assess t-helper cell plasticity. Front. Bioeng. Biotechnol, 2, 2015.
- [3] V. Alimguzhin, F. Mari, I. Melatti, I. Salvo, and E. Tronci. Linearizing discrete-time hybrid systems. *IEEE TAC*, 62(10), 2017.
- [4] G. Arellano, J. Argil, E. Azpeitia, M. Benítez, M. Carrillo, P. Góngora, D. Rosenblueth, and E. Alvarez-Buylla. Antelope: A hybrid-logic model checker for branching-time boolean grn analysis. *BMC Bioinf.*, 12(1), 2011.
- [5] J. Barnat, L. Brim, D. Šafránek, and M. Vejnár. Parameter scanning by parallel model checking with applications in systems biology. In *PDMC-HIBI 2010*. IEEE, 2010.
- [6] G. Batt, D. Ropers, H. De Jong, J. Geiselmann, R. Mateescu, M. Page, and D. Schneider. Validation of qualitative models of genetic regulatory networks by model checking: analysis of the nutritional stress response in escherichia coli. *Bioinformatics*, 21(suppl\_1), 2005.
- [7] N. Chabrier and F. Fages. Symbolic model checking of biochemical networks. In CMSB 2003, vol. 2602 of LNCS. Springer, 2003.
- [8] O.-T. Chis, J. Banga, and E. Balsa-Canto. Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE*, 6(11), 2011.
- [9] P. Docherty, J. Chase, and T. David. Characterisation of the iterative integral parameter identification method. *Med. Biol. Eng. Comput.*, 50(2):127–134, 2012.

- [10] J. Garcia-Tirado, C. Zuluaga-Bedoya, and M. Breton. Identifiability analysis of three control-oriented models for use in artificial pancreas systems. JDST, 12(5):937–952, 2018.
- [11] H. Gong, P. Zuliani, Q. Wang, and E. Clarke. Formal analysis for logical models of pancreatic cancer. In CDC 2011. IEEE, 2011.
- [12] R. Grosu and S. Smolka. Monte Carlo model checking. In TACAS 2005, vol. 3440 of LNCS. Springer, 2005.
- [13] W. Guo, G. Yang, W. Wu, L. He, and M. Sun. A parallel attractor finding algorithm based on boolean satisfiability for genetic regulatory networks. Plos One, 9(4):e94258, 2014.
- [14] B. Hayes, I. Melatti, T. Mancini, M. Prodanovic, and E. Tronci. Residential demand management using individualised demand aware price policies. *IEEE Trans. Smart Grid*, 8(3), 2017.
- [15] J. Heath, M. Kwiatkowska, G. Norman, D. Parker, and O. Tymchyshyn. Probabilistic model checking of complex biological pathways. TCS, 391(3), 2008.
- [16] 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.
- [17] C. F. Higham and D. Husmeier. A bayesian approach for parameter estimation in the extended clock gene circuit of arabidopsis thaliana. BMC Bioinf., 14(S-10), 2013.
- [18] I. e. a. Irurzun-Arana. Advanced boolean modeling of biological networks applied to systems pharmacology. Bioinf., 2017.
- [19] S. Ito, T. Ichinose, M. Shimakawa, N. Izumi, S. Hagihara, and N. Yonezaki. Qualitative analysis of gene regulatory networks by temporal logic. *Theor. Comput. Sci.*, 594:151–179, 2015.
- [20] K. A. Kim, S. L. Spencer, J. G. Albeck, J. M. Burke, P. K. Sorger, S. Gaudet, and D. H. Kim. Systematic calibration of a cell signaling network model. *BMC Bioinf.*, 11, 2010.
- [21] A. Kramer, B. Calderhead, and N. Radde. Hamiltonian monte carlo methods for efficient parameter estimation in steady state dynamical systems. *BMC Bioinf.*, 15, 2014.
- [22] M. Krauss, R. Burghaus, J. Lippert, M. Niemi, P. Neuvonen, A. Schuppert, S. Willmann, L. Kuepfer, and L. Görlitz. Using bayesian-pbpk modeling for assessment of inter-individual variability and subgroup stratification. *In Silico Pharmacol.*, 1(1), 2013.
- [23] 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.
- [24] 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.
- [25] L. Ljung. System identification: theory for the user. Prentice-Hall, 1987.
- [26] J. e. a. Lu. An in-silico model of lipoprotein metabolism and kinetics for the evaluation of targets and biomarkers in the reverse cholesterol transport pathway. *PLoS Comp. Biol.*, 2014.
- [27] T. Mancini, M. Cadoli, D. Micaletto, and F. Patrizi. Evaluating ASP and commercial solvers on the CSPLib. Constraints, 13(4), 2008.
- [28] T. Mancini, P. Flener, and J. Pearson. Combinatorial problem solving over relational databases: View synthesis through constraint-based local search. In SAC 2012. ACM, 2012.
- [29] 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, vol. 8044 of LNCS. Springer, 2013.
- [30] 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.
- [31] 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.
- [32] 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.

- [33] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. SyLVaaS: System level formal verification as a service. Fundam. Inform., 1–2, 2016.
- [34] T. Mancini, F. Mari, I. Melatti, I. Salvo, and E. Tronci. An efficient algorithm for network vulnerability analysis under malicious attacks. In ISMIS 2018. Springer, 2018.
- [35] 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.
- [36] 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.
- [37] 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*, vol. 9044 of *LNCS*. Springer, 2015.
- [38] T. Mancini, E. Tronci, A. Scialanca, F. Lanciotti, A. Finzi, R. Guarneri, and S. Di Pompeo. Optimal fault-tolerant placement of relay nodes in a mission critical wireless network. In RCRA 2018, 2018.
- [39] F. Mari, I. Melatti, I. Salvo, and E. Tronci. Model based synthesis of control software from system level formal specifications. ACM TOSEM, 23(1), 2014.
- [40] A. Miró, C. Pozo, G. Guillén-Gosálbez, J. A. Egea, and L. Jiménez. Deterministic global optimization algorithm based on outer approximation for the parameter estimation of nonlinear dynamic biological systems. BMC Bioinf., 13, 2012.
- [41] N. Miskov-Zivanov, P. Zuliani, E. Clarke, and J. Faeder. Studies of biological networks with statistical model checking: Application to immune system cells. In *ACM-BCB 2013*. ACM, 2013.
- [42] R. Moss, T. Grosse, I. Marchant, N. Lassau, F. Gueyffier, and S. R. Thomas. Virtual patients and sensitivity analysis of the guyton model of blood pressure regulation: towards individualized models of whole-body physiology. *PLoS Comput. Biol.*, 8(6), 2012.
- [43] M. Nobile, A. Tangherloni, L. Rundo, S. Spolaor, D. Besozzi, G. Mauri, and P. Cazzaniga. Computational intelligence for parameter estimation of biochemical systems. In *IEEE CEC*, pages 1–8. IEEE, 2018.
- [44] N. L. Novre. Quantitative and logic modelling of molecular and gene networks. Nature Rev. Gen., 2015.
- [45] Y. Pantazis, M. A. Katsoulakis, and D. G. Vlachos. Parametric sensitivity analysis for biochemical reaction networks based on pathwise information theory. BMC Bioinf., 14, 2013.
- [46] S. Röblitz, C. Stötzel, P. Deuflhard, H. Jones, D.-O. Azulay, P. van der Graaf, and S. Martin. A mathematical model of the human menstrual cycle for the administration of GnRH analogues. J. Theor. Biol., 321, 2013.
- [47] S. Schaller, S. Willmann, J. Lippert, L. Schaupp, T. Pieber, A. Schuppert, and T. Eissing. A generic integrated physiologically based whole-body model of the glucose-insulin-glucagon regulatory system. *CPT: Pharmacometr. Sys. Pharmacol.*, 2(8), 2013.
- [48] E. Shockley, J. Vrugt, and C. Lopez. PyDREAM: high-dimensional parameter inference for biological models in Python. *Bioinf.*, 34(4):695–697, 2018.
- [49] P. Stapor, D. Weindl, B. Ballnus, S. Hug, C. Loos, A. Fiedler, S. Krause, S. Hroß, F. Fröhlich, and J. Hasenauer. PESTO: Parameter Estimation TOolbox. *Bioinf.*, 34(4):705–707, 2018.
- [50] R. Stricker, R. Eberhart, M. Chevailler, F. Quinn, P. Bischof, and R. Stricker. Establishment of detailed reference values for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the abbott architect analyzer. *CCLM*, 44(7), 2006.
- [51] J. Sun, J. Garibaldi, and C. Hodgman. Parameter estimation using metaheuristics in systems biology: A comprehensive review. *IEEE/ACM Trans. Comp. Biol. Bioinf.*, 9(1), 2012.
- [52] 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.
- [53] N. Tsiantis, E. Balsa-Canto, and J. Banga. Optimality and identification of dynamic models in systems biology: an inverse optimal control framework. *Bioinf.*, 34(14):2433–2440, 2018.
- [54] A. Villaverde and J. Banga. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. J. Royal Soc. Interface, 11(91), 2014.
- [55] A. e. a. Villaverde. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. J. R. Soc. Interf., 2013.
- [56] W. H. Wu, F. S. Wang, and M. S. Chang. Dynamic sensitivity analysis of biological systems. BMC Bioinf., 9(S-12), 2008.
- [57] C. Zhan and L. F. Yeung. Parameter estimation in systems biology models using spline approximation. BMC Sys. Biol., 5(1), 2011.