A Vision and Strategy for the VPH in 2010 and beyond

Contents

1. The VPH vision

The Physiome, Systems Biology, the Virtual Physiological Human, Personal Health Systems, Biomedical Informatics, Life Science e-Infrastructures, Systems Pharmacology. All of these domains share one common issue: the need for integration. To implement biomedical research outputs into clinical practice and healthcare industries we need to integrate data, information, knowledge and wisdom. We need to integrate data of the same patient stored in different hospitals in different member states or in clinical research databases; we need to integrate the information related to various parts and processes of the human body into a systemic understanding of pathophysiology; we need to integrate the knowledge digitally captured into metadata, ontologies and models in order to respond to the combinatorial explosion of cognitive complexity that integrative research is producing; and we need to integrate the wisdom produced in the research laboratories and in clinical practice, which will be formalised in guidelines, standards and protocols and used to promote translation of basic science and integrative models into healthcare benefits.

This is a huge challenge that, if met, will have a tremendous impact on the life of our citizens, and on the European and international economy. The vision of a "digital me" that contains all my healthcare information, safely managed for access by the various biomedical professionals with my approval, communicated with all my wearable and implanted technology to constantly monitor my health status and informing me, my family and friends, or my healthcare providers of alarming events, supporting the collaboration of various specialists around my complex systemic diseases, and used with all my data to predict the future development of my health in order to facilitate disease prevention and a fully self-aware lifestyle, is a powerful vision. But the challenges are huge.

The Virtual Physiological Human (VPH) FP7 project is addressing this challenge by promoting and facilitating the use of computational models, software tools and web services. The goal is to achieve a more efficient and effective 21st century healthcare system and to create new economic opportunities for European healthcare industries. In common with other areas of application of modern scientific methods, medical practice will benefit from technologies in which digital data enables predictable outcomes through quantitative models that integrate physical processes across spatial scales down to the molecular level. We need to promote personalised, predictive, integrative and evidence-based approaches to medicine. These will use computational tools to link individual patient data with virtual population databases via the knowledge of biological processes encoded in mathematical models. The biomedical community also now has the opportunity, thanks to the adoption of new model and data standards and a common set of reference ontologies, to assemble the molecular pieces from 50 years of reductionist science in order to understand genotype-phenotype relationships by linking databases of genetic and proteomic data to anatomy and function at the cell, tissue and organ levels. Biophysically based computational modelling of the human body, applied to human physiology and the diagnosis and treatment of disease, will revolutionise 21st century bio-sciences and medicine. The success of this exciting opportunity is highly dependent both on the development, adoption and integration of ICT and eHealth infrastructures throughout Europe, and on the coordination of this effort with other related international initiatives such as the Physiome Project.

Note that the NoE is concerned primarily with ICT infrastructure, coordination and training for the VPH and the VPH-I projects themselves are primarily focussed on developing and implementing biophysically based computational models into clinical environments via industrial partners. The success of all of these endeavours is of course dependent on the continued progress of (separately funded) biomedical science in revealing the biophysical mechanisms underlying structure and function at all spatial scales.

A Roadmap for the VPH project was laid out in 2006 by the STEP coordinated action [1]. The outcome of the first FP7 VPH funding round in 2007 (Call 2) was the VPH Network of Excellence (NoE), three Integrated Projects (IPs), nine Specific Targeted Research Projects (STREPs) and two Cooperative Actions (CAs), all of which form the core of the European VPH Initiative (VPH-I). With

nearly three years of experience behind us (including STEP), it is time to assess our achievements and plan for the short, medium and long term future of the VPH and the NoE.

2. What have we achieved so far?

The importance of establishing a solid foundation for the VPH by creating model and data standards, together with mechanisms for achieving model reproducibility and reuse, was recognised in the STEP Roadmap. This, together with the development of plans for dissemination, training and outreach to the communities of researchers, physicians, patients, students, European industry and the public in general, has been the primary focus for the first year of the NoE. Direct engagement with the other VPH projects and clear examples of how standards-based models, software tools and web based services can be used to facilitate clinical outcomes are the next top priority target. These goals are discussed below, along with proposals for the additional resources and engagement needed to establish digital, personalised, and predictive medicine in Europe.

Standards, tools and services
The first stage of the NoE project has been largely concerned with establishing standards for models and data, building model and data repositories for published models, and assembling a toolbox of existing software programmes (many, but not all, open source) that are relevant to the VPH. A key role for the NoE has been writing the Application Programming Interfaces (APIs) for the markup languages that allow the application software packages to read models and data from the repositories. These developments are described in detail in the accompanying report. The markup languages provide the *syntax* (the grammar) for encoding models and data. Equally important are the *semantic* 'metadata' that give biological and biophysical meaning to the models and data via biological and biophysical *ontologies* (structured vocabularies). Mapping a number of existing ontologies onto the modelling framework is therefore another important thread in the immediate future, since it will enable unambiguous links to gene and protein databases. Note that considerable progress has been made over the last 12 months in coordinating the development of VPH standards such as CellML and FieldML with those of the systems biology community, in particular SBML.

Dissemination, training and outreach
Many communities will be impacted by the VPH, including: in the short term, biomedical researchers and students; in the medium term, healthcare workers and European industry; and, in the longer term, patients and the general public. Dissemination and training is therefore a major responsibility and one that has to be tailored separately for these various audiences. The initial focus for the NoE has been establishing a website containing full descriptions of the VPH projects and giving access to the VPH modelling and data resources and training programs, suitable for the first target community. The NoE website is already receiving 13,000 visits per month and the newsletters, which are published at 6 monthly intervals, have been extremely well received. There is clearly a large and growing interest in the VPH, and this needs to be addressed now in the way outreach, lobbying, and dissemination are handled.

Training is considered a fundamental part of the dissemination strategy, where the VPH NoE will target young and experienced researchers alike. Training has been identified not only as a key development area for the next 12 months, but also as one that has the potential to help sustain the VPH initiative in the long term. VPH training should be embedded in basic and applied research methods courses for higher degree students and as part of informatics education and training for the health workforce

Other efforts are contributing to the formation of the VPH community: the BiomedTown on-line community, which hosted the consensus process of the STEP action, has been constantly growing since then and now has nearly 2000 members. Also, the larger VPH projects are contributing to the dissemination of the VPH vision. In addition to their project web sites, projects like VPHOP or euHEART are publishing periodic newsletters that reach thousands of stakeholders worldwide.

International linkages
Internationally the WIRI agreement [2] and the Osaka Accord [3] have established a worldwide agenda for physiome research under the patronage of the European VPH initiative and the IUPS Physiome Project. Four other important events in the last 12 months have been the participation of a European delegation at the IMAG[1] symposium in Montreal in 2008; the CellML/SBML [4] combined meeting in Auckland, NZ, in April 2009; the *Virtual Tissue* conference organised by the USA Environmental Protection Agency in Spring 2009; and the satellite symposium at the IUPS 2009 conference in Kyoto. All these events have been of considerable political relevance, and have strengthened the role of the European VPH community on the international research scene. Note that many of the VPH-I projects have international partners and the NoE itself has 'International General Members'. This formal recognition of international membership is also important for VPH-linked co-funding arrangements in countries outside Europe.

VPH-I projects
The goals of the current 14 VPH projects are summarised briefly below. Although these projects have been running now for only 18 months, it is quite clear that most consortia are on track to deliver what they promised and some preliminary results are already emerging. There are major technological achievements in various areas, including: data collection, management and integration; processing and curation of data into information; reductionist and integrative modelling of pathophysiological processes; presentation, deployment and end-user applications. It is also notable that there is already an active involvement of companies participating in the VPH consortia, both SME and large corporations, and that the involvement is rapidly moving from their R&D departments to their strategic management as the first business scenarios emerge. The clinical partners are providing a vital contribution to many of the VPH projects, participating enthusiastically and with considerable commitment. Note that the NoE clinical advisory board will be playing a more active role from now on and will provide a mechanism for generalising the lessons learned from the clinical partners of the individual VPH-I projects.

The initial (Call 2) VPH projects are targeted as follows:
euHeart (IP) is a developing open source codes and multiscale/multi-physics models of heart electromechanics in clinical cardiac diagnostic and device development applications.
VPH2 (STREP) is developing decision support tools for heart disease.
preDiCT (STREP) is developing models of cardiac electrophysiology for drug design and toxicity testing.
ARTreat (IP) is developing an interventional decision support system for stenting procedures based on multiscale patient specific models of atherosclerotic disease.
ARCH (STREP) is developing clinical decision support tools based on patient-specific predictive modelling of vascular pathologies.
PASSPORT is developing an open source multiscale framework for diagnostics and surgical training in the liver, based on modelling liver cell regeneration.
IMPPACT (STREP) is developing minimally invasive, patient-specific treatment strategies for liver cancer based on bioengineering multiscale modelling principles.
PredictAD (STREP) is developing an evidence based statistical framework for diagnosis of Alzheimer's disease.
NeoMARK (STREP) is implementing collaborative research networks and tools for the early detection of oral squamous cell carcinoma.
CONTRACANCRUM (STREP) is using multiscale modelling techniques to simulate patient specific cancer treatment outcomes.

---

[1] The Interagency Modelling and Analysis Group (IMAG) coordinates multi-scale modelling initiatives from various United States agencies including the National Institutes of Health, National Science Foundation, National Aeronautics and Space Administration, Department of Energy, Department of Defence, United States Department of Agriculture, and Unites States Department of Veteran Affairs.

VPHOP (IP) is developing a patient-specific, multiscale modelling framework for predicting osteoporotic fracture in elderly patients.

HAMAM (STREP) is establishing a database of curated and annotated imaging data and software tools for breast cancer diagnosis.

ACTION-Grid (CA) is promoting collaboration in medical/biomedical Informatics and grid technologies to promote the interface between ICT and nanotechnology.

RADICAL (CA) is investigating security and privacy issues for VPH applications and best practices for medical and genetic data protection in distributed environments.

Nearly all of these projects deal with challenges relating to patient-specific, multiscale modelling and the implementation of models and software in clinical environments.

DEISA VPH virtual community

The NoE has also taken steps to build a relationship between the NoE and the VPH-I research projects through the development of a DEISA VPH Virtual Community. The Virtual Community, applied for and managed by the NoE on behalf of the VPH-I, provides access to high performance computing facilites for any VPH-I research project which requires such a facility; currently, over 50% of the projects are being supported in this manner. Moreover, additional EU funded projects working in e-Health related domains are also being supported.

3. New initiatives

Much of the work currently underway in the NoE will continue and expand in scope over the next 12 months, but here we describe a number of new initiatives that are currently underway.

Workflows and driving clinical problems

Over the next 12 months the NoE will work closely with the other VPH-I projects to define how the models, data and tools are used in biomedical and clinical *workflows* – the sequence of operations that start, for example, with clinical data (complying with standards relevant for privacy, security and ethics) and, by using the tools and models, end with a clinically useful diagnostic index or treatment strategy. Six cases studies are presented in the Appendix to illustrate workflows currently being developed in a number of VPH-I projects. In some cases this will require resources to support the real-time computational needs of medical diagnoses, and the NoE is therefore currently trialling high performance grid-based compute services.

Training and dissemination

Another essential and urgent step is the implementation of workshops and summer schools to train people in the use of the VPH models and software. Such activities will form an important part of wider, pan-European process directed towards the introduction of systematic educational activities with the aim of ensuring that academia, medicine, and industry throughout Europe have a workforce that is appropriately equipped to meet the possibilities offered by this new and important discipline.

Key to this initiative will be the *VPH Virtual Academy*. This web-based facility, accessed by a portal via the VPH NoE website, will provide a focus for interaction between training providers, course developers, young researchers wishing to develop a career in physiological modelling, established researchers seeking training as part of a commitment to Life-Long Learning, and representatives from the major employment sectors. VPH training will need to be responsive to the changing needs of employers, and the Virtual Academy will provide an environment to engage with and obtain feedback from industry, healthcare, and professional bodies.

The aim for the next twelve month period is to foster a training community by involving a critical mass of students from institutions across Europe. In this initial phase, the Academy will be directed towards gathering information on training needs, promoting VPH-related training activities and courses, and providing a point of contact between potential employers, training providers, and newly qualified researchers. The BiomedTown VPH portal, using Web 2.0 technologies and approaches could be used to engage and build communities.

RICORDO project

A recently funded VPH project called RICORDO will research and prototype a communal ontology-based annotation and repository communication strategy that supports the interoperability of VPH data and models across different biological scales. A new community standard for the representation of multiscale biological entities will be developed and integrated into the VPH Toolkit development strategy. The medical relevance of such an advance resides in (i) the use of a coherent multiscale anatomy standard[2] to refer to any body structure or location regardless of scale (an aspect that is also of relevance to electronic health record keeping and Clinical Document Architecture standards), and (ii) the demonstration of an interoperability plan to connect patient specific radiological images to mathematical models of physiology, as well as to disease-related genomic and molecular data.

MSV project

Another recently funded VPH project called MSV (*Multi-Scale Visualisation*) will provide a formal conceptual taxonomy for multiscale spatiotemporal visualisation for both spatial multiscale and temporal multiscale problems, sometimes involving multiple datasets. It will also develop better visualisation algorithms for handling large multiscale data sets and implement these in an open source software library.

4. Sustaining the VPH

Fourteen new projects related to the VPH were launched in Call 2; five more were funded in Call 4, and many others will be funded in Call 6. Currently, the only coordination of these related endeavours is via the VPH NoE, which is also pursuing its own specific goals such as the VPH toolkit. In order to transform the VPH vision into a reality for European industries and healthcare services we need a long-term coordination action in order to:

- coherently strategise and periodically revise the concrete research and technological development goals that should make the vision come true;
- sustain the standardisation and interoperability efforts;
- sustain the further development, maintenance and provision of tools, services, databases and other infrastructure for common use;
- monitor the development, adoption, and impact of VPH technologies;
- sustain the global adoption of VPH-based protocols that have proved effective;
- provide training and re-training in the use of VPH technologies.

All these activities cannot be properly scoped by the NoE or by any other initiative that is limited in time; they require the attention of a permanent organisation, capable of ensuring continuity over actions that may last for decades. This requires that the issue of sustainability of the VPH initiative be addressed. In our view, the best way to achieve this will be to establish a non-profit European 'VPH Institute' with a mandate to support the maintenance of VPH databases and the continued development of standards and business-friendly open source software.

A VPH Institute

There is a growing consensus among VPH stakeholders that in order to ensure long-term sustainability of the VPH initiative, it is necessary to establish a collective identity in the form of an international, non-governmental, not-for-profit organisation, hereafter referred to as the *VPH Institute*.

The scope of this institute will be to sustain the coordination of all academic, governmental, industrial, and societal stakeholders' efforts toward the common goal of developing an integrative biomedical science and technology that makes the VPH vision practically possible, effective, sustainable, and ethical.

---

[2] Discussed extensively by the VPH NoE http://www.vph-noe.eu/vph-repository/doc_download/48-vph-anatomy-framework

This general scope will be pursued according to a detailed strategic plan periodically revised by the Institute's governing body, which will include actions in specific contexts to:

- Provide a strategic think-tank capable of conducting road-mapping exercises and of developing strategic recommendations for the European Research System around the concept of integrative biomedicine.
- Support standardisation processes to ensure interoperability of integrative technologies and the integrability of data, information, knowledge, and wisdom captured in digital format.
- Foster the collaborative development, assessment, deployment, and maintenance of tools, services, and methods for biomedical integrative research and clinical practice.
- Manage, directly or through its member organisations, shared infrastructures to support the integrative approach in research and clinical practice. This includes repositories of data and models, benchmark collections, infrastructures for the collaborative development, etc.
- Conduct, under the mandate of governmental agencies and research-funding charities, studies to assess the status, the level of adoption, and the impact of integrative methods and technologies.
- Organise actions for positive lobbying and public information around the concept of integrative medicine.
- Provide for the heterogeneous collection of stakeholders that are involved with research, technological development, and clinical practice, a single voice for the citizens of Europe, their elected representatives, and any other institutional body that can sustain the development, assessment and deployment of the integrative biomedicine vision.
- Offer 'VPH-certified' courses for using the VPH tools and services.

The VPH Institute will be an organisation whose members are other legally established organisations. Membership will be open to any organisation whose legal status and statutes are compatible with the membership rules defined by the VPH Institute statute.

To start the statutory process, a *pro tempore* board will be formed by inviting a small number of VPH researchers that have been involved with the VPH initiative from the very beginning, are recognised authorities by the community, and are willing to allocate significant time to the statutory process. The only duty of this pro tempore board will be the drafting of the statute on which the VPH institute will be incorporated. As soon as the incorporation is completed, a membership campaign will be launched, the pro tempore board will be disbanded, and a new Board of Directors will be appointed according to the statutory procedures.

We expect to have the pro tempore board formed by the end of 2010, and the draft statute by the end of 2011, and the VPH institute fully operational in 2012, with at least one year of overlapping with some VPH projects funded in Call 2, including the VPH NoE, which will ensure a smooth handover.

VPH conference series
The ICT-Bio meeting [5] held in Brussels in October 2008 was the kick-off meeting for the VPH. The ICCB2009 meeting held in Bologna in September 2009 [6] was a second integrative biosciences VPH meeting. The NoE steering group are now planning a VPH conference series that will provide an annual event for the VPH community and also help to coordinate international physiome activities.

5. What are the biomedical science challenges?

The VPH project will achieve important outcomes within the lifetime of the current NoE by introducing computational modelling into the diagnosis and treatment of some diseases (particularly cardiovascular and orthopaedic diseases), but the real impact in the long term will be to transform European healthcare into a more personalised, predictive, and preventative process (see next section). The resources needed to achieve this long term goal must be realistically assessed and, in particular, we must now instigate projects to fill identified gaps in the necessary know-how and infrastructure.

Physiological databases

One current gap is the lack of comprehensive web-accessible databases of physiological data, encoded with well-established data and metadata standards. Such data provides numerical parameters for use in computational models. This need was expressed in section 3.2.3 of the VPH STEP roadmap. One standard, DICOM, does exist for medical image data. Others such as C3D [7] are well established binary data formats for specialist communities (biomechanics, animation and gait analysis in the case of 3CD). A more general purpose metadata standard (BioSignalML) is being developed for annotating physiological time-dependent signals encoded in a wide variety of existing specialist standards. But even this represents a small fraction of what is needed. A major effort is now needed by the physiology community to identify the types of physiological data that are available and to begin the development of a broad range of data standards and data repositories; as a first step, example datasets are being collected from the VPH-I and VPH NoE-Exemplar projects. The tools for interpreting these data are being developed by the VPH and Physiome Projects.

Genetics

The challenge of connecting modern genetics theory and methodology with multiscale physiological models, such that one can address and understand the genetics of complex diseases in a population context, defines a large and ambitious research topic that needs to be given specific attention in the coming years if we are to hope for tailored clinical treatments based on simulation studies of the genetic profile of the individual requiring treatment. Many challenges in personalized medicine reflect a lack of understanding of what is called the genotype-phenotype map , i.e. the aggregated phenotypic effects across different length and time scales of different constellations of genetic variation (genotypes). The biomedical genetics community is now facing serious challenges concerning the overall applicability of the genome wide association study (GWAS) approach when it comes to drug development and personalized medicine. The VPH initiative may be of substantial help to remedy this situation by offering a new R&D paradigm that has the potential to provide mechanistic model descriptions of the phenotypic effects of genetic variation. Such causally cohesive genotype-phenotype models are very advanced multiscale physiological models with an explicit link to genetic information and with the capacity to describe how genetic variation manifests in phenotypic variation at various systemic levels up to the tissue, organ and whole-organism level.

Clinical data

The re-use of clinical data remains a key challenge. As pointed out in the STEP roadmap, there is need for a "Global VPH security that makes possible the federation of clinical databases located behind hospital firewalls into the VPH framework" (VPH Research Roadmap section 12.1.5). In a recent report [8] Deloitte writes: "Most Life Sciences (LS) Research & Development functions are under increasing pressure to improve innovation, reduce development inefficiencies and advance product safety. Patient-level data, collected through Electronic Health Record (EHR) systems, offers one promising avenue for redefining Research & Development (R&D) and revolutionizing the LS value chain. Globally–aggregated, patient-level data could support the identification of disease mechanisms and new discovery areas, accelerate the termination of unsuccessful compounds, decrease patient recruitment cycle times for clinical trials, and improve drug safety surveillance through continuous monitoring".

Interoperability is the key to the effective re-use of clinical data. Currently, data exchange tends to be *ad hoc,* and no facility exists to support organised data exchange between multiple independent repositories (clinical, industrial and research). Candidate technologies capable of providing a data infrastructure that can facilitate VPH-wide data exploration, exchange and interoperability need to be explored and evaluated. A viable data infrastructure must support many activities (curation etc) giving data prospectors the freedom to revolutionise clinical procedures from the data they obtain, and yet issuing data providers with necessary assurances that their data will not be abused. Facilities for secure download/upload must be supported and data providers require further assurances that data users are appropriately authenticated and authorised to use the data. The VPH Institute is one mechanism by which formal authorisation procedures could be managed and policed whilst also

providing guidance on recommended procedures (eg. anonymisation) that are important to the clinic. The Institute is also well placed to be at the vanguard of standardisation efforts, since it will be astutely aware of the needs of the community and could recommend appropriate strategies that maximise opportunities for data interoperability. For instance, SNOMED CT is a comprehensive clinical terminology that is starting to become a standard in many clinical communities, including adoption as the language of the Care Records Service of the NHS in the UK, and is thus a target for adoption by the VPH initiative. This is but one aspect of a wider initiative, which includes such fundamental efforts as the harmonisation of nomenclature (modelling and clinical) through adoption of formalised ontologies, recommendations for data formats and data interchange (eg. DICOM) and suitable representation on international standards committees (e.g. HL7). The standards activity is necessarily multifaceted, and must recognise the competing demands of research, industry and the clinic. This is an opportunity for VPH advancement which would be sensitive to the dangers of over-regulation.

Finally, there are important legal issues to be tackled. By its very nature, the VPH crosses scientific and national boundaries. Differing interpretations of data protection law (eg. EU directive 95/46/EC) between member states discourage collaborative sharing of data for patient benefit. This is compounded by jurisdictional uncertainty due to a lack of legal precedents in this area. The ethical considerations relating to sharing of patient data are also formidable and could easily become a stumbling block to the progress of the VPH. The Institute would be a proponent for change, guiding and focusing effort in both areas, ensuring that these issues are given the priority they deserve in order to secure the vision of personalised and integrative medicine that is the goal of the virtual physiological human.

6. What are the healthcare challenges?

Major diseases like cancer, neurological and cardiovascular diseases are complex in nature involving environmental, life style, ageing and genetic components. A major challenge for the future how is to integrate the knowledge of all these different components into robust and fully reliable computer models and "in silico" environments that will help the development and testing of new therapies and better disease prediction and prevention tools in healthcare. The progressive advance in computing power and associated information technology offers the potential to deliver tailored clinical treatments based on simulation studies of the genetic profile of the individual requiring treatment.

The needs
The European healthcare system, including its biomedical research and technological development component, is a huge, complex, and highly articulated system. Due to the peculiar political history of the European Union, it is not a surprise that such a system is highly fragmented, not only between members states, but also between regions, districts, and even single hospitals. However, in spite of this extreme heterogeneity, there are some common requirements that are emerging in a number of analysis documents produced by very different sources [8-16]. Such requirements can be summarised in three keywords: *Personalised*, *Predictive*, and *Integrative* healthcare. A fourth keyword, *affordable*, is implicit, as the sustainability of healthcare systems is becoming the number one issue in a number of member states dealing with a constantly aging population.

More specifically, these common needs are: to maximise the yield of biomedical research expenditure; to achieve personalised healthcare for individuals and groups (women, children, etc); to improve the reliability, repeatability, and the timeliness of medical decisions; to integrate digital health information on a global scale; to resolve the individual-society conflict around the privacy of health data. It should be noted that at this stage these needs are very hard to quantify because the information is fragmented over dozens of reports produced by dozens of different medical specialties and much effort is required to elaborate into a single coherent framework a detailed and quantifiable description of needs. To address these issues it might be appropriate for the European Commission to consider funding a specific support action to collect, organise, and compose all these evidences into a fully justified and quantified needs analysis.

The 'Digital Me' vision

The vision we have is of a "DIGITAL ME", a coherent digital representation of each patient that is used as an integrative framework for the consolidation within the European research system of fundamental and translational Integrative Biomedical Research and the provision to European Citizens of an affordable Personalised, Predictive, and Integrative Medicine.

Personalised, predictive, and integrative healthcare

A new generation of medical technologies are needed to integrate the data available on a patient to support a more personalised diagnosis, prognosis, treatment planning, and monitoring, and to develop new drugs, therapies, medical devices, assistive, and diagnostic technologies that are optimised for specific groups of patients (age, gender, co-morbidity, etc). Diagnostic workflows are required, not on pre-defined general protocols, but on the prediction of risk obtained by models that combine both population and patient-specific information.

Ageing

Age is still the best predictor for most complex diseases. This implies that if we are to create individualised models describing the development and maintenance of complex disease situations on given genetic backgrounds in a way that is of broad practical utility for medicine, we have to incorporate the effects of ageing in multiscale and multiphysics models. By offering in silico representations of ageing tissues and organs upon which one can systematically study the effects of various chemical treatments, such individualised models would also define the foundation for a new and sorely needed drug targeting paradigm. The making of multiscale physiological models capturing the ageing process defines a daunting long-term theoretical-experimental research programme. But activity in this direction should be invoked in the coming years such that one can get a clearer view of what are the major challenges.

7. What are the ICT challenges?

The VPH-Physiome Project aims to provide a systematic framework for understanding physiological processes in the human body in terms of anatomical structure and biophysical mechanisms at multiple length and time scales. The importance of establishing a solid foundation for the VPH by creating model and data standards, together with mechanisms for achieving model reproducibility and reuse, was recognised in the STEP Roadmap. The framework includes modelling languages for encoding systems of differential-algebraic equations (CellML [17] and SBML [18]) and the spatially varying fields used with systems of partial differential equations (FieldML [19]). In both cases the parameters and variables in the mathematical models are annotated with metadata that provides the biological meaning. The languages encourage modularization and have import mechanisms for creating complex models from modular components. Model repositories have been established, together with freely available open source software tools to create, visualize and execute the models. The CellML repository [20] now includes models for a wide variety of subcellular processes.

The models and simulation codes that we develop, by their nature, require a complex IT ecosystem in order to run. Depending on the model in question, these can range from a desktop workstation to the very largest supercomputers in Europe and beyond. These machines run models that depend on data, so suitable data sources (such as clinical databases) need to be accessible from the compute infrastructure. The models also can potentially both use and generate vast quantities of data, meaning that suitable storage facilities need to be provided. If they are to be usable, all of these facilities must be connected by high performance network connections; the main requirement is that the entire distributed infrastructure available is integrated and can be accessed seamlessly, by clinicians and researchers alike. Access to many of these facilities requires that resource providers (e.g. computational grids and academic network providers) develop suitable policies to support the VPH integrated style of working, which involves unconventional access policies including advanced reservation and urgent, on demand access to resources.

Model and data encoding standards

The development of a biophysically based mathematical model is a creative endeavor, often requiring a great deal of insight into the physical processes being modeled and personal judgment about the approximations needed to satisfy Einstein's maxim that 'a model should be as simple as possible but no simpler'. Once created, however, a model should stand independent of its creator and be reproducible and testable by others. The model and data files that together can demonstrate reproducibility of a model on an automated basis, are called the *reference description* of the model. The issue of robustness and reproducibility is particularly important when a model representing some small component of physiological function is incorporated into a more comprehensive model – and especially one that is to be used in a clinical setting. To be worthy of reuse in this fashion, each independently developed component should be demonstrably 'correct' for the function it claims to represent, in the sense both of *biological validity* – it matches some aspect of biological reality - and *mathematical validity* – for example, it has consistent units and does not violate physical principles such as conservation of mass or charge.

The general strategy for developing the modeling standards is as follows:

1. Develop markup languages (MLs) for encoding models, including metadata, and data.
2. Develop application programming interfaces (APIs) based on the MLs.
3. Develop libraries of open source tools that can read and write the ML encoded files.
4. Develop data and model repositories based on MLs.
5. Develop reference descriptions to demonstrate model reproducibility.
6. Implement web services for a variety of tasks including access to automated scripts to run the models and compare results against experimental data, optimize parameter values for new experimental data and provide sensitivity analyses for changes in model parameters.

A useful way of viewing the development of standards is shown in Table 1, where progress in developing a specification of the minimal requirements for data, models and the simulation experiment are shown, along with the standards for the syntax of the data, models or simulation experiments and the ontologies for annotating the semantic meaning of terms in the data, models or simulation experiments.

	Data	Models	Simulation
Minimal requirements	Not available	MIRIAM [21]	MIASE [22]
Standard formats	PDB [23], DICOM [24] BioSignalML [25]	SBML, CellML FieldML	SED-ML [26]
Ontologies	GO [27], Biopax [28], FMA [29] SBO [30], OPB [31]	GO, Biopax, FMA SBO, OPB	KiSAO [32]

Table 1.The minimum information standards, syntax and semantics being developed for data, models and simulation experiments.

Note that the best example of an eHealth technology that is already in widespread use is the *Picture Archiving Communications Systems* (PACS) , usually based on use of the DICOM image encoding standard. Others close to maturity are *Electronic Transfer of Prescriptions* (ETP), *Computer based Patient Records and Electronic Medical Records* (CPR/EMR) and *Electronic Health Records* (EHT).

Model curation

There must be a concerted effort towards reproducibility, interoperability and the re-use of VPH models, including both future models and legacy, published models. This requires adoption of a consensus set of standards for the metadata that describe the models and of markup languages for their mathematical description. An integral aspect of interoperability will be the tagging of model variables and parameters with identifiers from reference ontologies such as the Foundational Model of Anatomy (FMA), Gene Ontology (GO), and appropriate ontologies for units, physics-based

quantities, physiological processes, etc). These must be adopted not only by the VPH community but also by the curators of the massive existing gene- and protein-databases in order to enable vertical multi-scale linking of models at the physiological scale (organs, tissues, cells) to the wealth of medically relevant molecular data.

Model curation is a long-term task that spans a wide spectrum of disciplines and will require a major effort but it is crucial to the success of the VPH vision. Sustainability of model repositories and software (including version control, archiving, technical upgrades, and provision for updating and expansion) will also be a major expense and limiting factor in the community acceptance of VPH models.

Top down or Bottom up?
The VPH of course needs to encompass both 'top down' and 'bottom up' approaches. A good example of top down is the PKPD (Pharmaco-Kinetic-Pharmaco-Dynamic) modelling community. PK deals with the advection, distribution, metabolism and excretion (ADME) of drugs and PD deals with the dynamics of how the drugs affect receptors. PKPD models accommodate human variability in an empirical fashion and treat body compartments with highly lumped approximations. The 'bottom up' approach of modelling biophysical mechanisms at the subcellular level is the realm of the systems biology community. The anatomically and biophysically based approaches of the VPH project are being designed to link these approaches. The model repositories based on the CellML and SBML standards already contain many models of both types.

The challenges of integrative multi-scale modelling
Biological systems are characterized by multiple space and time scales and there is a substantial need for new multiscale modeling algorithms to help bridge between the large range of spatial and temporal scales involved in the VPH. For instance, very often the macroscopic properties of a tissue, such as the bone stress-strain curve, or the diffusion coefficient of a drug or a chemical in the interstitial space of skeletal muscle, are related to microscopic effects. In this case, one needs to complement the macroscopic phenomenological description with a microscopic view of the problem. The basic idea is to model the theoretical input to a coarse-grained model from a more detailed microscopic model, bypassing the necessity of empirical description. As an example, homogenization and volume averaging methods can be used to obtain the strain energy function for bone tissue starting from basic information about a "reference" microscopic cell. Analogously, mass transfer in biomaterials or within a polymeric scaffold for tissue growth, can be effectively described at the macroscale by using macroscopic parameters that are obtained on the basis of a microscopic "cell problem" (i.e. on the reference elementary volume describing the scaffold). Finally, biological systems are often characterized by processes lasting a few seconds, such as metabolic biochemical reactions, and processes that fully develop only after days, months or years, such as atherosclerosis, angiogenesis, or the onset of an aneurysm.

Biological systems are also characterized by the interaction of many different physical processes at each spatial scale. For example, at the organ/tissue level, analysis of the beating heart couples large deformation mechanics of the myocardium with the reaction-diffusion equations governing the spread of electrical excitation, and with the equations of fluid mechanics for blood flow both within the ventricles and within the coronary vasculature. At the sub-cellular level the analysis of cardiac myocyte function requires the coupling of ion channel electro-physiology, calcium transport, myofilament mechanics, pH regulation and complex networks for signal transduction, metabolism and gene regulation.

The challenge of model reduction
With minimal information and model encoding standards now in place, it is time to address the extremely important question of 'model reduction' – how to automatically reduce the number of variables and parameters in a model when it is used under prescribed conditions. Some processes, for example, may be sufficiently fast in comparison with the time scale of events being modelled that they can be assumed to be instantaneous. Others may be approximated as being at steady

state. Some groups of reactions could be considered as a single module. In some case there may be substantial computational gains to be made by automated reduction in the number of parameters or variables associated with describing a spatial field. Thus a complex model that captures the biophysical and anatomical detail of some aspect of the physiology and anatomy of the body could be reduced to a simpler and therefore computationally more tractable problem under certain conditions.

The challenge of dealing with stochastic processes

Another modeling challenge, that has so far received relatively little attention, is that of incorporating stochastic behavior into the multiscale VPH models. At a molecular level stochastic behavior can be a reflection of Brownian (thermal) motion but at higher spatial scales it can be a reflection of unknown mechanisms – i.e. ignorance. It is very important that the consequences of this uncertainty in, for example, parameter values, be quantified (see Case Study #6).

The challenge of multiscale simulation and visualization software

Visualisation of the output of complex systems models and the human computer interface issues inherent in user interaction with such models is a new and significant area of research. Complex systems models are likely to have many input and output variables and may produce non-intuitive data representing, for example, emergent behaviours that are not easily represented by classical graphical or text-based methods. This is a rapidly moving area of cross-disciplinary research that may need specific funding under future calls for VPH project proposals.

The above challenges provide worthwhile and challenging problems for the mathematics and computational science communities and some VPH funding should be directly targeted to attract their expertise.

8. A strategy for the VPH

Timelines

A timeline for the STEP project (an FP6 initiative), the NoE, and other VPH projects funded under the first Call 2 of FP7 and the recently funded international Call 4 projects, is shown in Figure 1 together with the anticipated future calls.The establishment of a VPH Institute is also indicated.The anticipated impact that the VPH activities will have is also illustrated, first on biomedical research, then on industry and finally on the general public.
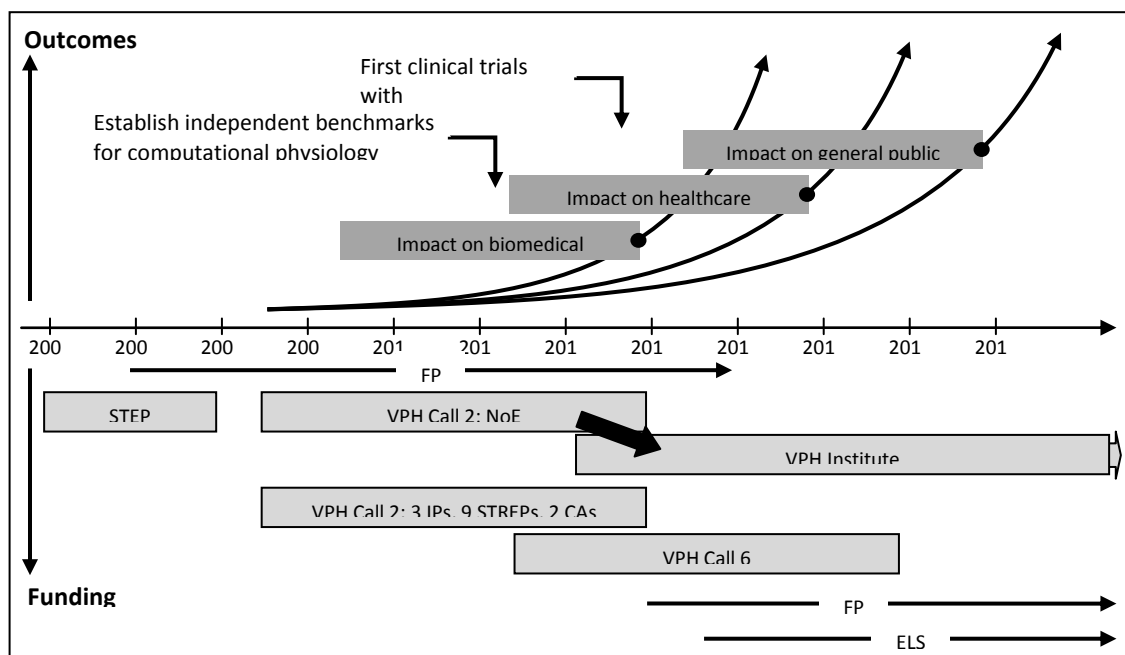
Figure 1. Timelines for VPH funding calls, the VPH projects and their impact on various communities. Sustainability of the NoE, and hence ongoing support for the VPH and the healthcare industries that depend on it, will be achieved through a VPH Institute.

The next steps
The VPH is a grand challenge. We propose the following specific actions.

2009-2010: Establish a collective identity
It is important that the multitude of players involved can speak with a single voice in a few strategic situations. This requires the creation of a collective identity around the VPH brand name. The VPH community is already working in this direction and will have a VPH institute fully operation in early 2011.

2010-2012: definition and quantification of needs
The STEP experience showed that when properly managed by a motivated consortium, and when embraced by a lively and motivated community, a road-mapping exercise can be of great value to capture and quantify needs and to develop a vision around them. In particular, we recommend three road-mapping actions:
- Road-mapping CSA[3] on VPH FET ('Future and Emerging Technologies' - Call 5)
- Road-mapping CSA in integrative health research
- Road-mapping CSA in health e-infrastructures

These actions should be sustained by other units of the European Commission (namely DGINFSO FET Proactive, DGRTD Systems Biology, and DGINFSO e-infrastructures as part of the Capacities program) and should involve significant portions of the traditional constituencies of these units, as it is necessary to include in the action a number of complementary expertises that are well represented in these constituencies.

Similar actions should be undertaken to push the VPH agenda as high as possible in those European institutions that fund fundamental research such as the ESF or the ERC. The general strategy to adopt toward these units and their constituencies should be inclusive, not invasive. We need to make clear that we do not want as VPH researchers to start designing e-infrastructures, running web-bench biology experiments, or develop fundamental research in computer science, mathematics, or physics. We recognise that there are neighbourhood communities that can do this much better than we can. What we offer is a common goal toward which all these skills and those we represent as VPH community can join forces. In the continuum of skills and interests, we need to find among VPH researchers those who are working closer to the fence with each of these communities, and support them as ambassadors, toward the formation of mixed consortia that can run these road-mapping exercises in a qualified and representative way. It is equally clear that in each of these neighbourhood communities we need to find the experts who are fascinated by the challenge we pose, and who are not afraid of the change that this would necessarily require to their research practice.

2011-2014: Disseminate and structure
As the results from these road-mapping exercises emerge, it will be necessary to coordinate and organise them into an operational plan to tackle this European Large Scale Action. This action will have to find substantial dedicated funding at the European level, but at the same time will be sustained and nurtured by a number of funding actions in the various neighbourhood domains that will take place as part of FP8. Another important dimension will be played by member state funding agencies that case by case will support horizontal or vertical initiatives. Preparing such composite and structured action requires in our opinion a specific coordination action, which will play a preparatory role in this direction.

2014-2019: European Large Scale Action on personalised, predictive, and integrative healthcare

---

[3] 'Coordinated support action'

These five years will make it possible to face this grand challenge at all levels (research, technological development, implementation, assessment, deployment) only if the necessary critical mass is reached in terms of skills, resources, and commitment of all stakeholders. We hope this document will serve as the first step toward the creation of such a critical mass.

Acknowledgments

References

1. www.europhysiome.org/roadmap
2. www.biomedtown.org/biomed_town/LHDL/Reception/lhpnews/wiri
3. www.biomedtown.org/biomed_town/VPH/wiri/OsakaAccord
4. www.cellml.org/community/events/workshop/2009
5. ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=3956
6. www.iccb2009.org
7. www.c3d.org
8. "Secondary uses of Electronic Health Record(EHR) data in Life Sciences", Deloitte Development LLC, 2009.
9. "The benefits from translating biomedical research into the health care system" Report to Bio21 Australia 2007
10. "Personalized health care: opportunities, pathways, resources" US DHHS 2007
11. "Pharma 2020: Virtual R&D", PWC 2008
12. Erik Brynjolfsson, "Investing In The IT That Makes A Competitive Difference", http://hbr.harvardbusiness.org/2008/07/investing-in-the-it-that-makes-a-competitive-difference/ar/1
13. Apoteket and Stockholm County Council, Sweden, "eRecept, an ePrescribing application ", http://ec.europa.eu/information_society/activities/health/docs/events/opendays2006/ehealth-impact-7-2.pdf
14. Ib Johansen, "E-Health and Implementation of EHR" http://www.ehealth-benchmarking.org/2006/images/stories/06_johansen_denmark.pdf
15. European Commission, Information Society, "eHealth is Worth it – the economic benefits of implemented eHealth solutions at ten European sites", http://ec.europa.eu/information_society/activities/health/docs/publications/ehealthimpactsept2006.pdf
16. Barry R. Hieb, Gartner, "Stop the Bleeding: Use IT to Achieve Sustained Value in Healthcare", http://www.gartner.com/DisplayDocument?doc_cd=144534

17. www.cellml.org
18. www.sbml.org
19. www.fieldml.org
20. www.cellml.org/models
21. www.ebi.ac.uk/miriam/
22. www.ebi.ac.uk/compneur-srv/miase
23. www.rcsb.org/pdb/home/home.do
24. medical.nema.org
25. www.embs.org/techcomm/tc-cbap/biosignal.html
26. www.ebi.ac.uk/compneur-srv/sed-ml
27. www.geneontology.org
28. www.biopax.org
29. sig.biostr.washington.edu/projects/fm/AboutFM
30. www.ebi.ac.uk/sbo
31. sig.biostr.washington.edu/projects/biosim/opb-intro
32. www.ebi.ac.uk/compneur-srv/kisao

Appendix: Case studies of clinical workflows in VPH-I projects

The following case studies of clinical workflows from some of the VPH-I projects are given here to indicate typical operations that are needed to link bio-science or patient data to clinical outcomes via the models, tools, and services being developed in the NoE. These workflows indicate how the VPH toolkit can be used in biomedical and clinical practice.

Case study #1. Cardiac modelling: Predicting clinical indices from models based on patient images

Many clinical diagnostic procedures that use patient images (MRI, CT, ultrasound, etc) could benefit from more quantitative model-based approaches to the evaluation of pathologies and the design of treatment strategies. euHeart (www.euheart.eu/) is a VPH-I project that is working with a diagnostic imaging company (Philips Healthcare), cardiac medical device companies and cardiologists to embed patient-specific computation models of the heart into a clinical workflow. The generic pipeline or workflow that represents the sequence of operations is shown in Figure 2.
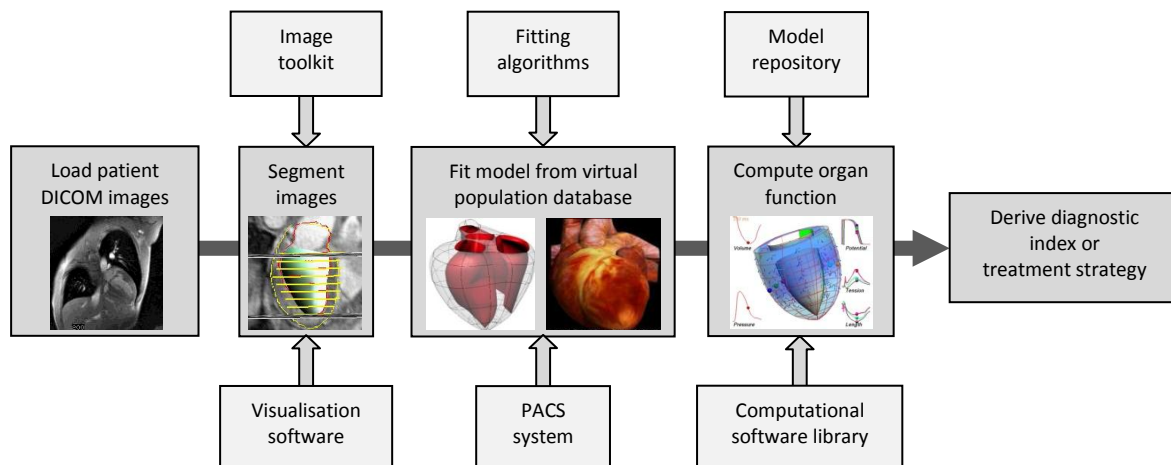
Figure 2: The sequence of steps used in the euHeart project in evaluating clinical indices or designing a treatment strategy by fitting models to patient images and then computing physiological function using biophysically based models. Illustrations here are for cardiac images and models from the euHeart project, but the workflow is equally applicable to all organ systems and many of the current VPH projects. Note that 'model repository' used in the final stage of the pipeline, incorporates results from biomedical research that includes molecular pathways (including signal transduction, gene regulation and metabolic pathways), ion channel electrophysiology and myofilament mechanics.

Case study #2 Using grid computing in time-critical medical diagnosis

The EU FP6 ViroLab project has developed techniques to tailor HIV treatments to the specific needs of an individual patient. To this end, they have developed a workflow execution system (called GSEngine) and a workflow (called the Virtual Patient Experiment) to automate the sequences of operations required to conduct a patient specific study that is able to rank the available drugs based on the specific characteristics of the HIV virus infecting a particular patient. Part of the Virtual Patient Experiment involves the development of models of a patient's HIV virus, which is then used to perform molecular dynamics simulations with the range of available HIV drugs, in order to predict each drug's effectiveness for the given patient mutations.

This workflow runs as follows: First, a genotypic assay is taken from an HIV patient and loaded into a database. Next, a clinician uses the Virtual Patient Experiment to query the database then produce a molecular dynamics model based on the individual patient's mutated HIV variant. Next, this interface is used to run a simulation of a range of available HIV drugs interacting with the HIV mutation. These simulations require large amounts of computing power and are typically run on a supercomputer. Once complete, a post-processing step is initiated to calculate the expected efficiency of each drug, then the clinician is notified of the ranked effectiveness of the available drugs.

This workflow can be readily generalised and applied to many different scenarios where a patient-specific molecular model can be constructed to rank the effectiveness of different drugs. The FP7 ContraCancrum project is reimplementing the workflow using many of the tools initially developed by ViroLab in order to rank the effectiveness of different cancer treatments.

Current projects that will use this workflow are: ViroLab, ContraCancrum

Case study #3 Grid-enabled neurosurgical imaging using simulation

The UK EPSRC funded GENIUS project is concerned with performing patient-specific brain blood flow simulations in support of clinical neurosurgery. To this end, the project has developed an efficient fluid solver based on the lattice-Boltzmann technique designed specifically to simulate blood flow in the brain, and has developed a bespoke client tool to automate the workflow of acquiring and segmenting data and executing the fluid solver application on one or more distributed high performance computing resources. The workflow is currently available from within the GENIUS problem solving environment, but could easily be modularised and made into a more generic plug-and-play application.
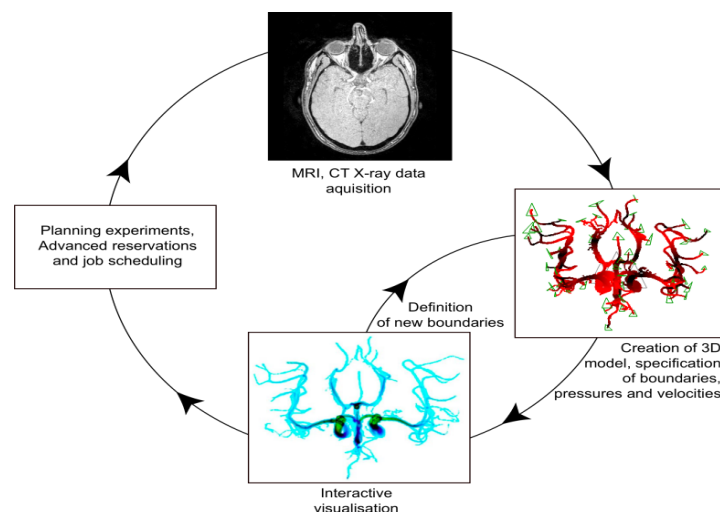

Figure 3: The GENIUS project workflow. Firstly, an MRI or CT brain scan takes place in a hospital, the scan is used to create a computational model, which is uploaded to a data repository. A clinical technician then uses the GENIUS toolkit to pull the model from the repository and use it to initiate a bloodflow simulation using a pre-reserved supercomputer class resource. Finally, a clinician uses the GENIUS toolkit interface to visualise and steer the simulation in preparation for a clinical procedure which is about to be conducted, altering parameters in the simulation as it is run, and visualising the results.

Case study #4 Multiscale modelling and application to in-stent restenosis

Almost every system that we seek to model in the Virtual Physiological Human initiative includes physical, chemical and biological processes that operate on multiple spatial and temporal scales. As we try to unravel the complex interactions between these processes, it is imperative that we are able to re-use the powerful models and descriptions that are available from decades of research on the single scale model components (e.g. cellular automata, finite element continua, agent-based models, ...). These models are inhomogeneous in nature and in structure, and are coded in a wide range of programming languages, reflecting the community in which they were developed. The need for the development of a formal theoretical framework for multiscale modelling, and for a flexible environment in which to couple models, was recognised by the COAST consortium (www.complex-automata.org), and partially funded by the Future and Emerging Technologies programme under contract EU-FP6-IST-FET- 033664. To demonstrate the power of the approach, the framework has been exercised in the context of the modelling of in-stent restenosis in coronary arteries, a common complication after a particular interventional procedure.

COAST has described an approach to the description of multiscale processes based on the construction of a scale separation map[4], providing an indication of which processes are scale-separated, when often simple hierarchical coupling approaches are effective and appropriate, and which are intimately coupled in scale, often requiring more complex coupling algorithms. It has developed and is distributing a Multiscale Coupling Library and Environment (MUSCLE), a platform independent agent system to couple multiscale simulations. MUSCLE provides the software framework to build simulations according to the complex automata theory[5,6].
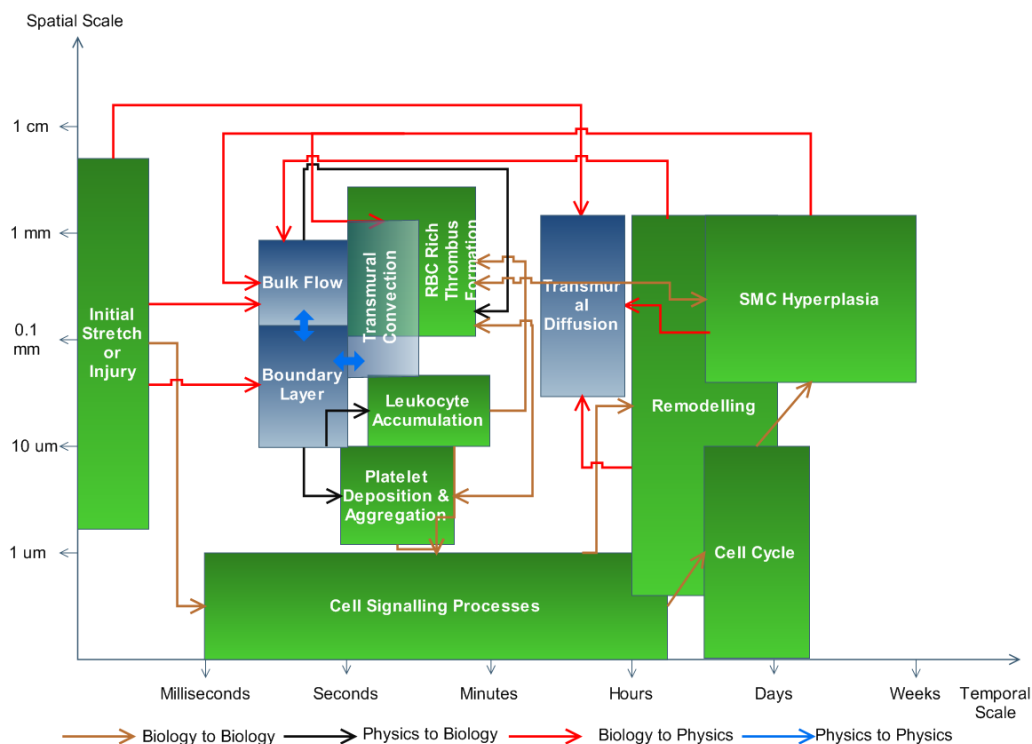

Figure 4a. Scale separation map illustrating physical process (blue) and biological processes (green) and interactions.

[4] DW Evans, PV Lawford, J Gunn, D Walker, DR Hose, RH Smallwood, B Chopard, M Krafczyk, J Bernsdorf, A Hoekstra: The Application of Multi-Scale Modelling to the Process of Development and Prevention of Stenosis in a Stented Coronary Artery, Philosophical Transactions of the Royal Society A.

[5] Jan Hegewald, Manfred Krafczyk, Jonas Tölke, Alfons Hoekstra and Bastien Chopard, An agent-based coupling platform for complex automata, M. Bubak et al. (Eds.): ICCS 2008, Part II, LNCS 5102, pp. 227–233, 2008.

[6] http://developer.berlios.de/projects/muscle/
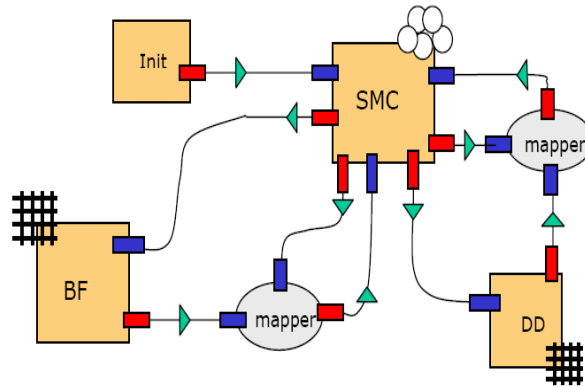
Figure 4b. Diagram illustrating individual processes (BF bulk flow, SMC smooth muscle cell model, DD drug diffusion) and smart conduits coupling the processes.

**Case study #5**: Integrated biomedical informatics for the management of cerebral aneurysms

About 2-3% of the population have aneurysms (local dilations of the blood vessels) in the cerebral circulation. These aneurysms are often silent (asymptomatic) and often stable, but when rupture occurs the consequences are devastating in up to two thirds of cases. The @neurIST project (IST-2004-027703) is predicated on the thesis that 'the process of disease diagnosis, treatment planning and treatment development is compromised by the fragmentation of the data that should underpin it'. Because of the devastating consequence of rupture, aneurysms are probably over-treated in current clinical practice. We need to be able to identify more accurately, in an individual, what their own risk of rupture might be.

@neurIST has developed an information infrastructure that allows us to collect together all of the relevant information about the patient, including demographic data, genetic information, lifestyle and other phenotypical information, image and radiological information, and even new information about the physical characteristics of the aneurysm derived from a complex image processing workflow. It has produced a clinical reference information model (CRIM) and a series of computational suites and practical tools to input, to access, and to process this information. This includes; @neuQuest, for entry of clinical data entry to the CRIM; @neuBrowser, for interrogation of distributed CRIM databases across multiple clinical centres; @neuFuse, embodying an effective and efficient workflow for visualisation and physical characterisation; @neuRisk, for personalised risk assessment; @neuEndo, for endovascular interventional planning; @neuLink for combining all personal and published information to reveal new information and to seek new associations that might contribute to our understanding of aneurismal disease as well as individual risk; and @neuCompute and @neuInfo that provide the infrastructure services to deliver data and computing resource to the point of use. @neurIST is a concrete realisation of the 'personalised medicine' objective of the Virtual Physiological Human.

One specific example of an effective workflow that has been built and implemented by @neurIST is that associated with the physical characterisation of the aneurysm. The challenge was to construct as system that was able to produce such characterisations with minimal user input and in a timely manner. The figure illustrates the operation of the workflow on a desktop computer, which might be adjacent to the clinical workstation (or even integrated with it). From @neuFuse the user is able to access a distributed clinical database (currently in most instances a mirrored database with all identifying personal data removed) and to select a case for study. The image is retrieved, in DICOM format, and visualised on the computer screen. A region of interest is selected and the aneurysm is segmented using an automatic process, without user intervention. A suite of tools is provided to perform a number of smoothing, clipping, repair and boundary condition application operations, to prepare the resulting geometrical model for analysis (shape analysis, stress analysis or computational fluid dynamics). An expert user can prepare an aneurysm for analysis in less than

fifteen minutes, and it has been demonstrated that a beginner can achieve results in approximately half an hour. The time for the simulation depends on the application (haemodynamic characterisation is most expensive) and on the hardware available. A transient haemodynamic characterisation might take three hours on relatively modest hardware resource. These times are certainly consistent with those available for an elective intervention case, for which the main issue is to minimise costly man-time in the analysis process. The resulting characterisations are saved, together with the model definition, in an 'abstract problem definition' file. This can be accessed by other @neurIST tools including @neuLink, @neuBrowser and @neuRisk, to seek associations between the physical characteristics and rupture or (in the longer term) to inform the interventional planning for the individual. It also provides an audit trail for the analysis process.

A major current focus of @neurIST is the dissemination and evaluation of its computational suites and tools. Early feedback, from an independent group of young clinicians engaged in a clinical training programme organised by the ESMINT foundation, has been favourable.
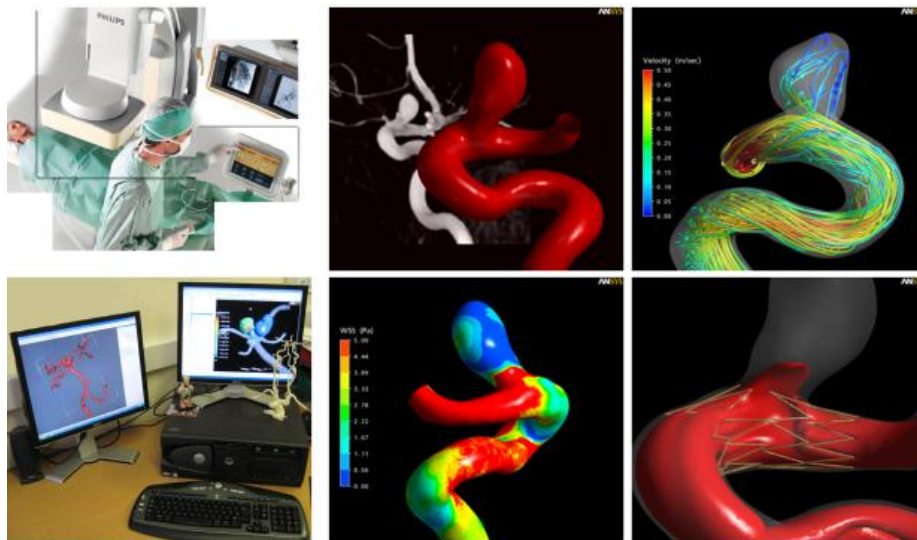
Figure 5: Top left: illustration of a modern interventional imaging system (courtesy Philips Healthcare). Bottom left: @neuFuse in operation. Right hand figures: Visualisation, haemodynamic characterisations and interventional planning.

Case study #6: Exploring parameter and model uncertainty using optimization algorithms for vertical systems – from genes, cells to composite tissues.

The grand VPH vision of a digitized body entails at least four hard conceptual and computational core problems. These are the integration of data and models between different (i) vertical (from gene to phenotype) and (ii) horizontal levels (between entities such as organs, cell and molecules). Moreover, for any given model capturing a biological process at a given level we are faced with the challenge of (iii) model uncertainty – the problem of deciding which elements (nodes, processes) and interactions (edges) that should be included in a model. Finally, for a given model structure there is a challenge (iv) to identify one or several sets of parameters which are consistent with available data.

It is important for the success of the VPH mission that significant research effort is expended on all of these core problems. To contribute to the last of these challenges, a workflow has been developed in the seed EP3 project (WHAM) and research at the Unit for Computational Medicine at the Karolinska Institutet. The workflow should be relevant to all VPH research projects aiming for causal understanding except those projects not mechanistically integrating across levels (i.e statistical models) or exceptional cases where biological knowledge can uniquely determine parameters and model structure.

To be specific and clinically relevant the workflow uses a time-dependent computational model capturing the plaque growth in an organ (the arterial wall), involving several cell-types and biological processes down the level of receptors, genes and proteins. Atherosclerosis is a major health problem and can induce a stroke or heart attack. The workflow runs as follows: (1) to review the state-of-art of biology of plaque growth; (2) to formulate a time-dependent mathematical model of the plaque formation where the interactions between elements are generically parameterized; (3) to search the parameter space using a particle swarm optimization (PSO) algorithm. Here we use physiological constraints and data from biological experiments to define seven evaluation functions (including lesion size) which therefore produce a fitness landscape corresponding to the "feasible sets of parameters"; (4) to analyze (statistics and machine learning algorithms) and identify different qualitative behaviours within the feasible sets of solutions from the simulations; (5) to formulate robust core predictions on the tissue (plaque level), cell and gene level which are based on groups of solutions; (6) to perform appropriate experimental validation.

From the ordinary differential equation model, containing 8 state-variables and 52 parameters, we computed more than 10000 different sets of parameters and approximately 10 % were considered feasible parameter sets. Surprisingly, we find that the behaviour of a high-dimensional (parameters) vertical model of a complex disease reduces to only three qualitatively different dynamical fingerprints (see figure). The underlying reason is that that the existence of several vertical layers (tissue, cell-to-cell interactions, genes) in essence strongly constrain feasible input-output conditions for each vertical model level. In this specific sense the integrative (i) vertical VPH challenge simplifies the (iv) parameter challenge. This finding makes the intrinsic parameter complexity in the system caused by our incomplete knowledge tractable for research. The model can now be used for predicting how to effectively reduce the size of the plaque at different points in time.

|(a)|(b)|(c)|
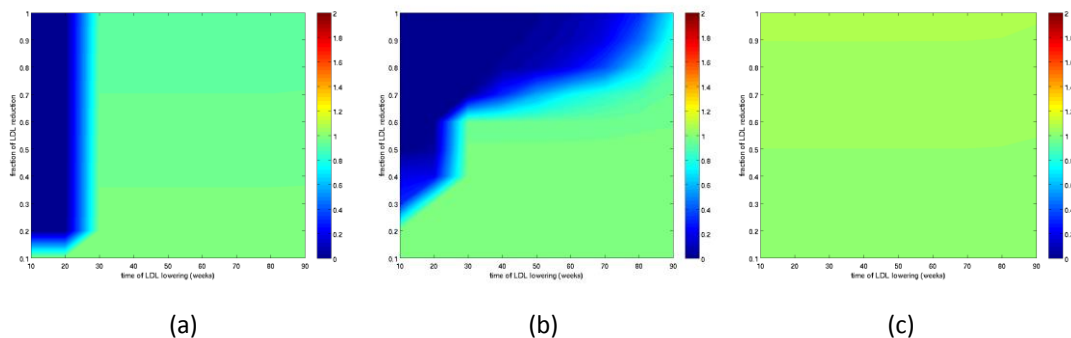
Figure 6. The colour scheme denotes the relative change of the plaque size as a function of time (x-axis) and size of the LDL reduction (y-axis). A blue points therefore indicates that there is no plaque at the end of the simulation if the LDL's level reduction is performed. A green point means that plaque will not be affected even in the case of a reduction. Among thousands of solutions we only observe three qualitative different behaviours.