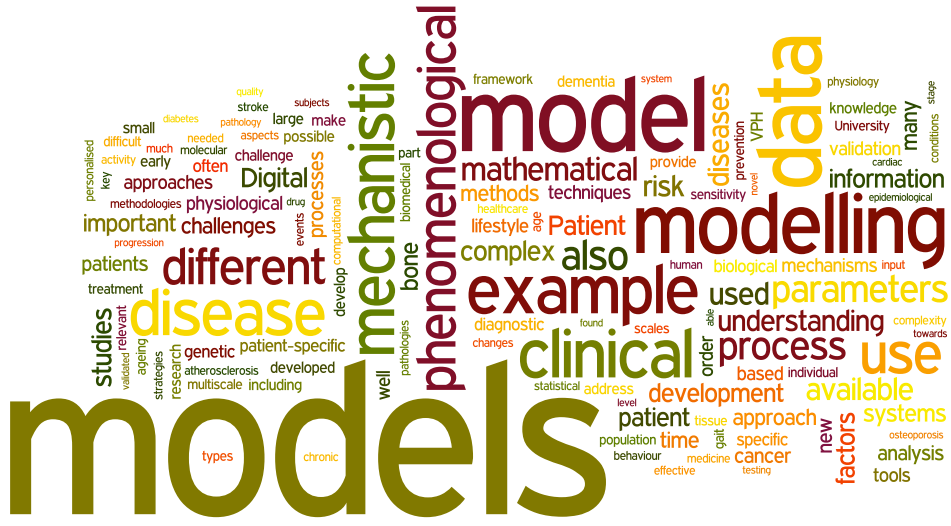
3 Mathematical modelling for the Digital Patient

Editorial team

Pietro Liò	University of Cambridge
Maria-Cruz Villa-Uriol	University of Sheffield
Vanessa Diaz	University College London
Liesbet Geris	University of Liège
Randy Thomas	IR4M UMR8081 CNRS and University Paris-Sud
Marcel Rutten	Technical University Eindhoven
Martijn Moné	VU University Amsterdam
Martin Nelson	University of Nottingham
Hans Westerhoff	University of Amsterdam, University of Manchester
William R. Taylor	ETH Zürich
Cesar Pichardo	InSCILICO, Ltd
Alfredo Hernandez	INSERM U-642, Université de Rennes 1
Paul Taylor	University College London

3.1 Overall introduction



Medicine has evolved rapidly in the last 25 years. Nowadays, it uses novel data gathering technologies to enable large-scale collection of observational data on complex biomedical systems, enabled by data-mining in order to exploit population studies, which are the basis of evidence-based medicine. . It is worth mentioning that there are recent technological breakthroughs on diagnostic and interventional imaging as well as a new generation of therapeutic devices; the first facilitates the creation of image-based models to further personalise patient models, while the second opens the possibility of exploring in parallel multiple treatments and their expected effects.

However, a new paradigm shift is about to occur: personalised medicine promises to revolutionise the practice of medicine, transform the global healthcare industry, and ultimately lead to longer and healthier lives. The Digital Patient is part of this new paradigm, grounded on the principle that it is possible to produce *predictive,* patient-specific mathematical models for *personalised* healthcare.

In many ways, modelling is able to complement data-mining, and the power of combining such approaches lies within leveraging their strengths for transforming observational data into knowledge. Building a model is a useful (and thrilling) scientific activity because it aims to reproduce the main features of a real system with the minimum number of parameters.

The objective is to gain a better understanding of how each of the different components contributes to the overall process. In this manner, modelling in medicine, fed by clinical research data, should aim to support clinical decision-making systems, and ultimately help the clinician to provide improved prevention and screening, diagnosis, prognosis, and/or a prediction of response to treatment.

The development of *robust* patient-specific models, as they become available, will significantly advance prevention and treatment of disease. Although models in healthcare can potentially have many uses, there is an increasing urgency to address the prevention of chronic diseases through lifestyle improvements as the best path to a healthier population. For example, chronic diseases are overwhelming western countries' healthcare systems. Between 70% and 80% of Europe's healthcare costs are spent on chronic care, amounting for some €700bn annually[86]. Chronic diseases account for over 86% of deaths in the EU. However, much of this disease burden is preventable, or may be delayed or diminished, through a combination of primary prevention measures, screening, and early intervention. While primary prevention focuses on healthy living, secondary prevention (early screening and diagnosis) and tertiary prevention (early intervention to slow the progression of diseases already identified) also play important roles in reducing the burden of chronic disease[87].

Ideally, we would have a full understanding of all biological processes in both health and disease, as well as of the relationships between structure and function at all scales, and this complete knowledge would be represented usefully in a collection of *inter-compatible* mathematical models at all relevant scales, *validated* against experimental and clinical data, *customizable* for individual patients, and *accessible* through a convenient and interactive user interface for use by health professionals and patients. With this resource, one would be able to simulate diseases and pathologies and to explore patient-specific *therapeutic intervention strategies*, *including evolution over time and realistic evaluation of prognosis*. This is, of course, the long-term vision of the Digital Patient.

In reality, our understanding at all scales is only partial, and available data is incomplete, as explained in other chapters of this document. Nonetheless,

(i) current knowledge of anatomy and physiology is extensive on many levels and much of it has already been successfully represented in mathematical models, and

(ii) relevant data is abundant, despite being incomplete and despite the many issues of standardisation, accessibility, and interoperability.

Given this state of affairs, this chapter focuses on the challenges that must be met for constitution of the constellation of mathematical models that will underpin the Digital Patient. Here, we address the goal of "generalization and wide use deployment of the concept of integrative modelling.[88]" At the outset, we point out that the "Digital Patient" will certainly not be a unified, monolithic, all-encompassing mathematical model of the human body from gene to organism. Rather, it will consist of many sorts of models that will be invoked as needed, depending on the nature of the question at hand, the types of data available, and the degree of understanding of the subject under scrutiny. It is important to say that in this chapter we describe the overall picture with only a limited focus on the details; however, in order to address this, we have included numerous useful references for parts we could only touch

[86] http://digitalresearch.eiu.com/extending-healthy-life-years/report/section/executive-summary
[87] http://www.epha.org/a/5131
[88] Hunter et al. 2013, Interface Focus 3:20130004

upon because of space constraints. Lastly, we acknowledge that some aspects are up for debate, since the field is a rapidly evolving one, with a wide range of contributions from many different disciplines. We would like to emphasise that what has been traded off in detail has been gained in richness of diversity, in the spirit of what is needed for the development of a Digital Patient.

Although clear boundaries among types of models are hard to define (which will facilitate the development of what will later be defined as "hybrid" modelling methodologies), this chapter will roughly demarcate two major modelling categories that will systematically appear throughout. This categorisation is not absolute, but rather delineates the terms of reference that are used when distinguishing between different types of models that embrace a large variety of modelling techniques and fit different purposes in the context of the Digital Patient:

> **Phenomenological models** are *related to the empirical observations of a phenomenon*, where a phenomenon is understood as an observable fact or event. When considering how to achieve the realisation of the Digital Patient, these models occupy a central position whenever a quick means is needed to represent pathologies quantitatively for both basic science and practical applications.
>
> **Mechanistic models**, on the other hand, aim at reaching a better understanding of the mechanisms that underlie the behaviour of the various endpoints of the biomedical process.
>
> Mechanistic models often investigate the molecular and cellular basis of biomedical processes through their physico-chemical properties. They are able to consider events at different orders of magnitude for both spatial scales (from intracellular to cell, tissue, and organ) and time scales (from the 10^{-14} s of the molecular interactions to the hours, months and years of the biomedical processes).

The diversity among the available models is clear, but their "integration" does not imply that they will all necessarily be interlinked. Further along in this chapter, we address several different pathologies and present modelling strategies for each of them. The models range from probabilistic, data-driven ("phenomenological") models with no "mechanistic" underpinning, to multi-scale, multi-physics models based on state-of-the-art understanding of the underlying anatomy and physiological mechanisms. However, while recognising that different communities favour one or the other approach, it would not serve the present purpose to focus on the relative advantages or disadvantages of "phenomenological" versus "mechanistic" modelling, since most models combine elements of both, and the real challenge lies in providing appropriate tools for quantitative exploration of a variety of clinical challenges. We will invoke both approaches, as appropriate, addressing the following key challenges to be faced for achievement of the integrated Digital Patient on the following topics:

- Selection of mathematical modelling approach
- Model personalisation in the face of multiscale complexity
- Impact of data quality on model personalisation
- Coping with complexity
- Formalisation and generalisation of model testing and validation strategies
- Translation and Clinical utilisation of models

3.2 Key challenges

3.2.1 Selection of mathematical modelling approach

To realise the Digital Patient, a first scientific challenge that needs to be addressed is the selection of the most adequate mathematical modelling approach. There is not a unique way of creating a model and many aspects determine this selection, such as model purpose and/or data availability. In a broad sense, the choice might be between phenomenological and mechanistic mathematical models, but a recurrent topic when modelling any disease is the strong link and dependency between these two modelling approaches. One example is the case of complex multi-omics structured data (embracing genotype information, metabolomics datasets, and subclinical and clinical phenotypes), which would use both data assimilation and novel mechanistic methodologies to elucidate pathological mechanisms.

Another particularly relevant example is faced when dealing with *the challenge of comorbidities*. Comorbidity is the term used to address diseases, often chronic ones, co-occurring in the same individual; i.e. an illness may develop, but health conditions also depend on another simultaneously occurring pathological process elsewhere, like for example inflammation, diabetes, or respiratory problems. As a result of this complexity, it is likely that modellers will often find themselves in between the two types of modelling methodologies, since there is no clear practical separation between phenomenological and mechanistic methods.

Finally, efforts from the modelling community are needed to use existing VPH reference ontologies to annotate the resulting models and to make them available through VPH common model repositories. This formalisation will enable interoperability between models in general. Examples of initiatives pursuing this objective are the two EU-funded VPH projects VPH-Share[89] and p-medicine[90]. These two projects are collaborating together and seeking complementarities. VPH-Share is working to provide the essential services and computational infrastructure for the sharing of clinical and research data and tools, facilitating the construction and operation of new VPH workflows, and collaborations between the members of the VPH community. In this project, evaluating the effectiveness and fitness-for-purpose of the infostructure and developing a thorough exploitation strategy are key activities to create confidence and engage the communities. P-medicine intends to go from data sharing and integration via VPH models to personalised medicine. The emphasis is on formulating an open, modular framework of tools and services, so that p-medicine can be adopted gradually, including efficient secure sharing and handling of large personalized data sets, enabling demanding Virtual Physiological Human (VPH) multiscale simulations (e.g., in silico oncology), building standards-compliant tools and models for VPH research, drawing on the VPH Toolkit[91] and providing tools for large-scale, privacy-preserving data and literature mining, a key component of VPH research.

[89] http://www.vph-share.eu
[90] http://www.p-medicine.eu
[91] http://toolkit.vph-noe.eu/

An important consideration valid for mechanistic and phenomenological models is that biomedical systems characterisation is more rationally and robustly addressed when driven by a comprehensive, rather than selective, use of all the available information, provided the varying degree of accuracy of the components of the evidence base is correctly recognized.

Among the wealth of methods available for phenomenological modelling, Bayesian techniques for multi-parameter evidence synthesis have demonstrated a rational and exhaustive use of the whole body of information available for decision models. [92]

Recently, methodologies resulting in a combination of Bayesian inference for partially observed Markov process models and nonlinear dynamical systems approaches have also been developed. [93]

Other techniques encompass formal methods. It is worth noting that novel concepts and terminologies originally developed in the theoretical computer science domain (for example executable models, expressivity, abstraction, statistical model checking, stabilization, reach ability analysis, formal verification), which are scarcely known by other modelling communities (for example engineers and physicists), are providing important insights and tools for modelling and analysing complex biological systems. The key concept here is the distinction between the mathematical model and the computational model. The two terms are tightly related, since the computational model is a mathematical model executed by a computer.

Other approaches (including process algebra, hybrid systems, Petri nets, state-charts) provide a battery of methodologies[94]. One example is the verification of a property representing a condition of illness or the effect of a drug; we could imagine that at the clinical level, computer-aided therapies and treatments will develop into intervention strategies undertaken under acute disease conditions or due to external factors (like infections) to contrast cascade effects. In non-acute states, predictive inference will propose prevention plans for comorbidity management.

Additional methodologies are based on dynamical systems theory, particularly chaos and fractals[95]. In healthy tissue, a full repertoire of receptors and ion channels respond to mechanical micro-stress events and generate small highly variable noise in a variety of physiological signals such as heartbeat, blood pressure, gait, and nephrons[96] etc. In disease or aging conditions, we commonly observe a reduction of such variability and more smoothness. This difference in biological signals has been used to extract useful information

[92] see an example in Ades AE, Welton NJ, Caldwell D, Price M, Goubar A, Lu G. Multiparameter evidence synthesis in epidemiology and medical decision-making. J Health Serv Res Policy. 2008 Suppl 3:12-22

[93] EDWARD L. IONIDES, ANINDYA BHADRA, YVES ATCHADÉ AND AARON KING Iterated Filtering, The Annals of Statistics 2011, Vol. 39, No. 3, 1776–1802

[94] Fisher J, Henzinger TA (2007) Executable cell biology. Nat Biotechnol 25: 1239-1249; Bartocci E, Corradini F, Di Berardini M, Merelli E, Tesei L (2010) Shape calculus. a spatial mobile calculus for 3 d shapes. Scientific Annals of Computer Science 20: 2010; Heiner M, Gilbert D, Donaldson R (2008) Petri nets for systems and synthetic biology. In: SFM. Springer, number 5016 in Lecture Notes in Computer Science, pp. 215-264; Setty Y, Cohen IR, Dor Y, Harel D (2008) Four-dimensional realistic modeling of pancreatic organogenesis. Proc Natl Acad Sci USA 105: 20374-20379; Bartocci E, Cherry EM, Glimm J, Grosu R, Smolka SA, et al. (2011) Toward real-time simulation of cardiac dynamics. In: Proceedings of the 9th International Conference on Computational Methods in Systems Biology. New York, NY, USA: ACM, CMSB '11, pp. 103-112. doi: 10.1145/2037509.2037525. URL http://doi.acm.org/10.1145/2037509.2037525

[95] Seely AJ, Macklem P. Fractal variability: an emergent property of complex dissipative systems. Chaos. 2012 Mar;22(1):013108. doi: 10.1063/1.3675622

[96] Laugesen JL, Mosekilde E, Holstein-Rathlou NH. Synchronization of period-doubling oscillations in vascular coupled nephrons. Chaos. 2011 Sep;21(3):033128. doi: 10.1063/1.3641828.

about the state of the patients and to create diagnostic tests. The mathematical characteristics of these signals resemble those found in deterministic chaos, fractals, and self-organizing emergent properties of complex dissipative systems; in many cases, the physiological signals are studied by means of wavelets. Cardiovascular physiological signals provide a rich literature of fractal and/or chaotic behaviour, particularly related to the His-Purkinje network of the heart[97] [98]. While the gait of healthy adults follows a scale-free law with long-range correlations extending over hundreds of strides, the fractal properties of gait are modified in Parkinson's disease. Notably, we observe a decrease in the gait correlation and changes in stride length and gait variability[99] [100] [101]. In the case of Alzheimer's disease, the analysis of the fractal dimension of the EEG is used to discriminate patients affected by the disease from control groups with an accuracy of 99.3%, sensitivity of 100%, and a specificity of 97.8%[102].

3.2.2 Model personalisation in the face of multiscale complexity

The formulation of mathematical models for medicine represents a real challenge, not only because we do not fully understand all the pathophysiological mechanisms, but also because many illnesses are prolonged in duration and, in the case of chronic diseases (like for instance stroke, diabetes, arthritis, osteoporosis, atherosclerosis, etc.), are generally managed rather than cured. Including ageing as one of the elements in personalised models is an example in itself of such complexity, and we have devoted a box to that (see Box at the end of the chapter).

One possible strategy to model disease over time is to use hypothesis-based models that combine mechanistic and phenomenological elements corresponding to the degree of understanding of the different components. These models should be adjusted to represent specific time-points in the patient's evolution in order to calculate expected progression of clinical indicators based on probabilistic models that are, in turn, rooted in population data. This approach takes advantage of state-of-the-art knowledge of the (patho-)physiology while also exploiting the mass of data becoming available from clinical trials and epidemiology, on one hand, and GWAS and other molecular data, on the other hand. In appropriate cases, and as pointed out in the VPH Vision document (Hunter al. 2013), "the VPH initiative may be of substantial help by providing mechanistic model descriptions of the phenotypic effects originating from genomic network variation[103]. Such causally cohesive genotype–phenotype models are very advanced multiscale physiological models with an explicit link to molecular information and with the capacity to describe, for example, how genetic variation manifests in phenotypic variation at various systemic levels up to the tissue, organ and whole-organism level."

---

[97] Sharma V. Deterministic chaos and fractal complexity in the dynamics of cardiovascular behavior: perspectives on a new frontier. Open Cardiovasc Med J. 2009 Sep 10;3:110-23. doi: 10.2174/1874192400903010110.

[98] Schmitt DT, Ivanov PCh. Fractal scale-invariant and nonlinear properties of cardiac dynamics remain stable with advanced age: a new mechanistic picture of cardiac control in healthy elderly. Am J Physiol Regul Integr Comp Physiol. 2007 Nov;293(5):R1923-37. Epub 2007 Aug 1.

[99] Hausdorff JM. Gait dynamics in Parkinson's disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling Chaos. 2009 Jun;19(2):026113. doi: 10.1063/1.3147408

[100] Hausdorff JM, Ashkenazy Y, Peng CK, Ivanov PC, Stanley HE, Goldberger AL. When human walking becomes random walking: fractal analysis and modeling of gait rhythm fluctuations. Physica A. 2001 Dec 15;302(1-4):138-47

[101] Beuter A, Modolo J. Delayed and lasting effects of deep brain stimulation on locomotion in Parkinson's disease. Chaos. 2009 Jun;19(2):026114. doi: 10.1063/1.3127585

[102] Ahmadlou M, Adeli H, Adeli A. Alzheimer Dis Assoc Disord. 2011 Jan-Mar;25(1):85-92. doi: 10.1097/WAD.0b013e3181ed1160. Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease

[103] Shublaq N, Sansom C, Coveney PV. 2013 Patient- specific modelling in drug design, development and selection including its role in clinical decision- making. Chem. Biol. Drug Des. 81, 5–12. (doi:10. 1111/j.1747-0285.2012.01444.x)

Consider also the amount of data routinely collected on the millions of patients treated in worldwide healthcare systems, not to mention the rapidly expanding knowledge of human genetics. We are now at a point where computing power and mathematical modelling are becoming able to make use of such vast amounts of information, despite the inevitable noise from 'random variation'. New perspectives may be required in which systems can learn from data, even generated at the individual level. One may even conjecture that it may be possible to develop individual disease models for each person. For example, consider the case of identical twins with similar lifestyles; one develops a chronic disease 30 years before the other. Is it possible to create models that are capable of accommodating such apparent discrepancies? The answer is positive in principle, because models can input lifestyle information as well as physiological, metabolic, and genomic data. Nevertheless, this will require not only the coupling between models at different time and length scales, as well as the description of different physical phenomena, but also between different (phenomenological and mechanistic) modelling paradigms. Until recently, the complexities of many diseases have made them almost intractable for modellers, biologists, and clinicians. Current computational capacities should enable effective modelling of systems that are relevant for therapy. A more systematic, all-encompassing, intelligent and informed update of the models would provide a way forward in order to overcome fragmentation and to address lack of data.

It should be stressed that this is also a technological challenge. For example, multiple interacting processes are typically well described individually by different modelling approaches, like ordinary vs partial differential equations (ODEs vs. PDEs). When such heterogeneous models are coupled naively, the resulting hypermodel[104] may become difficult or impossible to solve. It is thus necessary to develop better – and where possible generic – ways to deal with such coupling problems, so that they can be embedded in open access software libraries for common research use, even if individual researchers or teams lack a deep understanding of the underlying complex mathematical, numerical, and computational techniques involved.

Last but not least, it is essential to highlight the role of boundary conditions for models. These boundary conditions will have to be personalisable, robust and efficient.

3.2.3 Impact of data quality on model personalisation

Biomedical data sources commonly include incomplete entries mainly because of the difficulties at the data collection stage (see chapters 1 and 2 for more details). On the one hand, this complicates the modelling process that needs to cope with this situation, while on the other hand it reduces the confidence of the conclusions derived from the models that have been built. Bayesian statistical inference provides a mathematically consistent framework for encoding uncertain information at the time of observation by obtaining a posterior measure over the space of "trajectories" of the biomedical process. For example, although missing data and the difficulty of dealing with lifestyle or self-reported questionnaire data will decrease the quality of the data, the statistical inference provides a powerful means to constrain probability measures over the causal spaces. Bayesian methods could also

[104] Within this document we use the term 'hypermodel' to describe a concrete instance of an integrative model, built as the orchestration of multiple computer models that might run on different computers at different locations, using different simulation software stacks. Typically a hypermodel is a computational model that might operate on multiple spatial scales, perhaps from molecular through cellular to organ and patient level, and/or on multiple temporal scales, from acute response upwards, and/or might include descriptions of physical, chemical and biological processes.

make use of the available information to predict missing values[105]. The impact of using these methods would need to be evaluated during model validation.

There will always be issues with the quality of data collected routinely in hospitals, especially if the aim is to use models in clinical practice, because data quality depends on many factors, including human skill. This issue could potentially be addressed by the development of new methods able to *integrate* the knowledge from experts and the data collected.

3.2.4 Coping with complexity

The complexity of most diseases and the different answers that various types of data would provide remind of the popular tale “The Blind Men and the Elephant” by John Godfrey Saxe (1816-1887). The author writes about a group of blind men who touch an elephant. Each person feels only one part, such as the side or the tusk; so when they describe to each other what they have found, they are in complete disagreement.

For many diseases the most important step is the identification of key model parameters, which can often be measured with only limited accuracy. This issue becomes more critical when multiscale models exhibit nonlinear behaviour, where small variations in certain input parameters could produce significant differences in the output predictions. Clearly some parameters cannot be directly measured on the patient of interest, so one has to use values derived from estimated population mean and variance or from animals. When identification of key parameters that could actually be determined or estimated is problematic, non-parametric models might also be effective.

It is becoming evident that in order to approach the complexity, model order-reduction techniques are sorely needed, while to overcome the sparsity and the variable relevancy and quality of the data it is often important to consider coupling mechanistic with phenomenological modelling.

For single-scale models, researchers have developed a number of methods to account for uncertainties and variability, most of which require Monte Carlo techniques. However, in the case of large multiscale models intended as the orchestration of multiple submodels, transformation of some or all of these submodels into stochastic models leads to heavy computational costs. For example, in the VPHOP project referenced below in the Exemplars section, a full cell-to-organism multiscale musculo-skeletal model in which only the daily physical activity and the related risk of falling were modelled as a stochastic process. The estimation of the risk of bone fracture over 10 years required over 65k core-hours of calculations for each patient. While the final VPHOP hypermodel runs 50 times faster, this was achieved by introducing considerable simplifications in some of the most complex processes.

The applied mathematics community has developed a number of methods – such as Markov-Chain Monte Carlo[106] and the method of Morris[107] – that address aspects of this problem. Nonetheless, we need to target these general modelling techniques to the specific problems, validate them extensively, and make them available to the VPH research community in ways that make their widespread adoption possible given their considerable

[105] Nguyen, V. A., Koukolikova-Nicola, Z., Bagnoli, F., & Lio’, P. , 2009 Noise and non-linearities in high-throughput data. J STAT MECH THEORY doi:10.1088/1742-5468/2009/01/P01014

[106] Persi Diaconis. The Markov chain Monte Carlo revolution. Bulletin of the American Mathematical Society. 46 (2009), 179-205. MSC (2000): Primary 60J20. Posted: November 20, 2008

[107] Jeffrey S. Morris. "Statistical Methods for Proteomic Biomarker Discovery Based on Feature Extraction or Functional Modeling Approaches" Statistics and Its Interface 5.1 (2012): 117-136.

complexity. We also need to strengthen the stochastic physics background of our students and post-docs as our research sector develops in this direction.

In the case of phenomenological models, an illustrative example is the study of links between morbidities and risk evaluation for a specific pathology. The phenomenological model aims at organising the wealth of observations within a formal structure. One aspect of this challenge is that the connection between data availability and the creation of the phenomenological model is often bridged by human experts, causing a major bottleneck in their ability to understand and engineer complex biomedical systems.

3.2.5 Formalisation and generalisation of model testing and validation strategies

In general, scientific progress towards *creating* and *validating* any model generally relies on asking the right questions, and this is far from a banal statement. Different modelling methodologies often answer slightly different questions and, as a consequence, different studies use methodologies that are difficult to cross-compare. Differences between an approximate and an exact model are usually remarkably less than the disparity between the exact model and the real biological process[108]. In such cases, the knowledge that an "expert system" could provide to understand a specific pathology and predict its course is often placed into question rather than believed and built upon.

One fundamental aspect of personalised models of any kind is that they should always be subjected to a *sensitivity and robustness analysis*, concepts that are intimately linked to the notion of “validation”. The sensitivity analysis aims to identify the most influential model parameters, including dependencies among input parameters and between inputs and outputs. The robustness analysis aims to evaluate the probability that the model will not deviate much from a reference state in the face of perturbations. If a given biological process itself is robust to external perturbations, then analysis of successful models that represent that process will be valuable to further our understanding of the biological mechanisms underlying the real system's robustness. Moreover, models will only be adopted in the clinic once they have satisfied the sensitivity and robustness requirements that will make them useful *in practice*. Only by knowing their limitations and how well they are able to make credible predictions or diagnoses with small or large differences in the input data will clinicians feel confident enough to use them as a tool for personalised diagnosis and treatment, since this is directly linked to issues of patient safety. Here lies one of the main challenges of the Digital Patient: clinical acceptance of the patient models that will be developed in the future. This is further discussed in the sections below.

3.2.6 Translation and Clinical utilisation of models

Even though mechanistic models are complex in nature, some have already entered the clinical arena in the form of software applications embedded in diagnostic or therapeutic devices. Examples of such models are pressure wave propagation models as implemented in the Nexfin monitor, by BMeye B.V. to evaluate central blood pressure and cardiac output from finger plethysmography. Another example of a device is the pacemaker with IAD (Medtronic and others), where a model-based decision algorithm controls the defibrillation action. At an early stage in the promotion of models for clinical use, the more or less generic models used in the applications above could be expanded and personalized to increase the use of validated models in the clinic.

[108] D. J. Wilkinson, Stochastic modeling for quantitative description of heterogeneous biological systems, Nat. Rev. Genet., 2009, 10, 122-133

In pharma, clinical pharmacology is an integral part of clinical trials and the approval of a new drug. Empirical (non-mechanistic), semi-mechanistic (e.g. pharmacokinetic/ pharmaco-dynamic – PKPD models) and more mechanistic methods (e.g. physiologically based pharmacokinetics – PBPK models) have been part of quantitative pharmacology, or pharmacometrics. New mechanistic models in drug development are trying to include more information about the biology, pharmacology, disease, and physiology in order to describe and quantify the interactions between xenobiotics and patients, including beneficial effects and side effects that result from such interfaces. A new emerging area called systems pharmacology is being developed as an approach to translational medicine that combines computational and experimental methods to elucidate, validate, and apply new pharmacological concepts to the development and use of small molecule and biological drugs to determining mechanisms of action of new and existing drugs in preclinical and animal models, as well as in patients. Approaches related to pharmacometrics – in particular PKPD modelling – are increasingly being applied in the development of novel therapeutics. The impact of these investments is being supported by both pharmaceutical research organizations and regulatory agencies[109].

Non-mechanistic methods such as non-compartmental analysis (NCA) require regulatory approval for new drug application (NDA). NCA provides a framework to use statistical moment analysis to estimate pharmacokinetic parameters dependent on total drug exposure. Some of the parameters obtained from NCA have a practical meaning and can be interpreted, such as the volume of distribution or clearance. However, the parameters provide little insight into physiology, nor how patients will behave towards a different set of conditions. NCA still plays an important role in bioequivalence studies and rapid analysis, but the utility and impact of pharmacokinetic data has increased massively since the arrival of the more mechanistic population approaches.

In the cancer field, models of the MAP kinase pathway around the EGF receptor can be used for the individualization of the treatment of some cancers[110], and models of the Warburg effect advise the dynamic dosing of new glycolytic inhibitors of tumorigenesis[111,112]. A final example is the differential network-based drug design models for parasitic diseases such as malaria and trypanosomiasis[113].

These examples indicate that simple models that cover a relatively small part of a pathology or mechanical process are most likely to be adopted for clinical use soonest. As mentioned above, a key point in this process is the validation of the models before actual clinical adoption can occur on a larger scale. Proof of the specificity and sensitivity of models in, for instance, diagnostic tools, is and will continue to be crucial to this adoption process. In the textbook developed by the VPH Network of Excellence (to be published by OUP), Viceconti and Graf draft a framework for pre-clinical and clinical model validation, as well as for clinical acceptance. While such a framework may appear complex and demanding, we believe it is only in this way that the natural resistance against computer simulations will be overcome, proving conclusively that the Digital Patient technologies are accurate, robust, and clinically effective.

[109] Van der Graff, P., CPT: Pharmacometrics & Systems Pharmacology (2012) 1, e8

[110] K Oda, Y Matsuoka, A Funahashi & H Kitano, A comprehensive pathway map of epidermal growth factor receptor signaling, Molecular Systems Biology 1:2005.0010, 2005

[111] H Pelicano et al. Glycolysis inhibition for anticancer treatment, Oncogene (2006) 25, 4633-4646

[112] D. A. Tennant, R. V. Durán and E. Gottlieb, Targeting metabolic transformation for cancer therapy, NATURe RevIeWS | Cancer, 10, 267-277, 2010

[113] Tekwani BL: Current antiparasitic drug targets and a paradigm shift in discovery of new antiparasitic drugs and vaccines. Curr Drug Targets 2008, 9:921

An example of multiscale complexity: Ageing

In the latest VPH NoE strategy document (Hunter et al., VPH Vision 2013) it was written: “age is the dominant risk factor for most complex diseases. The making of multiscale physiological models capturing the ageing process defines a very ambitious long-term theoretical – experimental research programme of vital importance to the VPH vision.”

Ageing is a hurdle to overcome and its inclusion in personalised models for the Digital Patient is a challenge that multi-scale models will need to resolve. Already when merely considering the multifarious interconnections between ageing and lifestyle and genetic factors, one can appreciate the complexity of this dynamic process. The age factor is successfully used in epidemiological studies to specify, for example, the contact rate of the spread in an infectious disease, which summarizes the infectious effectiveness of contacts between susceptible and infectious subjects. However, even for population/epidemiological studies where functional biological relationships between the different causes and effects of ageing in the model are not relevant, epidemic models that take into account the age structure of a population are very intricate.

One possibility is to identify genomic markers most closely associated with age and related disease traits. In the VPH-FET roadmap[114], it was highlighted that “[...] combining genomic, proteomic, metabolomic and environmental factors may provide insights into pathogenomic mechanisms and lead to novel therapeutic targets”. In diseases precipitated by complex interplays of genetic predisposition and a broad spectrum of environmental and nutritional factors, the challenge is immense. In this context, epidemiological factors such as urbanicity, geographical distribution, migration behaviour, and maternal risk factors such as infections, malnutrition and adverse life events during pregnancy, have been suggested as being relevant to different extents. It is clear that a combination of different types of modelling paradigms will be necessary to establish the relationship between these factors and the interplay with genetic determinants, which thus far remains unknown. Integrated, system-based investigations are a promising approach to obtaining deeper insights into the disease aetiology and its management or cure.

In mechanistic-type models, the physiological aspects of ageing can be represented as time-dependent changes in relevant model parameters. On a population level, a number of such factors have been identified, including bone density in women, gradual reduction of renal function (i.e. falling GFR), reduced mobility (e.g., increased sitting-to-standing time), elevated pule arterial pressure, elevated TPR (total peripheral resistance), and left ventricular hypertrophy.. As the mechanisms responsible for these age-related changes become elucidated, the corresponding details in the models can be adjusted accordingly, and when the mechanisms are unknown these changes will be reflected in appropriate phenomenological model parameters. However, ageing is perhaps a case where phenomenological models are easier to build (taking advantage of the wealth of observational data available) compared to the mechanistic ones, since mechanistic aspects of ageing are less well known.

[114] https://www.biomedtown.org/biomed_town/VPHFET/reception/vphfetpublicrep/plfng_view

3.3 Modelling different pathologies - exemplars

Patho-physiological phenomena must be interpreted from clustering of extracted features, time evolution, and multi-parameter analysis. For “black box” or data-driven models, this might be enough; an additional step of formulation of cause-effect relationships is needed for “mechanistic models”. In any case, the lack of data covering decades hampers the effectiveness of most types of models. There are very few longitudinal studies available and this is one of the greatest challenges in modelling disease. Six examples in which modelling has been successfully used in clinical applications are described below in relative detail, but we also highlight areas in which gaps and unmet needs are evident. A comprehensive view of the six is presented in the table below and descriptions follow below in the text.

TABLE 1: COMPARATIVE OVERVIEW OF MAIN FEATURES OF SIX DIFFERENT PATHOLOGIES

Pathology	Comorbidities	Phenomenological modelling	Mechanistic modelling
Breast cancer	COPD; CHF; stroke	Tumor diameter growth; biopsies; histology; development of stage diagnosis	Cell invasiveness based on prognostic and diagnostic molecular markers
Osteoporo-sis	Several types of cancers (breast, prostate, multiple myelomas); endocrine unbalance; infections (HIV); therapies (HAART)	Bone mineral density; Wolff’s law; development of Frax tool.	Molecule-to-cell, cell-to-tissue coupling models, for example osteocytes, hormones
Atherosclerosis	Inflammation; obesity; diabetes	Imaging: CT, MRI or US 3D+T with resolution of 1 mm per voxel; arterial elasticity; plaque biomechanics in general; restenosis after stenting	Proliferation and migration of vascular smooth muscle cells; plasma lipoproteins (LDL and HDL),
Cardiomyo-pathy	Obesity; diabetes; coronary artery disease; hypertension; infection	ECG patterns; abundance of longitudinal studies (Framingham[115], Dawber, Busselton)	Energy metabolism based on glycolysis; mitochondrial functionality; lactate production; ionic (sodium) currents; excitation / contraction of single cells
Dementia	Stroke and heart failure prediction tools predict dementia (Kaffashian[116])	Cognitive tests and memory; EEG patterns; MRI; brain mapping; network models of atrophy; use of longitudinal data	β-Amyloid plaques; neurofibrillary tangles; tau phosphorylation
Stroke	Hypertension; coronary disease and diabetes; the Charlson comorbidity	Cognitive tests and memory; performance; Charlson index; neuroimaging	Based on oxidative DNA damage and repair; vasoconstrictor such as endothelin-1

[115] Thomas R. Dawber, M.D., Gilcin F. Meadors, M.D., M.P.H., and Felix E. Moore, Jr., National Heart Institute, National Institutes of Health, Public Health Service, Federal Security Agency, Washington, D. C., Epidemiological Approaches to Heart Disease: The Framingham Study Presented at a Joint Session of the Epidemiology, Health Officers, Medical Care, and Statistics Sections of the American Public Health Association, at the Seventy-eighth Annual Meeting in St. Louis, Mo., November 3, 1950.

[116] Kaffashian S, et al "Predicting cognitive decline: A dementia risk score vs the Framingham vascular risk scores" Neurology 2013; 80: 1300–1306.

3.3.1 Breast Cancer

Multistage cancer models are widely used to model solid tumours that appear to develop through well-defined phenomenological stages, including initiation, pseudo-tumoral and cancer transformation[117]. Here the histological analysis, screening, and clinical incidence data could be used to calibrate, validate and check the consistency of the several sources of evidence and define the stage of the cancer. The phenomenology of breast cancer disease is related to its aggressiveness, which stems from its rapid recurrence and metastasis positioning. Phenomenological models of breast cancer use a wide range of parameters related to imaging, pathology, basic research, clinical trials, clinical practice, genetic predisposition, and epidemiology. The most meaningful parameter set from these analyses could be wrapped to construct new phenomenological parameters to describe growth rhythms, growth delays, and time constants. This modus operandi introduces a vast simplification by turning a system with a large number of constituents specific to the used techniques into a limited number of effective degrees of freedom embedded in a few phenomenological parameters[118]. For example, in breast cancer, a diffuse redness provides evidence that inflammatory processes are involved in the pathogenesis of this disease, which is rare but the most aggressive form of breast cancer.

It is clear though that in order to understand the processes behind tumour growth and treatment, other (more mechanistic) approaches are required. For example, the ContraCancrum project[119] aimed at bringing together different levels of biocomplexity producing an integrated oncosimulator and validating it on two dedicated clinical studies concerning glioma and lung cancer. The project modelled and simulated cancer vs. normal tissue behaviour at different levels of biocomplexity, and also modelled a facet of the systemic circulation via pharmacokinetics, and synthesised models of hematological reactions to chemotherapy.[120,121]

One interesting proposition is to try harnessing the power of epidemiological studies in conjunction with a systemic mechanistic approach, as proposed by Sokhansanj and Wilson[122]. They describe a mathematical model that mimics the kinetics of base excision repair and thus permits them to investigate in silico the effects of genetic variation in this important DNA repair pathway. As written in [123] *"If one succeeds in constructing a mathematical model that reasonably represents the biochemical reality, the payoff is large. One can experiment with the model by increasing or decreasing inputs (corresponding, say, to changes in diet) or by raising or lowering activities of enzymes (corresponding to genetic*

[117] see an example in Wai-yuan Tan, Leonid Hanin Handbook Of Cancer Models With Applications (Series in Mathematical Biology and Medicine) World Scientific Pub Co Inc; 1 edition (August 11, 2008)

[118] Bastogne T, Samson A, Vallois P, Wantz-Mézières S, Pinel S, Bechet D, Barberi-Heyob M. Phenomenological modeling of tumor diameter growth based on a mixed effects model. J Theor Biol. 2010 Feb 7;262(3):544-52.)

[119] http://www.contracancrum.eu/

[120] A. Roniotis, K. Marias, V. Sakkalis, and G. Stamatakos "Mathematical guide for developing a heterogeneous, anisotropic and 3-dimensional glioma growth model using the diffusion equation", Information Technology Applications in Biomedicine (IEEE-ITAB 2009), Larnaca, Cyprus, 2009.

[121] G. S. Stamatakos, D. Dionysiou, S. Giatili, E. Kolokotroni, E. Georgiadi, A. Roniotis,V. Sakkalis, P. Coveney, S. Wan, S. Manos, S. Zasada, A. Folarin, P. Büchler, T. Bardyn, S. Bauer, M. Reyes, T. Bily, V. Bednar, M. Karasek, N. Graf, R. Bohle, E. Meese, Y.-J. Kim, H. Stenzhorn, G. Clapworthy, E. Liu, J. Sabczynski, and K. Marias, "The ContraCancrum Oncosimulator: Integrating Biomechanisms Across Scales in the Clinical Context", 4th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation, Athens, Greece, September 8-9, 2010.

[122] Sokhansanj BA, Wilson DM. Estimating the impact of human base excision repair protein variants on the repair of oxidative DNA base damage. *Cancer Epidemiol Biomarkers Prev* 2006;15:1000–8

[123] http://cebp.aacrjournals.org/content/15/5/827.full

polymorphisms), or eliminating entire reactions completely (corresponding to gene-knockout experiments). One can take apart and put back together the biochemical network piece by piece to determine how it works. In contrast to biological experiments, these in silico experiments are quick and inexpensive and, if done well, can give real insight into the genetic and molecular network".

3.3.2 Osteoporosis

Bone is one of the most adaptable tissues in the body. Accurate phenotypic descriptions of human skeletal phenomena are starting to accumulate[124]. During adulthood, there is a stable equilibrium (homeostasis) with the formation of new bone by the osteoblasts and the removal of older bone tissue by the bone-resorbing osteoclasts. This homeostasis can be perturbed by aging (osteoporosis), infections (osteomyelitis[125]), changes in physical activity, or through metabolism. Due to the deposition of collagen in particular directions by osteoblasts, bone acquires anisotropic properties with an alignment of the principal directions of the bone (trabeculae) with the principal direction of stresses, known as Wolff's law. Hence there is a direct relationship between bone adaptation and mechanical loading. This sensitive equilibrium is broken at a later stage in life when the osteoblast activity is reduced, leading to osteoporosis and an increased fragility of bone. The osteoporosis case shows an excellent example of mechanistic modelling put to the service of the clinical community.

The mechanistic approach lends itself with relative ease to the understanding of osteoporosis; bone and muscle have been active and successful research strands in biomechanics for decades. Multiscale modelling and simulation approaches have tried to bridge the spatial and temporal gaps involved. For example: by detailed modelling of musculoskeletal anatomy and neuromotor control that define the daily loading spectrum, including paraphysiological overloading events; by modelling fracture events as they occur at the organ level and are influenced by the elasticity and geometry of bone, which leads directly to the tissue scale as bone elasticity and geometry are determined by tissue morphology and finally reaching the cell, as cell activity changes tissue morphology and composition over time. Some examples of this are found in [126,127]

Several types of phenomenological models have also been proposed, for example based on PDE solvers using histological and micro-CT image information[128], a topological osteoactivity metric, i.e., the resorption-formation steady-state is represented as a torus in multidimensional phase space[129], a process-algebraic specification (for example, the space-defined Shape Calculus), which provides an effective multiscale description of the process. The phenomenological approaches make use of abundant bone mineral density data in health and pathology (for example osteoporosis[130]).

---

[124] see for example Groza T, Hunter J, Zankl A. Decomposing phenotype descriptions for the human skeletal phenome. Biomed Inform Insights. 2013;6:1-14. doi: 10.4137/BII.S10729. Epub 2013 Feb 4

[125] Liò P, Paoletti N, Moni M.A., Atwell K, Merelli E. and Viceconti M, Modelling osteomyelitis, BMC Bioinformatics, 13: S12, doi:10.1186/1471-2105-13-S14-S12.

[126] Gerhard FA, Webster DJ, van Lenthe GH, Müller R. In silico biology of bone modelling and remodelling: adaptation. Philos Trans A Math Phys Eng Sci. 2009 May 28;367(1895):2011-30. doi: 10.1098/rsta.2008.0297.

[127] Bonjour, J.P., et al., Peak bone mass and its regulation, in Pediatric Bone, Second Edition, F.H. Glorieux, J.M. Pettifor, and H. Jüppner, Editors. 2011, Academic Press Inc, Elsevier.

[128] Viceconti, M., Clapworthy, G., Testi, D., Taddei, F., and McFarlane, N. (2011). Multimodal fusion of biomedical data at different temporal and dimensional scales. Comput Methods Programs Biomed., 102:227–237.

[129] Moroz A, Crane MC, Smith G, Wimpenny DI. Phenomenological model of bone remodeling cycle containing osteocyte regulation loop. Biosystems. 2006 Jun;84(3):183-90

[130] Liò P, Merelli E.and Paoletti N, (2012) Disease processes as hybrid dynamical systems, Proceedings of the 1st International Workshop on Hybrid Systems and Biology (HSB 2012), EPTCS 92, pp. 152-166; Paoletti, N.,

One associated and interesting aspect of phenomenological modelling is its use in identifying the major reasons for osteoporotic fractures. While intensive work continues into evaluating bone loss and the aetiology of skeletal osteolysis throughout the ageing process, the single major cause for an osteoporotic fracture is the occurrence of falls. The accurate prediction of fracture risk can therefore only be achieved by observational studies that lead to an understanding of the factors that play a beneficial or detrimental role in modifying an individual's risk of fall. Fall risk assessment currently varies from questionnaire-based evaluation of health and medication factors to intensive laboratory measurements for quantification of gait and balance parameters[131]. However, most of these tools have been shown to discriminate poorly between fallers and non-fallers[132,133]. The best assessment tools currently achieve a sensitivity and specificity of around 75%[134,135]. In clinical assessments, the single best predictor for falls has been the existence of a previous fall[136,137,138]. While this increases the accuracy of fall risk assessment in retrospective studies where subjects have already fallen, identification of future fallers becomes challenging in prospective cases when a prognosis for a subject who has not yet fallen is required. In subjects with no previous falls, kinematic abnormalities during gait and balance seem to contain important information related to the likelihood of a future fall[139]. The successful identification and inclusion of such functional indices – including balance [140,141], temporal and spatial variability during gait [142,143], muscle strength[144] – is now thought to contribute towards

Lio', P., Merelli, E., & Viceconti, M. (2012). Multilevel computational modeling and quantitative analysis of bone remodeling.. IEEE/ACM Trans Comput Biol Bioinform, 9(5), 1366-1378, Liò, P., Merelli, E., Paoletti, N., & Viceconti, M. (2011). A combined process algebraic and stochastic approach to bone remodeling. Electronic Notes in Theoretical Computer Science, 277(1), 41-52

[131] Persad CC, Cook S, Giordani B: Assessing falls in the elderly: should we use simple screening tests or a comprehensive fall risk evaluation? European journal of physical and rehabilitation medicine 2010, 46(2):249-259

[132] Gates S, Smith LA, Fisher JD, Lamb SE: Systematic review of accuracy of screening instruments for predicting fall risk among independently living older adults. Journal of rehabilitation research and development 2008, 45(8):1105-1116.

[133] Oliver D, Papaioannou A, Giangregorio L, Thabane L, Reizgys K, Foster G: A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work? Age and ageing 2008, 37(6):621-627.

[134] Persad CC, Cook S, Giordani B: Assessing falls in the elderly: should we use simple screening tests or a comprehensive fall risk evaluation? European journal of physical and rehabilitation medicine 2010, 46(2):249-259

[135] Yamada MA, H.; Nagai, K.; Tanaka, B.; Uehara, T.; Aoyama, T.: Development of a New Fall Risk Assessment Index for Older Adults. International Journal of Gerontology 2012, 6:160-162

[136] Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ: Will my patient fall? JAMA : the journal of the American Medical Association 2007, 297(1):77-86

[137] Bongue B, Dupre C, Beauchet O, Rossat A, Fantino B, Colvez A: A screening tool with five risk factors was developed for fall-risk prediction in community-dwelling elderly. Journal of clinical epidemiology 2011, 64(10):1152-1160.

[138] Gerdhem P, Ringsberg KA, Akesson K, Obrant KJ: Clinical history and biologic age predicted falls better than objective functional tests. Journal of clinical epidemiology 2005, 58(3):226-232

[139] Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ: Will my patient fall? JAMA : the journal of the American Medical Association 2007, 297(1):77-86

[140] Sherrington C, Lord SR, Close JC, Barraclough E, Taylor M, O'Rourke S, Kurrle S, Tiedemann A, Cumming RG, Herbert RD: A simple tool predicted probability of falling after aged care inpatient rehabilitation. Journal of clinical epidemiology 2011, 64(7):779-786

[141] Swanenburg J, de Bruin ED, Uebelhart D, Mulder T: Falls prediction in elderly people: a 1-year prospective study. Gait & posture 2010, 31(3):317-321

[142] Taylor ME, Delbaere K, Mikolaizak AS, Lord SR, Close JC: Gait parameter risk factors for falls under simple and dual task conditions in cognitively impaired older people. Gait & posture 2013, 37(1):126-130

[143] Hamacher D, Singh NB, Van Dieen JH, Heller MO, Taylor WR: Kinematic measures for assessing gait stability in elderly individuals: a systematic review. Journal of the Royal Society, Interface / the Royal Society 2011, 8(65):1682-1698

accurate predictions of fall risk when combined with established clinical parameters (e.g. medication, cognition), and may therefore allow improved stratification of elderly subjects in a clinical setting. Here, by investigating the functional movement and muscular control characteristics that differentiate subjects who are most susceptible to falling, observational studies are important in improving our understanding of the aetiology of falls, but may well play a key role pushing the boundaries for the early clinical identification and stratification of subjects at risk of falls.

It is noteworthy that the *phenomenological modelling could in principle be applied to study the system*: human body, sensor networks, prostheses, which could be tested and validated in a very effective way without a precise mechanistic model.

3.3.3 Atherosclerosis

Atherosclerosis is a multifactorial disease in which not only genetic, biochemical, and physiological factors play a role, but also environmental and life-style factors. This pathology is a prime example of complex processes acting along multiple biological, length and time scales. In this disease, lifestyle is particularly important and its interaction with genetic components can be subtle; for example, a single locus of lipoprotein A appears to identify patients at risk of aortic and mitral valve calcification[145]. Investigations regarding atherosclerosis have focused on various aspects of the disease to improve risk assessment for cardiovascular events, studying biomarkers related to the onset and progression of atherosclerosis, or applying experimental methods to investigate underlying disease mechanisms. Modelling in systems biology is also particularly active[146,147]. From a mechanical perspective, it is well known that the development of atherosclerotic plaque is most prevalent in regions of low shear stress. Computational investigations have considered certain aspects of the development of atherosclerosis connected to specific haemodynamic conditions. Their aim is to study possible hypotheses regarding the main processes of arterial pathogenesis. These models often use non-linear reaction-diffusion equations describing the transport and reaction of various species involved in the process[148]. Recently, a first version of a platform-based prediction of atherosclerosis was published[149]. It applies diffusion-reaction equations based on a patient-specific reconstruction of arterial segments and predicts plaque growth.

However, time constraints – i.e. the disease may need a long time to develop – make mechanistic modelling difficult, and that is where statistical modelling often comes into play. There is a plethora of epidemiological studies and statistical modelling to predict risk linked to progression; for example, a recent study showed that sedentary participants had a 22%

[144] Delbaere K, Van den Noortgate N, Bourgois J, Vanderstraeten G, Tine W, Cambier D: The Physical Performance Test as a predictor of frequent fallers: a prospective community-based cohort study. Clinical rehabilitation 2006, 20(1):83-90

[145] Thanassoulis G, et al.2013.. http://muhc.ca/sites/default/files/nejm%20pre-publication%20copy.pdf

[146] Ramsey SA, Gold ES, Aderem A. A systems biology approach to understanding atherosclerosis. EMBO Mol Med. 2010 Mar;2(3):79-89. doi: 10.1002/emmm.201000063.

[147] Huan T, Zhang B, Wang Z, Joehanes R, Zhu J, Johnson AD, Ying S, Munson PJ, Raghavachari N, Wang R, Liu P, Courchesne P, Hwang SJ, Assimes TL, McPherson R, Samani NJ, Schunkert H; Coronary ARteryDIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) Consortium, International Consortium for Blood Pressure GWAS (ICBP), Meng Q, Suver C, O'Donnell CJ, Derry J, Yang X, Levy D. A Systems Biology Framework Identifies Molecular Underpinnings of Coronary Heart Disease. Arterioscler Thromb Vasc Biol. 2013 Mar 28. [Epub ahead of print]

[148] Di Tomaso, Diaz-Zuccarini, Pichardo-Almarza, IEEE transactions in biomedical Engineering, 2011

[149] Siogkas et al. In 2011

increased carotid atherosclerosis progression compared to active counterparts [150], and it is noteworthy that the statistical/epidemiological studies are the ones informing healthcare policy makers[151]. *Modelling should make use of all data available by using the best modelling paradigms fit for each purpose* and it is the *integration of these that will allow making substantial progress*. An interesting idea has been presented in[152], where hybrid mechanistic/data-driven approaches are proposed in order to overcome some of the limitations of mechanistic models via the use of machine learning (and vice-versa). The proposed framework attempts to develop a modelling workflow in which, instead of learning in the space of data, intelligent machines will learn in the space of mechanistic models. It is noteworthy that much of the data on atherosclerosis come from autopsies, since control data from a healthy population over long periods of time prove difficult to obtain. It would be ideal to consider phenomenological models as a way to augment the imaging information content. The challenge is to characterize the role of personalised modelling and how to integrate physiological, environmental, and lifestyle data.

3.3.4 Cardiomyopathy

Modelling techniques have been used with success in describing human anatomy, physiology, and disease. The use of novel technologies harnessing the power of mathematical models has progressed towards predictive cardio-patho/physiology from patient-specific measurements, for example[153] in order to improve diagnosis, treatment planning and delivery, and optimization of implantable devices by making cardiac models patient-specific using clinical measurements. Advanced cardiac models, for example[154], have been used as a starting point and used state-of-the-art clinical imaging to develop new and personalized models of individual cardiac physiology. There are interesting and promising results in this area, ranging from arrhythmias to myocardial deformation, cardiac wall motion, and patient-specific tissue information such as myocardial scar location[155,156,157,158].

Current genome technologies may enable insights into personal behaviour and stress conditions that produce changes in DNA methylation in different tissues, like the heart[159].

---

[150] Palatini P, Puato M, Rattazzi M, Pauletto P. Effect of regular physical activity on carotid intima-media thickness. Results from a 6-year prospective study in the early stage of hypertension. Blood Press. 2011 Feb;20(1):37-44.

[151] Consensus Report of the European Atherosclerosis Society. Lipoprotein(a) as a cardiovascular risk factor: current status. Eur Heart J (2010) doi: 10.1093/eurheartj/ehq386

[152] O.M. Doyle, K. Tsaneva-Atansaova, J. Harte, P. A. Tiffin, P. Tino and V. Diaz-Zuccarini, "Bridging Paradigms: Hybrid Mechanistic-Discriminative Predictive Models.", IEEE Transactions on Biomedical Engineering

[153] www.euheart.eu

[154] Hunter, P., et al., Mech Ageing Dev, 2005. 126(1): 187-92, Smith, N., et al., J Exp Biol, 2007. 210: 1576-1583

[155] M. W. Krueger et al. *Towards Personalized Clinical in-silico Modeling of Atrial Anatomy and Electrophysiology, Medical & Biological Engineering & Computing, Springer, 2012, in press*

[156] J. Weese et al, *Generating Anatomical Models of the Heart and the Aorta from Medical Images for Personalized Physiological Simulations*, Medical and Biological Engineering and Computing, Springer, 2013, in press

[157] S.A. Gaeta, T. Krogh-Madsen, and D.J. Christini. Feedback-control induced pattern formation in cardiac myocytes: A mathematical modeling study. J. Theor. Biol. 266:408-418, 2010.

[158] Li W, Kohl P & Trayanova N. Induction of ventricular arrhythmias following a mechanical impact: a simulation study in 3D. Journal of Molecular Histology 2004/35:679-686.

[159] see for example Movassagh, M., Choy, M. K., Knowles, D. A., Cordeddu, L., Haider, S., Down, T., Lio, P, Foo, R. S. (2011, November 29). Distinct epigenomic features in end-stage failing human hearts. Circulation, 124(22), 2411-2422. doi:10.1161/CIRCULATIONAHA.111.040071; Haider, S., Cordeddu, L., Robinson, E., Movassagh, M., Siggens, L., Vujic, A.,Lio', P,. Foo, R. (2012, October 3). The landscape of DNA repeat elements in human heart failure. Genome Biol, 13(10), R90. doi:10.1186/gb-2012-13-10-r90

Challenges in phenomenological modelling could also look at medical and surgical interventions (for instance stents) and disease early predictors, as exemplified in[160] and making use of signal-based analyses. One challenge would be to consider phenomenological models to include comorbidities lie diabetes, as well as prior knowledge such as medication or medication history. This is reflected by the finding that patients that are scheduled for a peripheral artery intervention do much better if they are already on statins and aspirin[161]. Regular aspirin use is associated with an elevated risk for neovascular age-related macular degeneration[162]. Other factors that are difficult to include in a mechanistic assessment are, for example, ethnicity, gender, and lifestyle.

3.3.5 Dementia

Dementia is not a single disease, but is rather an umbrella syndrome that includes many different forms[163]. All neurodegenerative diseases share a number of common distinctive pathological hallmarks, such as extensive neuronal death and clinical symptoms like compromised function in the affected brain regions. Although in many cases few proteins are found to have significantly different concentrations between healthy and diseased neurons, the basic mechanism of dementia is still unclear[164]. Effective pharmaceutical treatment of dementia is currently not available.

Mechanistic models can provide an essential and much needed platform for improved understanding of dementia.. A clear exemplar is the case of vascular dementia, in which the use of a patient's anatomical and physiological characteristics and mechanistic models of plaque progression could lead to the development of powerful tools to help to elucidate the relationship between progression of disease in time, and cognitive impairment. There is also much to gain in better capturing the mechanistic complexities of the blood-brain barrier and its relationship to neural behaviour. Another compelling case is the role of detailed modelling and analysis of the microvasculature and its relationship with stroke and Alzheimer's disease, which has been recently addressed by [165].

Mechanistic models based on molecular data are, however, challenged by results from epidemiological studies that point to lifestyle factors, such as poor diet and physical and cognitive inactivity. This is an area where phenomenological models could consider social parameters which are difficult to incorporate in mechanistic contexts; for example, (1) as a person-centric model highlighting the context of a patient's significant relationships; (2) as a disability approach, according to which people with dementia are people with cognitive disabilities; or (3) as a medical approach, in which people with dementia have a neurological disease. Lifestyle is very important, as shown by the finding that individuals with the highest

[160] Hock Ong ME, Lee Ng CH, Goh K, Liu N, Koh ZX, Shahidah N, Zhang TT, Fook-Chong S, Lin Z. Prediction of cardiac arrest in critically ill patients presenting to the emergency department using a machine learning score incorporating heart rate variability compared with the modified early warning score. Crit Care. 2012 Jun 21;16(3):R108. [Epub ahead of print].

[161] Ardati A, et al "The quality and impact of risk factor control in patients with stable claudication presenting for peripheral vascular interventions" Circ Cardiovasc Interv 2012; DOI: 10.1161/CIRCINTERVENTIONS.112.975862.

[162] Liew G, et al "The association of aspirin use with age-related macular degeneration" JAMA Intern Med 2013; DOI: 10.1001/jamainternmed.2013.1583

[163] http://www.webmd.com/alzheimers/guide/alzheimers-dementia

[164] Lage K, Karlberg EO, Størling ZM, Olason PI, Pedersen AG, Rigina O, Hinsby AM, Tümer Z, Pociot F, Tommerup N, Moreau Y, Brunak S. A human phenome-interactome network of protein complexes implicated in genetic disorders. Nat Biotechnol. 2007 Mar;25(3):309-16

[165] Chang Sub Park and Stephen J. Payne. A generalized mathematical framework for estimating the residue function for arbitrary vascular networks. Interface Focus 6 April 2013 vol. 3 no. 2 20120078.

levels of cardiorespiratory fitness during middle age were significantly less likely to develop dementia in their senior years[166]. Thus, introduction of physical activity can reduce the risk of cognitive impairment in old age. Metabolic syndrome and diabetes are also associated to dementia[167].

3.3.6 Stroke

Stroke is yet another case in which modelling could offer much needed help. The rapid loss of brain function due to disturbances in the blood supply to the brain can be due to ischemia (lack of blood flow) caused by blockage (thrombosis, arterial embolism) or by a haemorrhage, which are all suited for rigorous analysis via *in silico* tools[168]. Stroke is one of the leading causes of death and acquired handicap. There is already work in the literature[169] developing physiology-based models for acute ischemic stroke. This is a case where most of the clinical trials have failed, contrasting with promising results during preclinical stages. This continuing discrepancy suggests some misconceptions in the understanding of acute ischemic stroke, and this is where modelling techniques can provide assistance for understanding its underlying mechanisms. One possible method for identifying the shortcomings of present-day approaches is to integrate all relevant knowledge into a single mathematical model and to subject that model to challenges via simulations with available experimental data.

Several phenomenological models have been proposed that account for stopping of the blood flow in some part of the brain (ischemia), reduced oxygen levels, and damage to cells[170]. Recent models have focused on studying the influence of spreading depression on the death of the cells; this is like a transient suppression of all neuronal activities spreading slowly across a large region of the brain[171]. Future models may take age into account: almost half of children with haemorrhagic stroke had seizures at presentation or within a week of onset[172]. Also lifestyle plays a key role. An exercise program such as tai chi that focuses specifically on balance was found to reduce the incidence of falls[173]. Phenomenological models should consider nutrition and cooking methods; for example, diets that are heavy in fried and salty foods could be the most dangerous in terms of stroke risk.

---

166 DeFina L, et al "The association between midlife cardiorespiratory fitness levels and later-life dementia: A cohort study" Annals Intern Med 2013

167 Frisardi V, Solfrizzi V, Seripa D, Capurso C, Santamato A, Sancarlo D, Vendemiale G, Pilotto A, Panza F. Metabolic-cognitive syndrome: a cross-talk between metabolic syndrome and Alzheimer's disease. Ageing Res Rev. 2010 Oct;9(4):399-417. doi: 10.1016/j.arr.2010.04.007. Epub 2010 May 2

168 Sims NR, Muyderman H (September 2009). Mitochondria, oxidative metabolism and cell death in stroke. Biochimica et Biophysica Acta 1802 (1): 80–91. doi:10.1016/j.bbadis.2009.09.003. PMID 19751827.

169 Duval V, Chabaud S, Girard P, Cucherat M, Hommel M, Boissel JP. Physiologically based model of acute ischemic stroke. J Cereb Blood Flow Metab. 2002 Aug;22(8):1010-8.

170 Duval, V.; Chabaud, S.; Girard, P.; Cucherat, M.; Hommel, M.; Boissel, J.P., 2002. Physiologically based model of acute ischemic stroke. Journal of Cerebral Blood Flow and Metabolism, 22, 1010-1018; Dronne, M.A.; Boissel, J.P.; Grenier, E.; Gilquin, H.; Cucherat, M.; Hommel, M.; Barbier, E.; Bricca, G., 2004. Mathematical modelling of an ischemic stroke: an integrative approach. Acta Biotheoritica, 52, 255-272

171 G. Chapuisat, M.A. Dronne, E. Grenier, M. Hommel, H. Gilquin, J.P. Boissel A global phenomenological model of ischemic stroke with stress on spreading depressions Prog Biophys Mol Biol. 2008 May;97(1):4-27. Epub 2007 Nov 1

172 Beslow L, et al "Pediatric intracerebral hemorrhage: Acute symptomatic seizures and epilepsy" JAMA Neurol 2013; DOI: 10.1001/jamaneurol.2013.1033

173 Taylor-Piliae, RE, et al "Stroke survivors in a 12-week Yang-style tai chi intervention have fewer falls" ISC 2013; Presentation W P362.

3.4 Timeline and impact

In this section, we try to summarise the long-, mid-, and short-term challenges for modelling in the Digital Patient framework. The challenges are ranked according to the developments that are required to meet them (short- / mid- / long-term) as well as by their impact (benefit for patients).

Short-term challenges

- Formalisation of model testing and validation strategies, determining how selection of testing strategies should be made independent of model development.
- To immediately strengthen collaboration between modellers and clinicians, and improve uptake and testing of models despite an on-going development process. A recommendation in this respect is to call for small focused projects that address early stages of the disease modelling process with mixed teams with the goal of early testing in small cohorts of patients.
- Encourage the development of hybrid paradigms in order to capitalise on the potential of modelling as a whole for personalised medicine.
- Development of relatively simple models (see examples provided in previous section for cardiovascular diseases and cancer) that address specific topics in patient studies, for the expansion of diagnostic methods and therapies in the clinic.
- Expansion of the set of models that can be applied clinically, with existing models applied in particular areas of diagnostics: models describing a small part of physiology, with a limited number of inputs and outputs, directed towards a specific disease or diagnostic method.

Mid term challenges

- Creation of online repositories to house disease- and patient-specific data, through which mechanistic model inputs may be linked to patient lifestyle factors (age, fitness, diet, etc.)
- Development of mechanistic models as tools to integrate data into structures that enable computation of their implications.
- Development and validation of customized models for specific pathologies, with patient-specific inputs and outputs.
- Development of hybrid strategies for the combined use of phenomenological and mechanistic models
- Development of mathematical formalisms for multi-scale processes
- Automatic debugging and systematic testing tools for patient-specific models, possibly in combination of machine learning techniques and artificial intelligence

Long-term challenges

- Combination of specific models into a large-scale patient model encompassing larger, multifactorial pathologies such as heart failure, renal failure, etc.
- Personalise not only anatomical data but also the physiological/pathological processes taking place (multiscale) by linking model parameters to easily obtainable

patient data, leading to an individual patient model rather than a statistical patient model.

3.5 Summary and conclusions

Clinical utilization of models: reasons for optimism

In this chapter we have made an attempt to cover the mathematical modelling challenges that scientists and technologists will need to face in the short-, mid- and long-term to enable the realisation of the Digital Patient. The first one to be addressed is the *Selection of the most adequate mathematical modelling approach*. There is not a unique way of creating a model and a categorisation typically used to classify the mathematical models by the scientific community is distinguishing between phenomenological and mechanistic models. The first are built purely based on empirical observations of a phenomenon, while the second aim to represent the underlying mechanisms of a biomedical process. Making these models widely available in online semantically annotated repositories should enable the development of hybrid approaches able to customise and combine (even automatically) phenomenological and mechanistic models for use in the Digital Patient. The creation of such models also encompasses the challenge of *Personalising and extending them to cover multiple scales*, and including ageing is a representative example of this difficulty. Models are created using *real (and non-ideal) biomedical data sources* that commonly include incomplete entries, available in repositories that are non-standard, difficult to access or that lack interoperability features. This complicates even further the modelling process, because it needs to handle the uncertainty introduced by the often-incomplete input data and estimated parameters. As a result, any future model should come together with an estimated valid range of operation and a measure of confidence on the results.

Nevertheless, as already mentioned during the discussion of the *Translation and Clinical utilisation of models* challenge, despite the complexity of mathematical models of bodily functions, some have already **entered the clinical arena** in the form of software applications embedded in diagnostic or therapeutic devices. This indicates that in the short-term, **simple models** that cover a relatively small part of a pathology or process are *most likely to be adopted for clinical use* early on. Prior to actual clinical adoption on a larger scale, another key challenge especially relevant in the mid- and long-term is *Automating, generalising and formalising the process of model testing and validation*.

Concrete recommendations include the following:

a) Creation of online repositories to house and share disease- and patient-specific data and models to enhance collaboration within the VPH community, providing ubiquitous access

b) Develop hybrid methods and strategies to automatically and seamlessly combine phenomenological and mechanistic models, exploiting the use of VPH ontologies and annotated online repositories containing well-documented and validated models

c) Develop surrogate modelling methods that make possible to replace computational demanding sub-models (typically large PDE models) with estimators developed on pre-computed solutions, to provide a fast estimate of the model outputs and an upper boundary of the estimation error

d) Develop integrative modelling frameworks that support the abduction cycle that applies inductive reasoning to observations to generate hypotheses on mechanistic relationships, verify these against reference observations, and where predictions are

in good agreement with observations, incorporate this new mechanistic understanding into the inductive reasoning, so facilitating new discovers

e) Develop fast numerical restart methods that make it possible to employ user exploration of the information space to re-run the model with different inputs at very low computational cost when compared to the first run

f) Develop a theoretical framework for the analysis of scale separation, and general homogenisation and distribution strategies to define space-time relations across scales

g) Develop strategies to formalise and generalise the testing and validation of mathematical models, providing accurate and automatic estimations on the impact that incomplete data has in the personalised models