Flagship Final Report
IT Future of Medicine

30 April 2012
1) Executive Summaries ........................................................................................................................................... 4
1.1 An ICT revolution that revolutionises medicine ......................................................................................... 4
1.2 Deciding about ICT and the future of medicine ......................................................................................... 6
1.3 A summary of the proposed Flagship Initiative specifically for the general public .............................. 8
2) S&T vision and methodology ......................................................................................................................... 10
   2.1 Scientific vision, unifying goal and main objectives of the flagship, also in comparison with the current state of the art ......................................................................................................................... 10
      2.1.1 Scientific vision ............................................................................................................................................. 10
      2.1.2 ICT challenges ........................................................................................................................................... 11
      2.1.3 Objectives ................................................................................................................................................. 16
   2.2 Matching of Flagship proposal with Flagship concept ............................................................................. 18
   2.3 Description of methodology, strategic research roadmap with milestones, and work plan for ramp-up phase ......................................................................................................................................................... 19
      Implementation plan ........................................................................................................................................ 20
   WP1 Medical Platform .................................................................................................................................... 21
   WP2 Analytical Platform: Analytical Technologies and ICT ....................................................................... 27
   WP3 Infrastructure Hardware and Software .................................................................................................. 32
   WP4 Data Pipelines ......................................................................................................................................... 37
   WP5 Computational methodologies: statistics, machine-learning, modelling and data-integration algorithms .......................................................................................................................................................... 41
   WP6 Integration platform ............................................................................................................................... 44
   WP7 Coordination and Management ............................................................................................................. 51
   2.4 Availability of resources to implement the roadmap .............................................................................. 54
   2.5 Metrics proposed to monitor progress in the Flagship ....................................................................... 56
   2.6 Coordination of activities and research communities ......................................................................... 57
3) Implementation ..................................................................................................................................................... 58
   3.1 Governance and scientific leadership ......................................................................................................... 58
   3.2 Management plan ....................................................................................................................................... 60
   3.3 Expertise of individual participants and quality of core consortium as a whole .................................. 61
   3.4 Approach to flexibility of partnership and involvement of relevant actors ........................................... 65
4) Impact ................................................................................................................................................................. 68
   4.1. Impact on Science and technology .......................................................................................................... 68
   4.2 Impact on economy and society ................................................................................................................ 70
       Potential contribution to EU competitiveness ......................................................................................... 70
       Impact on EU Health Policy Development and Policy Implementation .............................................. 72
4.3 Use of results and dissemination of knowledge ......................................................... 73
Internal Dissemination ........................................................................................................ 74
Engaging regulatory and governmental agencies ............................................................ 75
Exploitation of results ......................................................................................................... 75
Management of IPR ............................................................................................................. 76
4.4 Education and training at European level .................................................................... 78
4.5 Potential ethical and legal implications ......................................................................... 80
Supplementary Material (in annex) ..................................................................................... 87
Annex II: Overview participants and associated partners .................................................... 87
Annex III: Glossary .............................................................................................................. 87
Annex IV: Multimedia Material .......................................................................................... 87
1) Executive Summaries

1.1 An ICT revolution that revolutionises medicine

Europe ages. Its health budgets soar. Europeans suffer from ill-managed diseases. And all of this against a background of extremely successful life sciences with little societal impact, and revolutionary ICT that seems to affect all areas of society except for medicine.

What will Europe do? Wait until the problems disappear? Wait until solutions originate in other countries? Or realise that it is best posed to provide a solution; a revolutionary ICT that revolutionises medicine:

ICT for the Future of Medicine (ITFoM)

ITFoM (www.itfom.eu) is a flagship sailing under the flag of the European Union that, empowered by ICT sails, sails from state-of-the-art medicine of 2012 to a truly individualised medicine of 2025.

Much of life is computation: based on a program blueprinted in the genome but instantiated by information communicated to the genome from the environment, every living cell computes its behaviour, its response to stimuli, and its communication with other cells. It does this by using a biological information and communications strategy that gives rise to behaviour of tissues, organs and the organism that is of unrivalled robustness and functionality. Designed and optimised by billions of years of evolution, the computing is carried out by interacting elements (genes, proteins, transcripts, metabolites, cells, tissues): a ‘biological ICT machine’ continuously computing responses to communicated information.

The results of these biological ‘computations’ are of crucial importance for us humans. They determine if we live or die, if we are ill or healthy, and if a specific therapy will help to cure us, or make us sicker. At present medicine has the arduous task of dealing with the outcomes of these computations without access to the algorithms or to the enabling ‘bio ICT’. Medicine uses the experience and knowledge in the individual patient’s physician’s brain, most of which gets lost when that physician retires. Medicine receives little support from the life sciences that produce oceans of data that are nicely tucked away in journals or databases. Due to the tremendous complexity of the human body the data cannot be connected in ways that predict how to treat the individual patient, taking into account everything that we can measure to find out what makes him/her different from similar patients: the genome, the proteins in his blood, the metabolites in his urine, but also detailed images of his body, or advanced sensors that continuously analyses the functioning of the heart.

For every new patient, instead, medicine has to go by the experience obtained from thousands of other patients, culminating in the concept of an average patient suffering from an average disease of a specific type. Clinical trials of new medicines also follow this concept, at best improved by dividing patients into two or three different groups. Consequently, multifactorial diseases such as cancer and diabetes are sometimes managed, but rarely cured. For this reason, two different people can well respond completely differently to the same therapy of ‘the same’ disease. And they respond differently to similar nutrition or to having an identical DNA mutation. Having up to 80% non-responders to specific medicines is no exception, at tremendous cost to the healthcare system and with tragic delays in treatment. From their genetic makeup, two individuals differ almost in every
tenth protein, and dietary and other environmental differences lead to an even greater difference in network functions. Often their diseases are very different on a molecular level, whilst looking similar to the unassisted eye of the doctor. The diversity in response is well understood. Why is this understanding then not empowering medicine?

The reason is that the networks in the human body are far too complex to be understood by the unaided human mind. ITFoM calls ICT to the rescue. We propose for modern ICT – which has already revolutionised so many other aspects of human life – to enable a revolution in medicine to pave the way to truly individualised medicine. This should be done by putting together all the relevant biomedical information for individual people to generate personalised computational health models. Individualised therapies can then be designed and safely tested on such ‘virtual patient’ computer replicas before being applied to the real patient.

However, existing ICT cannot provide for this at present. For each individual the calculation would require about two thousand of the fastest supercomputers that would, moreover, cripple a power grid if switched on. ITFoM therefore proposes a second revolution, now in ICT itself: enabling the calculation of health and disease for millions of (consenting) individuals. To enable this revolution, ITFoM proposes a two-tiered mechanism. The first tier maximises the use of existing and developing ICT and analytical technologies. The second tier develops and implements a new type of ICT based on the concept that if the human body itself is able to perform the computation of health and disease, then an ICT that is structured as a replica of the human body should be able to achieve the same.

ITFoM will develop the new technologies in four ICT workpackages. WP3 develops the novel hardware and software required; constructing predictive ‘virtual individuals’ in a form that would scale to large fractions of the population in Europe. WP4 builds pipelines of information and data between the various data sources, producers and consumers, basic researchers and hospitals. WP5 produces the new computational ICT that is required for the biology-driven ICT. The approach depends on smooth interfacing of WP1 with the medical world of patients, physicians, nurses, healthcare assessors and policy makers, as well as with clinical reference material. And it has to ensure that the massive influx of analytical data becomes organised through new ICT (WP2) so that it can actually be fed into the integrators of WP6. WP6 integrates all activities at ICT platforms producing interoperability of maps, data, models and tools. It also uses four different modelling strategies to integrate data into simulations of the dynamic behaviour of the virtual patient models. A fifth sub-workpackage then combines the results of these and possible other modelling strategies to produce the optimal medical predictions for the patient, to be obtained from taking into account all his/her relevant biomedical information.

WP6 will construct an initial set of reference models using the four different modelling pipelines, based on detailed analysis of patient samples by a large range of –omics techniques (genomics, epigenomics, transcriptomics, proteomics, metabolomics, metagenomics etc.) in WP2, developed further on the basis of data from the personal genome project (PGP), which aims to carry out a detailed analysis of up to 100,000 volunteers with a wide range of analytical techniques. These models will then be individualised with detailed –omics, imaging, sensor, nutrition and lifestyle data. The outcome of different therapeutic or preventive options will be calculated, using an appropriate mix of the four modelling strategies, with the results combined in a statistically clean fashion to provide the optimal information for the doctor treating the patient. The models can also be used (in anonymised form) to conduct ‘virtual clinical trials’, speeding up the development of new generations of drugs for specific subgroups of patients, for which no therapy options are currently

---

available, or, for instance, to model the spread of pandemics in specific areas in Europe. Models can be used in interactions between doctors and patients to increase understanding of disease, compliance with therapy and input of patient and doctor knowledge into the disease model.

The personalised health models will not remain academic. They are produced in the public private partnership of ITFoM and will be used for a number of medical use cases covering the areas of cancer, metabolic syndrome, cardiovascular disease, and infectious and neurodegenerative diseases. Rare diseases, healthy ageing and microbiomics are engaged as cross cutting areas. Industry is associated to ensure that the resulting ICT models can be implemented in a plethora of applications, ranging from portable personal healthcare assistants to intelligent sensors and personal dietary designers.

Accepting that the final governance model is subject to requirements imposed by the European Commission, ITFoM proposes an EEIG (European Economic Interest Grouping) type of structure with a governance board consisting of shareholders (funders), a steering committee functioning as the board of directors with an elected chairperson, and a CEO (Chief Executive Officer), CFO (Chief Financial Officer) and COO (Chief Operating Officer) that are controlled by the steering committee with respect to R&D content.

Trillions of euros are spent annually in worldwide healthcare. Also the relevant ICT industry is immense. ITFoM will therefore focus on integrating the R&D activities that already exist by (i) promoting standardisation and interoperability, (ii) creating reference maps, data, models and tools, (iii) ensuring that the necessary hardware, software, computational ICT and pipelines are engaged or developed, (iv) resolving ethical, policy-making, IP and related issues in ways that are attractive to all stakeholders, and (v) making sure that the job of producing and implementing individualised medicine gets done.

ITFoM has associated around it an R&D cloud of some €300bn and expects to require some €1.3bn of core funding to carry out the proposed integration into the European capability to build millions of ICT models for individual humans to help design individualised therapies and healthcare advice. It expects part of the core funding to come from the FP7 and H2020 programs, and the rest from integrating national and other H2020 programs.

ITFoM will also produce an unprecedented level of integration of organisations. First it will help integrate R&D programs of multiple European nations by association to the ITFoM model-driven integration strategy. Second, it will integrate medicine and ICT activities; medicine will be proposed by the newest ICT. And then it will integrate the molecular systems biomedicine (MSB) with the Virtual Physiological Human communities. Moreover, it will link nature with nurture, resolving the old controversy between these paradigms.

The resulting innovations will answer to the European digital agenda challenges, answer to major health and societal challenges by revitalising European industrial value chains such as (1) drug development, (2) diagnostics and prevention, (3) health devices and Point of Care (PoC), (4) hardware and software together with robust and secure pipelines to support the flow of medical data circulating. We are expecting a strong return on investment (ROI) for both European countries, and ITFoM stakeholders with new products and usages, and new entrepreneurial opportunities in the related industries. ITFoM will lead the way to a better and innovative healthcare system at sustainable costs. See also: http://www.itfomthemovie.eu/.

1.2 Deciding about ICT and the future of medicine

The evidence-based medicine of today is supported by the use of drugs that were developed without considering the great inherent variability in the response of individual patients to these drugs. The
clinical trials involved patients that were selected on the basis of only some criteria to constitute a homogeneous group, which was then split into two subgroups. Progresses in genomics and systems biology have shown that such groups are inherently heterogeneous; every individual differs substantially from every other individual in her/his molecular networks. Most of the drugs that make it to market have low to medium efficacies and it is almost impossible to determine which individual patients will respond well to the treatment and which may develop adverse events. The most likely interpretation is that some individuals respond extremely well whereas others respond poorly. Normally, medical doctors can only make educated guesses that a treatment will either or not be effective by having to rely on inadequate evidence (imaging, blood samples, and medical records), frustrating both medical doctors and their patients. They mostly have to rely on the concept of an average patient, rather than on a broad population of diverse patients. ITFoM (www.itfom.eu) will move medicine away from such average treatment toward precision treatments based on ICT-empowered predictions of individual responses to various possible treatment scenarios. In further sophistication of disease treatment, for example, a patient will receive an optimised cocktail of medicines affecting his/her molecular networks with maximum accuracy, e.g. efficiently killing tumour cells whilst leaving most bystander cells intact.

At the European level there is no systematic health and medical data digitalisation yet, ITFoM will foster and coordinate innovative ICT solutions to (1) support secure and reliable online access to their personal health information for consenting patients (by 2015 for the HER), 2) support ‘home-safe data security’ by allowing patients offline access to model results, and (3) support medical professionals in running intelligible model simulations to assist them in improving prevention diagnostics and therapies. Moreover, ITFoM solutions will help control and reduce the cost/benefit ratio related to energy consumption and the novel healthcare that they enable.

Context

The European economic health landscape is facing drastic dynamical changes due to a number of synergistic factors: (i) macro-economic, with the economic downturn, the emerging countries’ competitiveness and new business models, (ii) health sector driven, with strong reimbursement pressure, a complex regulatory landscape and a decrease in R&D efforts within both pharmaceutical and device industries, (iii) societal, with population ageing and the transition from acute to chronic diseases. Bottom-up pressures from new challenges in health data digitalisation, technological developments and the flow of biomedical data arising from clinical and molecular –omics analysis, top-down constraints with needs for innovative and targeted drugs, smart devices and other new technological support systems that answer to systems medicine challenges, and patients’ personal engagement are all paving the way toward individualised medicine applications. Increasing healthcare costs are a significant component of public spending spiralling out of control in many countries in Europe and beyond. They are increasingly considered as an important component in determining the credit rating of countries and can therefore significantly affect the national credit costs. In addition, ineffective healthcare systems reduce both the quality of life of older citizens and also hamper their ability to actively participate in all levels of society. This, in turn, significantly increases the costs of pension systems in times of demographic change. The cost and quality of our healthcare systems not only impact the health and wellbeing of European citizens, but are also key factors in the many unresolved long-term societal problems that Europe faces.

Truly personalised healthcare does offer vast opportunities. With annual healthcare expenses of the OECD and BRIC (Brazil, Russia, India, China) countries alone approaching seven trillion dollars, a truly personalised data-rich and computation-driven medicine of the future that replaces a significant fraction of current expenses by novel high-tech approaches will be able to redistribute significant fractions of the cost of current healthcare to new markets, generating both medical and economic benefits.
ITFoM answers to this unprecedented societal challenge by federating the top specialists of ICT and medicine by proposing an applied programme that fosters innovative solutions where both academics and industries (including SME's) will lead the way to deliver a better healthcare for Europe and the world.

**Bottlenecks**

Standard and interoperability of health and medical data, cross-border management, ethical and security issues or heterogeneity of hospital IT infrastructures, are bottlenecks for the development and the deployment of personalised medicine applications. Others are related to private or public reimbursement and readiness, the degree to which those involved are individually and collectively primed, motivated, and technically capable of executing the changes (patients, clinicians, regulatory bodies, payers...).

ITFoM will address the huge IT implications of population-wide individualised patient care using high-resolution analytics. The project outcomes will enable the calculation of health, disease and therapy for individual patients. This will revolutionise healthcare with (i) vast benefits for individuals through prevention, (ii) vast benefits for physicians through better diagnostic and therapeutic support, and (iii) new commercial opportunities in the ICT, analytic and drug development sectors. The socio-economic impact is enormous. This project will change the understanding of diseases and lead to a paradigm shift of prevention and therapeutic approaches. It will prepare the way for the inclusion of eHealth records in a protected and safe data environment. The pharmaceutical and biotech industries will benefit from new tools for drug development and will be able to come more quickly and less expensively with their products to market. And finally the ICT sector will benefit substantially because of the massive new market areas relating to the manifold aspects of personalised healthcare emerging, but also because the innovative technologies will be applicable in many other fields of engineering, such as the automotive and transport industries and the financial markets. Indeed, biology-driven ICT may well provide a great boost for the computation of complex systems in general. See also [http://www.itfomthemovie.eu/](http://www.itfomthemovie.eu/).

**1.3 A summary of the proposed Flagship Initiative specifically for the general public**

For many years there have been discussions about how health systems across Europe must change. With a lack of progress in treating the majority of complex diseases, major issues in drug toxicity and side effects, an ageing population and seemingly unsustainable costs, much of the discussion has focused on the necessity for change in healthcare systems to prevent their collapse. To drive this change it has become clear that the future of European healthcare depends upon major breakthroughs in science and technology. IT Future of Medicine (ITFoM, www.itfom.eu) is a project that will generate these breakthroughs. ITFoM will show that it is possible to streamline the delivery of medicines, improve patient outcomes and afford the patient greater control over and insight into their own health. It is, in fact, possible to do so without having to rely on more medical personnel and increasing health budgets. We must move away from one-size-fits-all and organise medicine and medical technologies so as to ensure that patients are always at the centre.

From the notion that we are all unique individuals all the way down to our molecular makeup, ITFoM aims to create integrated computer models for health tailored to every person that take all these differences into account. These “Virtual Patient” models will predict likely illnesses and suggest appropriate tests and treatments. We will achieve this ambitious goal by gathering and systematically integrating all accessible and relevant data of patients into personalised models. The highly detailed information about an individual’s genetic makeup, for instance, can be enriched with
data gathered from other modern analytical techniques that are rapidly becoming cheaper and more advanced, giving unprecedented insight into the functioning of our cells, tissues, organs, and our body as a whole. But the core technologies and their secure realisation must still be developed and implemented, since merely obtaining large amounts of detailed biological data for individual people does not turn this information into actual knowledge that can help us in taking medical and lifestyle decisions.

The idea of personalising medicine is already becoming increasingly important in all fields of healthcare, yet we are not currently incorporating all the technological innovations that have recently emerged and which could greatly improve our understanding of human health and disease. While the past decades have seen some crucial efforts in dealing with mainly infectious diseases, a similar impact is not yet evident in many other areas. For example, progress has been comparatively slow with treating diseases like cancer, diabetes, Alzheimer’s, Parkinson’s, osteoporosis and cardiovascular diseases. Diseases like these are still treated with generic blockbuster medication, putting many patients at risk of suffering from severe overdoses. Today about 80% of all cancer patients receive medication that will not have a positive impact on their health, whilst the medication itself produces severe toxic side effects. This is due to the fact that every tumour is different and subject to each individual patient’s unique characteristics. These facts have to be taken into account to determine optimal medical treatment on a case-to-case basis. And not merely in cancer; the same is true for most other complex illnesses.

So what are these “Virtual Patient” health models? In a simplified way they resemble gathering meteorological data to run simulations that forecast tomorrow’s weather; something that we have become increasingly good at. Moreover, scientists and engineers already use simulations everywhere, for instance to test cars and airplanes before building them. These simulations are faster, less expensive and – most importantly – less dangerous than trying it in reality first. Enormous improvements in techniques to characterise every single patient in great detail as well as the huge developments in information and communications technologies (ICT) have brought performing similar simulations for personal health and medical treatment within our reach.

It is a significant challenge that has been compared to the Apollo space programme in scale and complexity. Since everybody is different, the computational models will need to be tailored to each individual to reflect their own unique anatomical, physiological and genetic makeup. The huge amount of data generated for each individual is only one issue to be solved when tackling this challenge. We must set up a massive digital infrastructure with sufficient capacities to securely process, store, integrate and analyse vast quantities of information. This will result in new computational ideas and the need for novel human-computer interfaces that, in turn, will also pave the way for a broad spectrum of innovations and economic opportunities in Europe. We will see the need for improved instrumentation and automation of techniques to gather, analyse and integrate all the information required. Intuitive visualisations for each patient will then allow doctors to operate the system and receive guidance on tests and treatment designed specifically for the individual patient. In addition, healthy people will benefit from their personal health models by receiving lifestyle advice, helping them to circumvent health risks and assure their optimal condition and wellbeing. This visionary project will combine expertise in ICT with professionals from the fields of medicine, life sciences and public health. Together – and for the first time – they will create a truly personalised healthcare. In the long run, ITFoM is set to revolutionise our current idea and clinical practice of prevention and therapy, resulting in a new standard for individualised and effective medicine for the future. See also http://www.itfomthemovie.eu/.
2) S&T vision and methodology

2.1 Scientific vision, unifying goal and main objectives of the flagship, also in comparison with the current state of the art

2.1.1 Scientific vision

In its essence, life can be considered a computational process. Based on the program in the genome, and influenced by the environment, every cell computes its behaviour, its response to stimuli, and its interactions with other cells, ultimately giving rise to tissues, organs and the entire organism. As illustrated by the high degree of similarity between identical twins – sharing the same genome – this computation appears to be to a large extent deterministic. In contrast to computation in digital computers, it is however carried out by interacting elements (genes, proteins, transcripts, metabolites, cells, tissues etc.), with both elements and their interactions formed and shaped by billions of years of evolution of life on earth. A human, with 100 trillion cells in his body, each programmed by a genome of three billion bases, can therefore be considered a highly robust and efficient computation device.

The result of the computation is of central importance for us. It determines if we live or die, if we are sick or healthy, and if some specific therapy will actually cure us of some disease, or instead make us even sicker. The latter outcome is, unfortunately, not rare. On average, only a quarter of cancer patients that receive a specific mechanistic therapy will respond to the drug. Three quarter, treated with that (expensive) drug, will show often-debilitating side effects without any objective benefits from the treatment. Even more unfortunately so, cancer is not an exception. Many patients within a wide range of diseases respond very differently to therapies. Many deaths are caused by side effects, with a large number of avoidable deaths caused by the medical treatment itself.

We propose here to solve this problem by a straightforward concept. A concept, moreover, that has already proved its power in many other areas: the simulation of potentially dangerous, harmful or expensive procedures by in detailed computer models. If we design a new car or if we teach pilots how to fly a new type of airplane, we use computer models to eliminate dangers, prevent losses and cut costs. Any pilot switching to a type of airplane will first be trained extensively in a simulator that accurately represents the plane that they will fly in, prior to being allowed to fly the real plane. Every patient coming to the doctor might well be considered the equivalent of a new type of airplane, but instead will in most respects not be treated as a unique individual, but as an average member of a large group of supposedly similar, but actually very different, patients. Patients that, moreover, suffer from diseases that often differ greatly when zoomed in on their molecular bases.

To make such models of complex processes we need three basic components,

1. detailed knowledge of the individual state of the – complex – biological system, provided by detailed –omics, imaging and sensor data for every individual person,

2. detailed information on the rules that control the behaviour of each individual component of the system, and

3. the computing power required to predict the evolution of the complex system under specific circumstances and perturbations.
We are now in the fortunate situation that all these components have become available, offering a unique opportunity to develop models of individuals as the basis of a new personalised, data-rich computation-intensive medicine/prevention strategies.

**Dramatic progress in analytical/diagnostic technologies**

We have made huge progress in analytical techniques, allowing us to analyse the individual patient in enormous detail. We are now able to know more about every individual patient than we knew only a few years ago about the biology of man in general. The most dramatic development in this direction has been the enormous progress in DNA sequencing. The sequence of the human genome was completed roughly 10 years ago, at an estimated cost of 1-3 billion Dollars, in an international project lasting ten years. Relative to the human genome project, the cost of sequencing a genome has dropped by close to a million fold, and is still dropping at a phenomenal rate, opening the way to the use of deep sequencing as diagnostic tool in medical practice. Similarly other –omics technologies, developed originally for systematic analyses in high end research applications (proteomics, metabolomics etc.) could become useful for routine diagnostic applications, providing more detailed information for modelling individual patients.

**The rapidly increasing knowledge of human biology**

Intense support of basic research worldwide over many decades has given unprecedented insights into biological networks in man and model organisms in health and disease. It has, for example, been estimated, that, since the beginning of the ‘war on cancer’ in the USA in the early seventies, more than a trillion dollars has been spent on cancer research alone, as part of an annual expenditure of up to a trillion Dollars for biomedical research worldwide. Knowledge gained from this vast effort can now be used to guide the construction of mechanistic models covering different diseases areas. In turn, the comparison between prediction and actual response of millions of patients will, in the future, be a major driving force in improving our understanding of biological processes in man, driving further improvements in the prediction power of the models.

**Sustained rapid increases in ICT capabilities**

Similarly, and for a long time, computing power has risen for many decades, best illustrated by the development of leading supercomputers. As described in WP3, we expect exascale computing (machines able to carry out ten to the eighteen floating point operations per second) to become available around 2018, 5 years after the beginning of ITFoM. By the anticipated end of the flagship project, we might already be well on the way to zettascale computing. Alternative computing approaches (e.g. WP3.5) as well as algorithmic improvements (WPs 3, 4, 5 and 6) offer additional possibilities to resolve remaining computational bottlenecks.

**The goal**

On the basis of these dramatic developments, we propose to develop the basic infrastructure required to generate integrated molecular/physiological/anatomical models of every individual in the healthcare system, as the basis of truly personalised healthcare, with enormous repercussions for ICT developments and our understanding of biological and disease processes.

### 2.1.2 ICT challenges

**ICT, the future of medicine**

Making truly individualised medicine possible for every individual in the healthcare system will require an enormous biomedical effort. ITFoM recognises that it is not a biomedical effort per se that will overcome the major limitations and challenges; the major challenge is in developing and
implementing the required ICT. ICT has progressed immensely over the last two decades and now greatly improves many aspects of our lives. It has not penetrated similarly however into the inner workings of biomedical research and medicine. This is to a large extent due the enormous computational challenges that cannot be conquered by current computing standards. For this to happen we propose to deal with what seemed impossible in terms of complexity and scale by developing novel computing paradigms that mimic the organisation of life as a computational process targeted to the particularities of the human organism. This new ICT absorbs the rich experience of physicians and patients and to put this to use in computational models for individual patients. It also needs to seamlessly integrate with molecular analytical techniques (WP2). This requires extensive and sophisticated hardware and software solutions (WP3), fast and reliable data pipelines (WP4), and smart, adaptive and statistically reliable computational methods (WP5) for all data to be integrated into extremely complex mathematical models of the human body using linguæ francæ to facilitate sharing and synergism (WP6). Within 10 years, the combined ICT effort will be able to integrate numerous ICT and biomedicine programs into a process that enables truly individualised medicine: virtual patient models that will revolutionise ICT and medicine, supporting patients, general practitioners, clinicians and health policy makers alike (engaged through WP1).

Huge integration challenges

ITFoM builds on the sequencing of the human genome, yet aims far beyond: understanding the dynamically expressed human genome in its equally dynamic environment. Thanks to new –omics, imaging and sensor techniques, vast amounts of extremely heterogeneous data confront limited understanding of the molecular, pathway, cellular, anatomical and contextual organisation of human health. Data are stored in repositories that, in addition to often being standalone, are specific and optimal for said data type, even though for any health assessment or disease treatment the dynamic integration of subsets of data of diverse data types are what matters. There is a substantial disconnect between the patient, the clinician, the molecular geneticist, the computer scientist, the physiologist, the mathematical modeller, the bioinformatician and the public health official; they all need to do their utmost to be expert in their own fields. The ambition of making realistic computational ICT models of the human that relate molecules to pathophysiology, and the ambition to make the instantiations of those models for millions of individual humans that enable truly individualised medicine, are currently always pushed to future horizons. There is great need for an ambitious integrative project that drives and embraces new technologies to work to the convergence of these disciplines in order to break the current deferral to establish a crucial breakthrough; and this for the Horizon 2020.

The challenge quantified

The challenge is enormous. From the given that the human body contains more than 3x10^24 non-trivial molecules and that the fastest relevant time scale for biomolecular behaviour in networks is about 1 ms, one can calculate that performing a brute-force all-out high-resolution computation of the full dynamics of the human body would require 10^38 flops for just a 1-second time interval, not even taking into account the huge range of potential interactions. This is far beyond what will ever be feasible in terms of storage computation and even storage.

Particularly challenging is already the required computational processing and storage of advanced imaging information that spans the macroscopic (e.g. whole body) to microscopic range. Also other types of information need to be included, such as the PK/PD (pharmacokinetic/pharmacodynamic) of drugs, the nutritional status and ‘feelings’ of the patient, and the expertise of the clinician. Although less voluminous, here the challenge is that information is of a very different type and thereby hard to integrate with the molecular information.
Even less ambitious state-of-the-art models will cause major computing challenges. Modelling the response of one individual to e.g. 200 different drugs or drug combinations, taking into account the response of 200 different interacting cell types in the body, and carrying out the analysis 500 times to cover different parameter values (Monte Carlo runs) will already require approximately 20 million modelling runs. Doing this once a year for every one of 500 million Europeans using current technology, would require trillions of cores, and use approximately €60bn worth of electricity per year alone. This clearly calls for the development of innovative hardware and software solutions, as well as new computing paradigms in order to make the virtual-patient paradigm a sustainable choice for personalising medicine throughout the world.

The data challenges are equally daunting. One can easily imagine that it will be cost effective to routinely measure the full set of blood-accessible biomolecules using RNA, protein and metabolite measurements for every patient every time they visit a physician. In addition, for specific disease scenarios more invasive (e.g. biopsy based) samples will have to be analysed. The raw information would comprise molecular data in a dynamic range of $10^5$ to $10^9$, and so requiring the sampling of around $10^7$ molecules per assay, per sample, per health episode, per patient. A 10 million-fold sampling of RNA is around 0.5 TB of storage (assuming fully sequenced RNA, and the current routine storage of RNA analytes at 20 bits/base). In addition, current high-density imaging will increase in both resolution and utility. This means that without the ICT technology development and implementation that ITFoM will initiate, the total storage requirement would naively be of high exascale to low zettascale across Europe. To allow Europe even to consider individualised medicine requires both generic developments in storage and data-aware computing and specific routines, such as specialised compression and fast reliable analysis methodologies, to transform the problem into a challenging but feasible low to mid exascale problem.

ITFoM proposes ICT solutions to these limitations. It will first propel the ICT that is already moving and, on top of that, lead a second development towards revolutionary novel ICT.

**Propulsion: stepwise improvement of working prototypes (especially WPs 1, 2, 3.2, 3.5, 6.2)**

The first step will build on current advances in the best European research labs with an incremental, data-driven approach aiming to exploit the results available to date by bringing them to medicine. To be able to predict, for example, optimal treatment for an individual patient, we might not have to simulate every cell and every tissue in exquisite detail. Just like simulations of airplanes to be tested for aerodynamic behaviour do not require the exact details of the seats in business class. We therefore plan to incrementally explore the biomedical modelling for individual patients in increasing detail in a series of ‘working prototypes’ that can improve patient treatment decisions and drug development in some areas of medicine (e.g. specific types of cancer) early on. Individualised models that initially only contain the parts of complex biological networks in tumour cells that are most relevant to predict the response of the tumour to specific drugs (pathways containing most of the genes/proteins showing somatic changes in tumours or interacting with the drugs used in cancer treatment) should already be able to predict effects and side effects of specific drugs for specific cancer patients. Similarly, models will initially only represent limited number of cell types as separate, interacting molecular sub-models, i.e. different cell types in the tumour, liver cells metabolising the drug, other cell types in the body to identify possible side effects). With every progress we will make in solving the many computational bottlenecks, as our knowledge about the intricate processes in each cell type and their multifarious interactions will increase accordingly, we will be able to expand the prediction power, the range of diseases, as well as the range of patients receiving better, more effective medicine. This will occur in a series of working prototypes, validated and improved continuously in pre-clinical and clinical tests. These models will be fostered such that everybody’s virtual patient model will be a natural companion that empowers each individual in health and disease; from the patient in intensive-care, to a future where personal models assist on
predicting the effects of various lifestyle choices, or monitoring and managing a potentially adverse condition, aided by advanced healthcare devices and applications.

Novel biology-driven ICT (WPs 1, 2, 3, 5, 6.1, 6.3 and 6.4)

ITFoM will also develop new biology-driven ICT, through a model-driven bottom-up strategy that builds on while expanding an approach demonstrated (e.g.) in differential network-based drug design against sleeping sickness in Amsterdam and the German ‘virtual liver’ as well as in a number of VPH (Virtual Physiological Human) projects: the building of exact replicas of molecular and anatomical structures of the human body, up to a ‘Digital/Virtual Human’, a complete model representing the individual body of every patient in exquisite detail.

As calculated above, an increase in computing power following Moore’s law alone will not make ICT meet this challenge in any foreseeable future. An ICT revolution will be required that makes the computations much more efficient than the above data-driven strategy. Here ITFoM takes inspiration from considering the human body itself as an ICT ‘machine’ that computes its own behaviour in real time. Apparently, billions of years of biological evolution have selected for a robust data flow and processing organisation within living organisms, unique computational machinery that deals with the otherwise impossible computation challenge. ITFoM will learn from this principle, which intimately binds the ‘software’ and ‘hardware’ of the human organism, which will serve as Leitmotif for the development of a replica for the organisation of its ICT in a manner that should enable ITFoM to model the human body in great detail.

Based on latest advances in biology, in this approach we aim to organise the current ICT and computing elements, such as data, networks and computing models in terms of recognisable components of the human body, its networks, and their interactions. Our multi-scale approach to modelling the organism will be inspired by the concept of semi-autonomous self-organising entities that can be both autonomous and cooperative at different levels of resolution. We start from the premise that a cell as well as an organ as well as a whole organism receive and integrate information from the outside concerning nutritional and environmental status, communicates this information to the genome which in turn induces the creation of new information that defines its new biochemical functioning. This in turn leads to the activation of intracellular monitors of the information which is further transferred to the outside respectively to other cells, organs or organism. Following this multi-scale modelling perspective we will build our algorithms for functional classes of macromolecules (enzymes, mRNA, genes), for pathways and for subcellular compartments, for important cell types and for organs. These models will subsequently be integrated into a series of models of the human organism from a biochemical, physiological, anatomical perspective.

To deal with information uncertainty ITFoM will apply machine-learning approaches in the developments of a series of models. The modeller will be given the option to select between alternative models, using criteria such as robustness and fitness of the human in accordance with known experimental data, and relevance to the question asked. Here the ICT modelling will itself become self-organising, in a supervised machine-learning process. An important ICT challenge is to put also this self-organisation mechanism in place within the confines of the genome-sequence and the individual’s environment information, whilst ensuring that the criteria used remain under the control of the patient and his physician.

Information flows in “biology-driven” ICT

We regard ICT as the ensemble of technologies that optimise the communication of information. That communication often consists of a linear superposition of information packages. However, there are many cases of profound human interest in which information itself acts nonlinearly such that new phenomena emerge. ICT enabled several such emerging phenomena through the acceleration of communication of information and access to knowledge. The effects of instant
sharing of information through ICT is often much more valuable than the sum of the information parts distributed across a collective. ICT has proven to accelerate ‘collective intelligence’ and to enable coordinated action through, the nonlinear effect leading to new phenomena in many areas of our ‘eSociety’ which impacted economies of scale (e.g. through influence spreading across ICT-enabled social networks) to social innovation generation and political pressure. ITFoM aims to integrate these principles for collecting and communicating information. We could even see the structure of the data flow as “anthropomorphic”; it is meant to have the same organisation as the human body. Since for most aspects the information flow and regulatory circuits of the human body are robust (asymptotically stable in the sense of Lyapunov), and they will take into account the relevant, large but limited set of parameters, the expectation is that we can learn from this in building robust algorithms and hardware standards. These new ICT paradigms may indeed be also applied in computing the weather, economical phenomena or global warming. Here however, such approach will be used to try to perform the possibly most complex of all computations, the simulation of individual health and disease that takes into account all environmental effects.

The system of interacting molecules and tissues that constitute the human body can be regarded as an information communication process. It computes on the basis of a plethora of input information and as a function of millions of parameter values denoted in genome sequence, epigenomics and molecular/physiological state functions of the organism. But it is more than a computer in that it integrates information into predictions of the future and decisions on optimal behaviour. It thereby is intrinsically a highly nonlinear process that is organised in a hierarchical manner.

Nonlinear phenomena such as heartbeat and hurricanes can be computed by using vast linear models, requiring enormous numbers of computation cycles. The more robust of these phenomena (i.e. heartbeat but perhaps not hurricanes) can however also be computed using nonlinear models that are inherently more complex but at the same time less extensive. Much detailed knowledge is still required in such strategies, but the replication aspect of the computation is greatly reduced. In many respects the human body is organised precisely in this way: its different cells and tissues have virtually the same genome and follow almost identical algorithms in their computations with only some parameter values differing between them.

ITFoM’s biology-driven ICT models will run quite naturally on our digital computers. When digital computers were compared with the human brain, it was overlooked that the human body itself is also a computer, a computer much closer mechanistically to digital machines, since it is essentially molecular and granular. However, ITFoM will not model the brain as a computer or ICT machine. Indeed, the efficacy of digital computers to calculate human psychology are expected to be orders of magnitudes higher compared to their ability to model biological processes within the human body. Nevertheless, calculating the somatic aspects of the human brain is feasible and ITFoM will engage with this vis-à-vis Parkinson’s and Alzheimer’s disease, as well as other neurological diseases, as soon as mechanistic models can be formulated. Biology-driven ICT will also model the adaptive properties of networks in the human body and here interactions with adaptive hardware, modelled after the biological networks. In this way, this new ICT may begin to form a stepping stone for future brain-computer interfaces.

ITFoM modelling will be based on knowledge embedded in individual humans. They will thereby take into account the necessary detail. We know from the different response of human individuals to medicines or food, that such details are important for us. ITFoM will therefore NOT use the classical physics paradigm of simplifying away from the details, i.e. models will be informed to all detail that matters. Of course the models will share with the classical physics paradigm the attempt to leave out the detail that is not necessary.

**Multi-scale modelling in biology-driven ICT**
Our novel paradigms naturally overcome the ICT challenge of the required computation that will be relayed to the actual complexity of the organisation of the human body and thereby greatly reduced. Here we give some examples: Using the thermodynamic approach, only one equation will be needed for each macromolecular chemical species, rather than one for each individual molecule of that chemical species. That equation only needs to describe how the activity of the macromolecule depends on just a few (perhaps 4) concentrations. A number may be needed to identify which equation is relevant. The equation may require some 7 parameter values per gene product. For each organelle, cell and tissue type one needs the expression level of each (iso-) gene product, the kinetic parameter values remaining the same. If there are 200 such essentially different compartments in the body, this increases the above number to 207 parameters per gene product. When gene expression itself (including splicing) will also be modelled, this may ultimately be reduced to 30, leading to a total number of 1 million parameters per human individual. Because only a limited number of gene variants are frequent and because these are shared between groups of human individuals, such system will be modelled using multi-scale approaches.

Also the required computation power should be much reduced when using the biology-driven multi-scale approach. For most classes of gene products the relevant time scale of behaviour is more than a thousand times slower than 1 ms, and also here the ensemble, temporal, spatial and network compartmentalisation / averaging approaches may be used. For 30,000 gene products to be calculated once per minute in a quasi-steady state approach, for 3 important compartments, using 100 floating-point operations per gene product, should require only 150 kflops per human individual.

This envisaged drastic reduction of the computation power challenge, requires a strict organisation of the ICT for the virtual human however, which may not be achievable, attainable, or desirable (e.g. for reasons of confidentiality or IP). Moreover the above description of the organisation of human biology is an oversimplification, which may not suffice ultimately. Intracellular and intercellular heterogeneity, channelling of mass and information flux, epigenetics and microRNA-based regulation are only examples of the complications. Combinations of calculations that do use vast computer power should serve to examine where thermodynamic simplifications can be made. Novel improved hardware and superior algorithms that adapt dynamically as the action in the living organism shifts from fast and small to slow and large, intelligent hybrids between Bayesian and ODE models, new ways of making models learn from a mix of epidemiological, patient-opinion, DNA mutations and experimental kinetic data with variable uncertainty, are all likely to help make individualised medicine through ICT models possible.

There is a number of existing ICT approaches that could be confounded with biology-defined ICT. Biology-driven ICT that we propose here is neither bioinspired computing nor artificial life. Bioinspired computing takes characteristics of living systems, such as evolution and adaptation, as inspiration for new algorithms. These algorithms are inspired by biology, but do not replicate the real biology. ITFoM will go beyond that and look closer at the real human biology with variegated but unprecedented precision; only this will enable the calculation of how molecules acting on a set of precise macromolecular targets affect whole body function. ITFoM’s computing is not bioinspired yet “biodefined” by the actual human biology in a multi-scale model. ‘Artificial Life’, on the other hand, is the adventure of designing software that is self-replicating and thereby lives in computer systems. Again there, there is no tendency to mimic the actual functioning of the human body from molecules up in a multi-scale approach.

2.1.3 Objectives

ITFoM has a range of 10-year objectives, including:
1. Transformation of medicine and healthcare toward sustainable ICT-driven personalised medicine

2. Development and implementation of the revolutionary ICT required for objective 1

3. Development of complementary modelling platforms with different areas of strength and different planning horizons as the basis of the data and computation driven, personalised medicine of the future

4. Development of integrated molecular/physiological/anatomical reference models of human biology at multiple scales and levels of abstraction based on a combination of biomedical – omics and imaging data from samples of a limited number of individuals

5. Expansion of these reference models to additional individuals and data sets through data from the personal genome project (www.personalgenomes.org) combining molecular and medical data of up to 100000 volunteers

6. Individualisation of the reference models based on –omics, imaging and sensor data of patients in a limited number of disease areas/transversal topics (cancer, metabolic and rare diseases, heart disease etc.) as the basis of a novel, model driven personalised medicine, designed to the individual optimisation of therapy and prevention

7. Further development of analytical/ICT technologies currently used on small scale in cutting edge research applications to ultimately allow large-scale clinical use. Development of a limited number of new approaches to close critical gaps

8. Development of new mathematical, statistical and computational strategies to analyse and integrate highly complex data sets from individual patients, driving the generation of personal health to simulate and predict effects and side effects of different therapy choices for the individual

9. Optimisation of existing and development of new ICT to allow the construction of increasingly detailed and predictive models for an increasing number of individuals, scalable to potentially cover a major fraction of patients in European healthcare systems

10. Development of ethical, legal and computational frameworks to allow safe and effective handling of personal data on a European basis, and supporting the new paradigm of personalised medicine in medical practice and drug development and approval

11. Development of new interfaces and visualisation strategies, to optimise the interaction with medical professionals, as well as for the engaged individual with his own personal data

12. Development of new ICT necessary to enable the ITFoM vision to be scaled up and deployed, focussing on new biomedical and clinical interfaces, e-Health architecture, storage and processing, data gathering infrastructure, statistics and medical informatics, software engineering, IT regulation, security and privacy

13. Development of a Europe-wide infrastructure able to foster adoption of the new data and computation driven personalised medicine paradigm in European healthcare, feeding back (anonymised) treatment results to drive the improvement of modelling platforms for the benefit of patients

14. Continuing extension of the application areas of the novel personalised healthcare to new medical areas, with an increasing focus on prevention

15. Continuing development of new ICT, medical and societal mechanisms leading to improve healthcare, lower costs, and empower the individual to have more engagement in and control over his/her health status
In the ramp-up phase, the work will focus on the establishment and validation of the development of the new and key parts of the ICT and infrastructure, reference models, the demonstration of the individualisation of these models, the validation of the concept, and the demonstration of the integration of life science activities through new ICT on a limited number of use cases (especially cancer, rare diseases, metabolic and heart diseases).

2.2 Matching of Flagship proposal with Flagship concept

The IT Future of Medicine – ITFoM – Flagship initiative represents a visionary, highly interdisciplinary programme to enable explanation-based medicine on an individual basis. The concept of ITFoM is a science-driven approach using models that integrate all available relevant information concerning an individual patient to deliver a precise recommendation for the correct therapy at the right time.

ITFoM brings together European experts from a range of different scientific disciplines that are needed to achieve the ambitious aim to transform current medical practice into explanation-based personalised medicine that is based on integrated molecular/physiological/anatomical models of individuals in the healthcare system. This promise can only be fulfilled via a large-scale approach bringing together high-level expertise from all across Europe and beyond, combining leading centres in medicine, analytics, hardware and software development, pan-European (bio-)informatics infrastructures, statistics and machine learning, and systems biology and modelling. Key drivers are: our newly acquired ability to fundamentally characterise individual patients due to the fantastic progresses in analytical techniques, the accumulated body of knowledge on biological networks in health and disease, and major breakthroughs in ICT. The –omics, imaging and sensor technologies proposed here allow the collection of different types of molecular information and lead to a far more comprehensive understanding of a patients’ personal blueprint and current health or disease status, providing the basis for a new model-driven personalised medicine. These technologies will produce an increasing amount of data that will soon vastly existing simulations capacities in ICT, such as weather forecasting. Rapid developments in ICT in many areas (e.g. storage, speed, HPC, mobile computing) are an essential prerequisite for the success of the flagship initiative. In this respect the future of medicine has two essential key drivers: ICT and biomedical analytical technologies, that together will revolutionise medicine and allow the implementation of virtual patient models to be widely accessible to citizens in Europe and worldwide. This can be realised only through a federated effort between disciplines ranging from medicine to novel computing paradigms, from quantum sensors and new genome sequencing technologies, to new, energy-efficient computer hardware designs, and novel security and privacy concepts. Building on a solid European healthcare infrastructure, ITFoM will generate a productive and coherent ecosystem to overcome the fragmentation of the current healthcare sector towards a consistent standard for a personalised system. ITFoM has already mobilised experts from academia and industry all across Europe and created alliances with important initiatives worldwide that are essential for the realisation of the virtual patient.

Commensurate with the magnitude of problems we propose to address, as well as with the enormous impact on science, technology and society, is the large size of ITFoM, which needs to be realised through a federated effort of EC and national programmes and industrial involvement. By addressing a crucial global problem, ITFoM will drive European leadership and innovation to ultimately benefit healthcare for everyone. Implementing ITFoM as a pan-European initiative will cause Europe to gain leadership in the world in the development and implementation of the new tools and systems enabling the virtual patient for the optimisation of medicine. Europe provides the ideal environment for the ITFoM initiative: Europe comprises most of the leading experts in the fields addressed in ITFoM, but also the necessary infrastructures and scientific basis that have been
instigated over the last several years in a joint effort between member states and the European Commission. European industry is ready to take up the innovation that will be created within ITFoM, as well as the economic, political, social and healthcare system. The route to precision medicine via the virtual patient needs a pan-European effort, but also the linking with the rest of the world to incorporate the necessary expertise that complements the European landscape.

To achieve this, ITFoM will start from the strong momentum of many of the cutting-edge research areas currently funded in European (and worldwide) research, and will benefit from and provide benefits to numerous of the European infrastructures: in computing (including EDGeS, Enabling Desktop Grids for e-Science; EGI, European Grid Infrastructure), in medicine (including EATRIS, European Advanced Translational Research Infrastructure in Medicine; BBMRI, Biobanking and Biomolecular Resources Research Infrastructure), in systems biology (e.g. ISBE, Infrastructure for Systems Biology Europe) or EU research projects (PHGEN, Public Health Genomics European Network; SPIDIA, Standardisation and Improvement of generic Pre-analytical tools and Procedures for in vitro Diagnostics; HLS-EU, European Health Literacy Survey; etc.), and in uniting researchers from Europe, Canada, the USA and New Zealand, aiming for outreach on a wider global scale in the future. In combining forces from research centres, universities and industries including hardware design, software development, social media, health informatics, systems biology and instrument development, as well as including health sector and patient organisations, we will attempt to change the world by synergising the efforts of many individuals and of major organisations, while transforming them and improving healthy life in the process.

With its scope and ambition ITFoM will have a huge impact on science and technology development in Europe, but with a worldwide focus that leads the way global opportunities for innovative products and services related to ITFoM. This innovation strategy will lead to completely new opportunities for SMEs and will create the demand for highly qualified personnel in Europe. With the anticipated revolution of medicine, this flagship initiative will impact the whole healthcare system, ultimately refocusing it toward managing wellness rather than illness, and substantially improving the quality of life and healthy ageing.

ITFoM will generate substantial economic benefits through delivering major improvements in the healthcare ecosystem that currently costs approximately 7 trillion US dollars in the OECD and BRIC countries alone. In promising to halt or revert the otherwise unstoppable rise of healthcare costs driven by ageing populations, ITFoM addresses some of the most vexing and to date unsolvable problems of European governments: the budget crises in many European countries, driven by, among others, the enormous costs of the healthcare systems, as well as early retirement driven by unresolved health problems in older people.

2.3 Description of methodology, strategic research roadmap with milestones, and work plan for ramp-up phase

The problem

The development of information and communication technology has been fueled by various large-scale projects, including large physics and astronomy. Medicine and biology, in contrast, have profited relatively little from the increasing potential of ICT with the exception of large-scale genomics research. This is in contrast to the big problems that we are facing in these areas, ranging from ineffective healthcare systems that swallow an increasing fraction of the gross national products in ageing societies, to the need to feed a continuously increasing population in the face of demographic change and possible lifestyle changes.
In this programme, we specifically address the problem of human health. The solutions that we will develop – a combination of high-resolution analytical techniques to characterise the biological processes in an individual in detail, combined with new computer modelling approaches and social innovations – will nonetheless find applications in many other areas addressing the optimisation of biological processes and beyond. Medicine, at the moment, treats patients as parts of large, often homogeneous groups. In the case of cancer, we currently do not treat an individual cancer patient, yet a member of a group (of lung cancer patients for example), for which one specific therapy has been statistically shown to be somewhat more effective than others, even though this – often quite expensive – therapy might not benefit the individual in question yet could produce serious side effects.

The enormous costs of therapies are an important component of the huge increase in healthcare costs worldwide, soon expected to reach 20% of gross national product in the USA. As societies age, this load will continue to increase further, since personnel costs, which are already a major component of overall healthcare expenditure, will continue to increase further as fewer young people will have to support increasing number of older patients.

The solution

To be able to move our healthcare system to treat patients as individuals rather than as members of large divergent groups, we propose to develop a new, data and computation driven individualised medicine of the future, based on integrated molecular/physiological/anatomical virtual patient’ models for the benefit of individual people in the healthcare system. The establishment of such models methodologies is now becoming possible due to the enormous progress made in available analytical techniques, particularly in the so called ‘-omics’ technology areas, but also rapid progress in imaging and sensor technologies as well as the continuing developments in ICT, which, over the coming years, could make the virtual patient a key component in medicine and disease prevention.

Implementation plan

The implementation plan of IT Future of Medicine is based on 7 workpackages, as well as a number of use cases, with separate ICT development requirements, organised in a matrix structure.

The workpackage dimension in particular subdivides the overall project into technology areas (medicine/health informatics, analytical techniques/analytical informatics, ICT systems, ICT networks, statistics/machine learning, integration/modelling and coordination/implementation).

We anticipate four phases of implementation: the establishment of reference models (years 0 to 3), expansion to additional individuals and additional phenotypes in collaboration with the PGP project (years 1 to 4), establishment of disease specific models (years 0 to 10), small-scale tests in collaborative projects (years 1 to 10 for different disease areas), testing in pilot projects (years 3 to 10, depending on the disease), and identification of organisational changes (years 0 to 5) and translation into the healthcare system on a national and disease specific basis (years 5 to 10).

Use cases will initially focus in particular on cancer and metabolic disease, for which rapid progress is anticipated. Other important application areas will however be developed in parallel, based on close interactions with externally funded efforts.

These implementation phases will be prepared and supported by ethical and legal guidance in parallel, as well as public engagement programmes to ensure broad public acceptance of novel ICT concepts by all stakeholders involved in healthcare and public health.
**WP1 Medical Platform**

WP1 will provide continuous input of users’ needs into the ICT R&D pipeline of ITFoM and will prepare the implementation of the novel ICT solutions developed by ITFoM in healthcare and public health. WP1 will address, in particular, the provision of data and materials required to establish the reference models, will develop data management solutions for medical, life style and environmental information, define specific needs of user interfaces, establishes the link between ITFoM and the medical community and the citizen for the application areas, and will explore different mechanisms and concerns relevant to implementing the systems to be developed here as standard components in European and world wide healthcare systems. In addition, the medical platform will handle public engagement, health literacy and ethics tasks including data protection and privacy. All these duties and responsibilities include comprehensive interaction with other WPs, particularly WP2, WP3 and WP4.

The core elements of WP1 are to

- Provide the user needs form healthcare and citizens to ITFoM
- Define and control quality of health-related, life-style and environmental exposure data used for modelling
- Contribute to the development of data management and user interface solutions
- Address data privacy issues of ITFoM and its ICT models
- Enable the implementation of the ITFoM solutions into the healthcare systems by proving the development of evidence through use cases and pilot projects, preparing the necessary ethical, legal, economic and regulatory frameworks, and by setting up sustainable tools to support the constant transfer of knowledge to and from the health area.

**WP1.1 Bridge to current medicine and stakeholder engagement**

This sub-workpackage will focus on current medical practice, and acts as primary interface to the healthcare word. It is clear, that ITFoM cannot deal with all the hospital and medical research centres in Europe. Therefore we will use a small number of leading medical centres as the hubs and the centre of reference to the medical world. Such centres of excellence will join ITFoM at an institutional level in the context of strategic partnerships to (a) bring several research groups of the institution, and (b) act as a bridge to a medical domain. WP1.1 defines the selection procedure for these institutional partners (this group may change over the lifetime of the flagship project), provides the communication / information exchange platform and deals with all organisational tasks. There will be a very close cooperation with all the other sub-workpackages, especially WP1.3 (Reference data set) and WP1.7 (Coordination of use cases). The stakeholder engagement will be done through the involvement of the diverse scientific communities (European and international scientific associations), the private industrial sector and governmental and regulatory bodies related to ICT and biomedical research. ITFoM will set up liaisons with the leading research groups and projects by establishing a joint project coordination platform. The involvement of European patient organisations will be based on solutions developed in BBMRI. The involvement of the industry will be made partly through their association bodies, namely the European Federation of Pharmaceutical Industries and Associations (EFPIA), the European Federation of Biotechnology (EFB) and the European Diagnostic Manufacturers Association (EDMA) and the European Databank on Medical Devices (EUDAMED). Among the governmental bodies involved will be the European Health Technologies Agencies on regional and national level. The interaction with public bodies and political decision makers will promote regular funding streams directed to ITFoM related disciplines and activities.
WP1.2 Ethical and Societal aspects and associated public engagement

This sub-workpackage will explore the ethical and societal aspects of ITFoM, regarding both the underlying concept and its implementation, in close interaction with an ethical advisory committee, combining ethicists, lawyers, medical professionals and representatives of patient organisations.

The ethics of biomedicine has expanded, from an initial focus on the clinical encounter and the health-professional relationship, to include issues that are beyond the clinic, involving the impact of scientific and technological developments, in the postgenome era, for example, and a task for public health. Contemporary and future biomedicine projects such as ITFoM, however, are increasingly concerned with collection, storage, interpretation and use, of ever larger quantities of data, and it is here that there is a clear interaction with ICT ethics. Key ethical concepts invoked include privacy, and rights to know and not to know, in addition to issues of autonomy, justice and equity. It has become clear from the genomics context, however, that ethical concepts co-evolve with science. It is at the interface of ICT and biomedicine that we can expect further co-evolution, in the light of the potential implications for patients, professionals and society of ‘big data’ on the one hand, and the increasing ‘personalisation’ that the data make possible, on the other. Also important legal issues come up here such as the implications of the ITFoM model for access to social services, such as healthcare and life insurance. For the grand scale processing of sensitive data suitable representation formats (languages, models) and exchangeable standards for access restrictions both on the general level (institutions, countries) as well as on the individual level (individualised) consent need to be developed which allow policy based automatic enforcement of ELSI requirements. In collaboration with WP1.4 and WP7, WP1.2 will serve society with front running developments towards a new and dynamic ethics that is societally responsible, in line with ethics traditions, yet up to date in term of ICT empowered systems medicine.

Ethical monitoring and guidance will be evidence-based using scientific data on public perception collected during the running period of ITFoM. The aim of the ITFoM project is to contribute to a process of ‘revolutionising our healthcare system’ by creating a breakthrough in individualised medicine. Such fundamental changes mean that both the material conditions and the social norms that govern the conduct of a domain are altered in ways that will constitute deep-going challenges for the existing ways of understanding and operating medical research and healthcare with extensive implications for researchers, scientists, medical personnel, patients, pharmaceutical industry and society in general. Typically, revolutions are carried out outside the existing legal forms, outside of the traditional material basis, and outside the existing cultural normative system. If they happen they have deep impact and their central pre-condition and condition sine qua non is a good deal of ‘outside the box thinking’ as traditional concepts and methodologies limit the potentials for apprehending the transformations that are occurring. Informed consent and privacy are two examples of ways in which ethical thinking has been revised in the past fifteen years. In the early nineties the ethical considerations at stake included, primarily, individual autonomy, confidentiality, and beneficence. Since that time, with an increasing emphasis on public health, there has been a move towards solidarity and equity as important principles to consider. The second not unrelated, trend is towards rethinking the concept of privacy. It has been claimed that even if we could rely on public institutions to adhere completely to regulations on data protection, there is still a need to look at the issues in different ways. In the genomics context, there has been a suggestion that emphasis on privacy should be replaced by the concept of open consent. The concept of open consent is used in the context of the Personal Genome Project, which aims to build a framework for the development and evaluation of personal genome technologies. The advent of ‘celebrity genomes’ where key figures willingly publish their genomes to illustrate a different attitude towards privacy and disclosure has caused some concern. In the current context, however, arguably the most ‘revolutionary’ developments will arise from the interface of ICT ethics and biomedical ethics. We
can expect the emergence of IT tools (e.g. mobile devices and Apps) which might “enhance” ethical, legal and social rights of research and data subjects. These applications are being designed to serve a dual purpose. They enable participants to concretely exercise/enforce their ethical-legal rights. But they also allow for self-reporting by participants, thus contributing substantially to projects like ITFoM.

**WP1.3 Reference data set**

The reliability of any computational modelling approach will critically rely on the quality of data used for modelling. Several international projects, such as the International Cancer Genome Consortium (ICGC) or the 1000 Genomes Project have produced large high quality data sets. However, these data are restricted to certain organs or health/disease conditions and would not be suitable for integrated modelling of biological systems across different organs and disease conditions on an individual basis. The ITFoM reference data set will consist of unique biological samples representing most tissues/cell types of man as source of molecular data as generated by WP2 together with detailed related medical, lifestyle and environmental exposure data. This reference data set will be established by interacting with leading medical centres in Europe (defined by WP1.1). Based on work performed in BBMRI and the FP7-funded large integrated project SPIDIA that develops European standards for sample pre-analytics for molecular diagnostics together with WP2 and WP6.7 we will define consensus (‘gold’) standards for the collection procedures of samples (SOPs) and for clinical data, synergising the work carried out by leading consortia and industry. This “gold standard” will support and help drive long-term technological developments, at least for a ten years period, and will be constructed such that it will evolve with those developments. Accordingly, the reference collection will form a sustainable foundation for all analyses and modelling work done in ITFoM. Samples to be considered for the reference data are tissues, serum, plasma, peripheral blood mononuclear cells (PBMC), urine (repeated sampling), saliva (repeated sampling) and the microbiome (repeated sampling). If ever possible from one individual several tissue types and organs (brain, heart, kidney, lung, oesophagus, stomach, small / large intestine, liver, gall bladder, pancreas, salivary gland, thyroid, adrenal gland, testis, ovary, mammary gland, lymph node, spleen, tonsil, bone marrow, skin, cartilage, bone, skeletal muscle, smooth muscle, adipose tissue) will be collected. WP1.3 will consider both the technological, medical, epidemiological, as well as the ethical and legal aspects of collecting biological samples and the related medical data. The final reference collection will comprise at least 500 individuals, each with a comprehensive data collection of molecular and biological processes and in addition the medical, lifestyle and environmental exposure data in an unprecedented level of detail and standardisation. The data sets derived from consenting individuals, and a common denominator data set will be (i) continually extended as the patient goes through the treatment and follow-up cycle, and will (ii) be made available to the public in an anonymised fashion.

**WP1.4 Medical, lifestyle and environmental exposure data management**

Computational modelling within ITFoM is based on two major categories of data with different ethical, legal, and social implications (ELSI). Type 1 data are related to an individual and comprise a huge variety of medical data (including imaging data and physiological data) as well as data on lifestyle and environmental exposure. Type 2 data are molecular data that are generated by analysis of human biological samples using a variety of –omics technologies (data addressed in the analytical platform). Issues related to the management of type 1 data are addressed in WP1 since these data are typically collected and stored in the healthcare systems and have to consider special data protection and privacy requirements. Furthermore, standardisation of data quality including phenotype description, multi-language interoperability, and heterogeneous national ethical and legal frameworks for medical data are specific challenges to be addressed. All technical tasks will be done in close cooperation with WP3, WP4, WP5 and WP6. In WP1 the focus is on analysing and
representing the ELSI requirements for the exchange and processing of patient related data while the other WPs develop technical solutions for the processing of data (e.g. enforcing access restrictions). WP1 then will implement the tools and the interfaces in the information systems of the participants. Additionally, WP1 will develop technical solutions for augmenting data with provenance information according to the provenance exchange standards developed in WP4 and WP6 to enable data quality management along the whole data pipeline. Together with WP1.2 we will address the fact that large amounts of data related to healthy individuals and patients will be collected and stored. In addition to ICT challenges this brings up a number of ethical-legal issues and requires a co-evolution of biomedical ethics and ICT ethics which in the past have developed separately.

**WP1.5 e-Health interfaces**

WP1.5 will address the interface between ITFoM and e-Health to ensure, that data and models resulting from ITFoM can be efficiently accessed, understood, appraised and applied within the care process by all actors including scientists, clinicians, patients and citizens (e-Health Literacy). As more clinical information becomes available through e-Health systems, this will help to improve the models, provide input into the epidemiological analyses, and introduce efficient community platforms including mobile-health solutions (m-health), social media, social networks and (digital) health thereby helping to introduce new ICT into the healthcare systems. WP1.5 will address human computer interaction (HCI) challenges for different stakeholders (scientists, doctors, patients) including topics from information visualisation and augmented reality. New interfaces are necessary to augment medical data by different stakeholders and combine different data sets with global records. Clinicians and patients will also be involved to influence the design of the interfaces and ITFoM’s (WP3.6, 6.2) new visual languages. Application scenarios will include (medical) decision support systems, visual analytics, personal health records, search interfaces and awareness, with large potential for the associated industries. For instance, where interfaces are developed for mobile devices including medical sensors we will develop new HCI paradigms connecting them to WP6 computation through WP3 technologies. Sensors can capture simple responses like pace and pulse up to very general data such as logging of the daily food intake or specific measurements, e.g. data about insulin injections or concrete metabolites (see WP6.3 ICT models). When dealing with sensor data it is crucial to consider privacy and security issues, which will be done together with WP3.8. WP1.5 will also work in close cooperation with WP3.6 (interactive 3D visualisation) and use the data/model/tool aggregation of WP6.7/6.8/6.9 for prioritising what/how/when to show data on a screen and in which context as well as with WP1.6 and WP1.8 regarding the integration into healthcare systems.

**WP1.6 Economic and regulatory aspects**

To improve the integration of new technologies in the area of health and to decrease the time from final development to implementation and potential reimbursement, health product development strategies should not ignore the need to generate data that will be used to inform reimbursement and pricing decisions. Data requirements to support product assessments are evolving, heightening the need for early dialogue with decision makers to ascertain what evidence they are looking for in the decision-making process on a particular kind of technology. Taking into account the different areas of medical, economic, legal, social, ethical and organisational issues of a new developed technology, the decision whether and when to invest in a new technology is getting more complex and involves more stakeholders. In assessing medical issues, it is important to realise the differentiation between the regulatory (market access) and reimbursement needs (service provision). Based upon this scheme, this sub-workpackage will focus on the barriers and enablers of introducing technologies developed under ITFoM into the different European healthcare systems and beyond. It will assess and tackle the regulatory and reimbursement issues on the global, European, EU Member
States, regional (e.g. NUTS 2) and local levels to prevent problems in the implementation of beneficial new ICT products in healthcare. This will include the identification of organisational changes needed for an active citizen-centred and need-driven health services environment based on concepts such as health literacy. The relations among the different stakeholders that interact on the final successful implementation of innovations should be taken into account. In this sense Health Needs Assessment (HNA) and Health Technology Assessment (HTA) are part of the bridge between producers of innovations and providers of services. They have to deal with the requirements of the healthcare systems to timely, effectively and efficiently incorporate those innovations by increasing awareness, reducing potential harms and produce added value for a system and the users. The LAL (Learning Adapting Levelling) model (Lal et al., Journal of Translational Medicine 2011, 9:207) will be tested and used as a framework towards the goal of implementation of the in ITFoM developed products for healthcare systems from the first idea to application. The core of the LAL model framework resolves around a relative parallel initiation of the Technology Transfer activities and the Public Health Assessment Tools (HNA, HTA, Health Impact Assessment (HIA)). HTA and HNA tasks among others can develop independently without forethought to the LAL model but consequently fall under the framework of the LAL model due to the concept of early involvement beside the point that the LAL model encompasses all these individual components.

**WP1.7 Coordination of Use Cases**

ITFoM considers health and disease from the biological systems perspective, which describes the functioning of the human body (human physiology) depending on networks. This networking occurs at the molecular, cellular, tissue and whole body-level, and diseases correspond to network malfunction. Most human diseases are thereby multifactorial by nature, i.e. co-determined by multiple genetic and environmental factors. Because the networks of life are all connected, different diseases are often affected by the same sub-networks, although to different extents. Detailed modelling of the integral network inside the human, as condensed in the form of an ICT model for each individual, should ultimately deliver a completely new level of understanding health and disease. To bring ITFoM to much earlier fruition whilst working towards this ultimate aim, ITFoM will put in place a number of ‘use cases’, the activities and results of which will be integrated (by WP6-10) into human ICT models of ever increasing connectivity, resolution and quality.

Each use case will be driven by a medical need defined by WP1.7 and will be coordinated with and supported by all ITFoM workpackages in a matrix type organisation. Starting with the definition of medical needs, use cases may begin to develop / improve one or more specific models with the help of the reference data sets of WP1.3 in order to address a set of medical questions. WP6.6-6.9 provides the infrastructure to integrate maps, data, models and tools within and for a use case. This will be especially important, as the models for any use case will implement different modelling strategies, such as those of the watchmaker’s, engineer’s, mechanic’s and learner’s categories and combinations thereof (see WP6.1-6.5).

ITFoM will focus on use cases within in the following categories:

- Cancers
- Diabetes and metabolic diseases
- Cardio-vascular diseases
- Immune reactions and immune-mediated diseases
- Infectious diseases
- Diseases of the central nervous system
The following crosscutting themes will be addressed specifically

- Rare disease
- Health and ageing, e.g. osteoporosis, neurodegenerative diseases and other age-related diseases

Every use case will need beside a common set of basic data a specific reference data set (WP1.3). Therefore we will also consider subgroups, (e.g. age, gender, ethnicity) in the composition of the reference data set. Each use case will contribute to the set of ICT challenges and provide input from the involved user community to research tasks in modelling, HCI (human computer interfaces) and data visualisation. Furthermore uses cases serve as first level validation platform for the ICT solutions developed within ITFoM. In the Annex we give an overview on the requirements for the reference data set, model categories, research groups, application scenarios and corresponding interfaces.

**WP1.8 Evaluation and service implementation**

Within the ITFoM project two different kinds of working fields will be implemented:

a. development and design of a proof-of-concept and proof-of-principle of use cases for computational modelling in the context of diagnostic and therapeutic support and development and its introduction into health systems (LAL model). Experiences obtained in these use cases will feed back in the R&D strategy of ITFoM.

b. testing of pilot projects for implementation in healthcare in specific countries/regions in line with “Going Local 2.0” of the Digital Agenda for Europe. Using Structural Funds it will be validated whether the application of new strategies under optimal conditions can achieve the wish reducing “time to customer”. In preparation for a widespread adoption in European and world wide healthcare systems two main examples will be considered: German (Bismarck-like) and Spanish (Beveridge-like) healthcare systems.

These pilots will be based on the best practices determined in the European funded project PHGEN I and PHGEN II, HLS-EU, EUnetHTA, EuroScan international network and INAHITA. WP1.8 provides a unique opportunity to generate evidence to support cross-border healthcare decisions at the European level.

**WP1.9 Epidemiology in clinical medicine and public health**

The modelling of the ‘virtual twin’ will pose two essential challenges to epidemiology:

- First, the integration of role of prognostic factors, health management and screening activities that would impact the predictive modelling of treatments developed within the virtual patient
- Second, the evaluation of efficacy of treatments based on individualised medicine.

The ‘virtual twin’ will require several key epidemiological data related to prognostic factors, health management and screening activities. This sub-workpackage will collect and evaluate, by systematic reviews of the literature and existing knowledge, the role of these factors in the prognostic of key major diseases. These elements would be further used in the Virtual patient modelling. Classical epidemiology and statistical methods used in clinical trials will not be applicable to the new treatment scheme that will be derived from the ‘virtual patient/virtual individual’. This ‘virtual patient/virtual individual’ based medicine will pose a challenge in the evaluation of efficacy of treatments and new statistical methods and epidemiological designs have to be developed. This sub-workpackage will enforce a long-term research program on new epidemiological and statistical
methodologies applied to individual medicine. This research effort will enable to provide the tools for evaluation of clinical medicine in the new scheme of treatment using the Virtual Patient. The new epidemiological designs and statistical methods will be developed by group of experts, and several test studies will be conducted. In addition, clinical epidemiological studies will provide key information of the costs aspects of individualised medicine and prevention (see also WP1.6). Finally, this sub-workpackage will provide expertise for evaluating the impact of major changes in healthcare services (e.g. screening, treatment, patient management) that occur in key diseases during the conduct of the project. We will make a systematic review of prognostic factors, patient management and screening activities, develop epidemiological designs and statistical methods for evaluation of personalised medicine and monitor changes in the healthcare services that impact modelling for major diseases (with WP1.6).

**WP2 Analytical Platform: Analytical Technologies and ICT**

WP2 has three main goals:

1) To conduct molecular analyses with maximised ICT support, on clinical samples provided by WP1, via sub-workpackages (SWPs) focused on different analytical modalities

2) To monitor and actively contribute to the further development of systems medicine technologies by empowering the latter with revolutionised, biology-directed ICT

3) To develop and implement ICT components that (i) integrate the information generated by WPs 1 and 2, into ‘virtual-individual’ data collections, (ii) enable the recognition of the multimodal information components, (iii) are optimal input to the model integration in WP6 into virtual individuals, and (iv) make best use of the known properties of the biological systems under investigations as contained in the data-models of WP6.1-6.3.

The partners in the different SWPs of WP2 will promote the development of both established and emerging technologies that can contribute to the aims of ITFoM. All across the multiple relevant scales of the human physiology (–omics, imaging, monitoring and diagnostic technologies, etc.) that produce the type-2 data (see WP1.4) of WP2, the analytics will be greatly enhanced by adding new ICT components, co-developed with WPs3, 5 and 6.9. This will drastically increase the signal to noise ratio in the analytical techniques and focus the latter on the aspects that are important to measure.

The technologies extend over multiple scales, i.e. from molecules to cells, tissues, organs and the whole body. At each of these scales, WP2 will also help WP1 in transferring existing and emerging ICT technologies from biomedical research to routine clinical practice by empowering these technologies with appropriate ICT solutions (WPs3-5). In collaboration with WPs3.1 and 6.9, SWPs of WP2 will also provide the ICT expertise to analyse a small number of highly characterised samples and type-1 reference data sets provided by the medical WP1, to add corresponding type-2 data to generate a reference data set underpinning the reference type-3 data models of WP6. WP2 will also participate in all the use cases (WP1.8 and WP6.10) and will be involved in the validation of the model predictions and results from the reverse engineering approach (e.g. WP6.4). The interface to WP3 will foster the development of hard- and software solutions (e.g. for storage, computing, rapid data transfer, and automation in the data processing pipeline). Gold standards (WP1.3 and 6.7) will be implemented for the sampling and sample preparation as well as for sample analysis. With WP1, WP2 will drive ICT empowered technology development according to the challenges arising from transfer into clinical practice and by requests by the WP6 integral models.
WP2.1 From molecules to cells: –omics technologies

WP2.1.1 Nucleic Acids

DNA sequencing is rapidly becoming the universal approach to analyse genomes, epigenomes, transcriptomes and metagenomes, driven by the enormous decrease of cost and increase of throughput of DNA sequencing. The technologies that are already available or expected to emerge differ in key aspects from what was there only three years ago (e.g. amount of required starting material, read length, type/rate of errors, equipment, personnel, running costs, samples preparation, run time, and downstream computational requirements). Particular challenges remain associated with the sequencing of very small amounts of material (e.g. cell-free DNA in serum or urine, single cell DNA/RNAs), and the direct detection of base modifications. The costs of sequencing a human genome went down from 50 million€ in 2007 to less than 4000€ in 2010 with a ten-fold increase in coverage. If current trends continue, we would expect to see a per-genome reagent cost in the region of tens of euros within the next ten years.

Genomics. The information in the genome provides key insights in unique biological features of the individual and, in oncology, of the tumour (SNVs, indels, translocations, deletions, ploidy changes, LOH). Analysis of free nucleic acids in serum and of the genomes of small number of cells will offer additional diagnostic possibilities to provide input data for individualisation.

Epigenomics. The epigenome informs on normal and pathologic activity states in cellular time and space. The relevant information is captured by measuring changes in chemical modifications at DNA and chromatin level, irrespective of whether these are inherited or acquired through exposure or life-style. Pathogenic modifications can be the hallmark of certain cancers and other common diseases.

Transcriptomics. The transcriptome provides essential information about the structure and abundance of all transcripts in an allele specific manner, the abundance of microRNAs and of other regulatory or structural RNA species. In interaction with WP6, we shall develop ICT to enhance the predictions of the transcription activity of cells and the structure and abundance of their proteins.

Metagenomics. DNA sequencing also serves to efficiently characterise complex microbial communities associated with the human (the human microbiome). Microbiota at several body sites have a known role in disease, health and well-being. Identification of the quantitative and dynamic constitution of the microbial populations will be enhanced by the biology-informed ICT we shall develop in collaboration with WPs3.1 and 3.2.

In situ sequence analysis. We plan to develop a spatially resolved deep sequencing for e.g. pathology slides of cancer samples to generate cellulary resolved information on the transcriptome, and - through proximity ligation assays - on proteins, protein modifications and protein complexes.

Deep sequencing of B and T cell immune repertoires. Since the state of the immune system reflects disease processes, we will develop a comprehensive deep sequencing approach to generate complete information on immunoglobulin and T cell repertoires in B and T cells of every patient. We will develop biology-inspired ICT methods to enhance the data analysis.

WP2.1.2 Proteins

Protein abundance, modification and interaction are key to understanding biological processes, e.g. of signalling pathways as a very important basis for the ‘virtual patient’ model. Partly because the ICT required for proteins analysis is more complex than that for the analysis of nucleic acids, it has been more challenging to improve the technology towards increased throughput, precision, sensitivity, robustness, and reproducibility from an extended range of sample sources. We will focus on
extensive proteome analyses, covering the systematic analysis of proteins, protein variants and modification states, as well as functional interactions of sets of proteins.

**Proteomics - Mass Spectrometry methods.** Most current proteomic measurements are based on mass spectrometry where peptides that uniquely identify a specific protein are fragmented to generate a distinctive fingerprint pattern (fragment ion spectrum) that uniquely identifies the peptide in question. Reference spectral libraries have been developed for 99% of the human proteome, and these libraries constitute specific assays that now allow quantification of human proteins with high accuracy and reproducibility in biological specimens. Targeted mass spectrometry approaches and emerging techniques such as SWATH-MS will be used to achieve high proteome coverage. These techniques permit also the analysis of large number of samples, wherefore we shall develop biology-driven ICT.

**Proteomics - Affinity-based methods.** New assay designs (e.g. the proximity ligation assay) greatly improve specificity and sensitivity of detection, and tools are emerging for highly parallel protein measurements down to the detection of single protein molecules. Systematic generation and use of antibodies to analyse protein distributions on histology slides and in high-resolution in tissue culture cells provides necessary information and localisation on protein expression patterns. Programs for producing sets of recombinant affinity reagents fulfilling defined specifications and standardisation of protein detection methods (i.e. SRM), are important prerequisites for extensive protein analyses in research and diagnostics, also in the context of model-driven experimental design.

**Proteomics - Kinetics.** High throughput assays for enzyme kinetics in vitro and in vivo at the single cell level are still missing, but are essential to elucidate the functioning of living cells and complement the predictive power of molecular models. To transfer this technology into clinical practice, advances in imaging, sensor, and cell sorting technologies as well as in the ICT (including automation) are necessary.

**WP2.1.3 Macro-molecular assembly and metabolites**

Analytical techniques for metabolites and other small molecules in human samples complement the spectrum to cover all levels of biological information. Similar to proteins, metabolites and other small molecules provide a detailed overview of the biological and disease state of an individual. The challenge of measuring these molecules robustly with minimal invasiveness is a challenge that can be alleviated by the ICT we shall develop, focusing on what may be expected on the basis of other data-models (WP6).

**Metabolomics.** NMR measurements of metabolites in biofluids or tissue and cell extracts, or, by solid state NMR, in intact tissue samples, provide an effective approach for systematic metabolite analysis. GC- and LC-MS based measurements provide complementary, highly sensitive metabolomics data. In collaboration with WP6.8 we shall develop model-informed ICT that extracts from the raw metabolomics data the information that is most germane for diagnosis. Using different ICT we shall facilitate the transfer of metabolomic analyses to routine clinical practice enabling individual mechanism-informed phenotypic metabolic fingerprints, providing the basis for model individualisation by ‘reverse engineering’.

**Other ‘-omics’ approaches - Glycomics, Lipidomics.** The capacity to synthesise a wide range of carbohydrates and glycosylated proteins, combined with approaches to generate and use selective binding reagents offers a yet unexplored window into disease processes in an individual patient. Similarly underdeveloped is the use of lipids as indicators for disease and physiological state.

**High throughput protocols to identify drug target specificity.** After binding of proteins (including both purified disease relevant proteins and protein extracts from disease relevant cell lines/tissues) to drug beads and elution of bound proteins by excess drug, proteins with drug binding
characteristics can be identified by mass spectrometry. This will be coupled to ICT models at the molecular scale to quantify the binding characteristics of a drug on a patient-by-patient basis.

**WP2.1.4 Cells**

With recent technological advancements, individual cells will become a new target for routine clinical applications, providing an important complement to biomolecules and a unique diagnostic tool, but requiring biology-inspired new ICT to beat the noise/signal limitation.

**Mutanome analyses.** To analyse the functional effects of common new protein variants suspected to be involved in disease processes, (inducible) wild type and mutant DNA constructs will be introduced into a set of recipient cell types, coupled, where necessary, with shutting off expression of the wild type gene using e.g. shRNA constructs. Cells in different induction states will be characterised systematically by –omics and phenotype analyses (high content screening microscopy, FACS sorting, etc.) providing essential input into the construction of mechanistic models.

**WP2.2 From Cells to organism: imaging technologies, diagnostic and monitoring systems**

**WP2.2.1 Cell–omics**

With multiparametric quantitative assessment of single cells using mass spectroscopy-based cytometry, a large number of different cell types can be monitored (e.g. T and B cells, monocytes, dendritic and stem cells) in a small blood sample, including information of their maturation and activation state. In addition new technologies (e.g. HLA ligand exchange) in identifying large panels of epitope specific T cells will be applied to mass spectroscopy-based cytometry. This allows rapid and high-throughput identification of T cells specific for a disease or pathogen (e.g. tumour, HCV). *Ex vivo* cellular imaging cytometry will complement the approach using the latest generation of the Amnis Image stream (high throughput multispectral cytometric analysis with high content images). Using new ICT, these data will be integrated with metabolomic assessment of cellular products.

**iPS–omics.** Induced pluripotent stem cells (iPS cells) represent a surrogate material for (inaccessible) tissues of the patient. iPS cells deriving from skin or other somatic cells will be differentiated into a number of different cell types with subsequent analysis via the –omics technologies. In addition to contributions to diagnostics, iPS or stem cells will enable cell based therapeutic strategies, but this will require advanced ICT to provide for robust characterisation of the cells, which is what we shall develop.

**WP2.2.2 Imaging Technologies**

A new challenge arising from technological innovations will be the study of the whole body. Different approaches will contribute to a more comprehensive understanding of the body functioning. Through new ICT, we shall focus on empowering imaging by immediate interpretation in terms of units of biological functions (with WP6.1).

**Tissue imaging.** Immunohistochemistry and fluorescence *in situ* hybridisation techniques and their many variants are established as research and clinical tools to image the distribution of proteins and nucleic acids in cells and tissues. Improved molecular probing integrated with improved data and image analysis ICT will permit even more specific protein, DNA and RNA features to be imaged at cellular and subcellular resolution in tissue sections, fine needle aspirates, fingerprint approaches or cell smears. It will then also be possible to target features such as mutant genes and transcripts, interacting sets of molecules including complexes of nucleic acids and proteins, as well as variants arising through chemical modifications of specific nucleic acids or proteins.
**Body imaging.** Quantitative analysis of PET or PET-MR images provides information of the localisation of a certain active substance in the body. The new integrated PET-MR scanners provide new possibilities beyond those of each individual technique. We shall address the challenge to generate integrated complex multimodal, multi-parameter, multi-resolution 3D+time data. The new developments in *e.g.* detection techniques and stochastic algorithms will improve both resolution and sensitivity by an order of magnitude or more, with reduced calculation cost. X-ray tomography provides volumetric information about the body or a particular organ. When used conjointly with the injection of a contrast agent and with the dynamic model-based analysis that we shall develop, it enables efficient cardiovascular, pulmonary and tumour imaging.

**WP2.2.3 Monitoring and diagnostic technologies/sensors**

The combined use of new families of diagnostic monitoring systems and ICT empowered data analysis will be necessary in (i) *in vitro* systems to provide precise and complex analysis, follow-up and diagnostics – they will be utilised in hospitals, in medical analysis laboratories or at home, under the control of professionals, (ii) *in vivo* systems to answer to specific needs of continuous monitoring, thanks to low cost and disposable systems, (iii) *in vivo* systems to be able to follow super critical parameters, as well as to monitor specific populations of patients having a limited access to the medical care (ageing population, dependent people, etc.).

**Monitoring and diagnostic of body fluids and gas.** Monitoring in urine, blood, saliva, sweat and breath enables the application for portable devices not requiring sophisticated transducing tools such as fluorescence microscopy, surface plasmon resonance or colorimetry. Such devices will lead *e.g.* to the ability to monitor body fluids in real time, either in domestic applications (*in vitro* and *in vivo*), while direct recordings can be interfaced to central systems for point of care monitoring as well as for medical surveying.

**Monitoring of the environment - Exposomics.** To further improve the prediction power of the model, information on life style and environment will be integrated in addition to molecular/physiological/anatomical information. Exposomics refers to the identification of effects of previous exposures to harmful agents in the environment (*e.g.* radiation or pollution) based on measurements in biological material of global sets of biomarkers of exposure. Systemic metabolic profiling in body fluids can contribute through the identification/investigation of biomarkers for exposure.

**Personalised diagnostics approaches for healthy living.** In order to provide individuals with a mean to optimise their behaviour towards nutrition and exercise and to live a healthier life, it is necessary to integrate behavioural data with metagenomics and potentially other medical and biological information. Metabolomics and cellomics represent powerful diagnostic tools that provide a snapshot of an organism’s metabolic and phenotypic state through systemic analysis. In addition, unique methodologies and algorithms will be developed to interpret metagenomics data in relation to behavioural inputs such as nutrition, exercise, sleep or other.

**WP2.3 Transversal technologies**

**WP2.3.1 Generic enabling technologies**

The underlying challenge will be to develop “3rd generation systems” with auto-calibration and homogeneous context acquisition. To reach this goal, critical ICT bottlenecks need to be addressed, impacting on most of the analytical technologies described above.

**Microfluidic development and standardisation.** The activities cover the standardisation of the microfluidic interconnects, the management of the embedded reagents, the qualification of the materials used in the devices and finally the identification and development of generic microfluidic
devices. This will allow to develop and propose complex microfluidic systems, integrating heterogeneous technologies and to be able to answer to the needs of the healthcare market.

**Miniaturisation and integration.** In order to decentralise sampling and analysis, the key issues will be the miniaturisation (even down to nanotechnology) and cost reduction of the instrumentation itself, covering a broad range of analytical technologies. This includes optical, electrical, mass spectroscopy, the integration of the information, and the implementation of model driven experimental design.

**WP2.3.2 Data analysis**

High-throughput biomolecular profiling is increasingly used in diagnostics and therapy in clinical routine. The challenges of building an infrastructure to facilitate this feedback loop are enormous due to the complexity and volume of these data. Integration and comparison to relevant data in real time requires ICT approaches such as definition of standards, data compression, feature extraction, etc., which while not specific to biomedicine need solutions targeted to it. One emphasis will be on automation of data processing, detection of redundancies and integration of heterogeneous and highly complex data (WP3.2, WP6.7)

**Computational predictions of regulatory changes.** In collaboration with WP6, we shall implement biology-driven ICT to empower homology-based prediction of changes in gene regulation due to variants in the genome and epigenome (and, in some cases, transcriptome). This will in turn fortify the role of genome sequences in the building of mechanistic models of biological processes in different tissues. Similarly, the effects of epigenetic changes will be predicted if DNA but not RNA is available for analysis.

**Protein structure modelling - predictions of function changes and impact on protein-drug interactions.** As computational structure-function predictions improve, there will be better opportunities to foresee effects of functional changes in genes or transcripts computationally. Developing new ICT connecting this work to the network models of WPs6.1-6.5, will improve the prediction of protein-protein and protein-drug interactions, leading to strengthening of models and therapies. The very large numbers of variants and interactions will be tackled using distributed and/or cloud computing (with WP3.1).

**Digital image processing and analysis.** Imaging is at the core of medical diagnosis and treatment planning. Cell, tissue and body imaging on the macroscopic level X-ray, CT (Computed Tomography), MR and US images serve to diagnose medical problems and plan treatment. Through research in digital image processing and analysis, methods that can supplement human visual inspection with quantitative assessments have become available. Such quantitative assessments of images will become essential to study also molecules in time and space resolution and are part of the foundations for the modelling of the patient in ITFoM. This sub-workpackage will develop the ICT that integrates image processing with the structural and functional components identified in WP6.6.

**WP3 Infrastructure Hardware and Software**

The demands of ITFoM go beyond the capabilities expected from roadmaps for general purpose HPC technology over the next decade. WP3 will therefore address the development of innovative hardware and software developments as well as new computing paradigms to propose innovative, robust and efficient personalised medicine solutions throughout Europe and the world.

**Workpackage advisory group**

To ensure highest level expertise within the project, we propose to establish a high level workpackage advisory group. Members of this group will receive progress reports on the work within the workpackage minimally twice a year, will participate in regular workpackage level teleconferences, and will be invited to attend the ITFoM Meetings planned annually.
**WP3.1 Processing Efficient Primary Data Analysis CPU/GPU/Dataflow: Scalability**

Methods for the generation of genomic, transcriptomic, epigenomic, proteomic and metabolomic data have evolved dramatically in the past decade. The increase of data production, processing requirements and data storage of nucleic acid sequencing, which is nowadays the main tool of data production of genomics, transcriptomics and epigenomics, is dramatically outpacing Moore’s law. Simply extrapolating the computational effort inherent in genome/epigenome/transcriptome sequencing to hundreds of millions participants in the healthcare system will generate completely unacceptable costs and energy demands largely stemming from the computational and data storage requirements (e.g. primary data processing of a cancer genome requires 5000 CPUh and 200 Gb of storage). Similarly, new, more powerful, proteomics, metabolomics, imaging and sensor based analyses will create computational challenges of their own, requiring new hardware and software concepts. Together with WP2, and the other SWPs of this WP, WP3.1 will address the question of computational scalability of the analytical/therapeutic procedures which we expect to be routinely used in clinical practice in the future. Another huge challenge faced today and with more emphasis in the next years is how to integrate complex and heterogeneous data. At the same time there is high heterogeneity of architectures, usage/culture of programming according to various communities of biologists and physicians generating data (e.g. WPs1.6, 2). Multi-processor architectures have become a de facto standard in most business form factors as well as scientific workstations and HPC servers. Multiple configurations may be found such as multicores, CPUs, GPUs, many-core accelerators and for hybrid solutions. In order to exploit the unprecedented performance of this hardware, several programming models have been either adapted or created from scratch such as MPI, OpenMP, STREAM, CUDA, OpenCL and MaxCompiler. From the user’s point of view the main issue deals with the adaptation of legacy source code or the design of new software taking benefit of the intrinsic performances of these architectures. This workpackage aims to address the challenges of the primary data analysis/reduction across all analytical/diagnostic procedures in WP2 in the context of the complexity and the heterogeneity of architectures and programming usages. It will focus on the development of tailored computational systems that can process data accurately, efficiently and close to the data-generating device. The tailoring will in part be achieved by implementing the concept of ‘natural ICT’: human biology is organised in terms of functional modules in which a great many processes occur, but between which the number of interactions is limited (e.g. the millions of molecules inside a cell do not communicate with the millions of molecules in another cell, except through a very small number of communicating molecules). By distributing the computation over cores in ways that are similar to the distribution of the biological processes over biological components not only an increase in computational efficiency may be achieved, but also an increase in comprehensibility of the computations and programming. Here the hardware structure may be adjusted to the particular part of the human model is being calculated.

**WP3.2 Data to Objects – Objects to Models**

To derive integrated molecular/physiological/anatomical models of every individual, we will work with WP2 and WP5 to preprocess data according to their level of heterogeneity, complexity and redundancy to establish reference models of (wo)man which will then be personalised based on the data generated from the individual. This will require converting data (e.g. sequences), into the information on objects (specific proteins, gene) each with a rich set of attributes and methods, describing the state of the object, and its possible functions/interactions. Changing object features in the reference model based on individual data will automatically trigger changes in the attributes of downstream objects. Much of the individual information will however identify changes in sets of objects (e.g. metabolites), which are the consequence of unknown changes in upstream objects.
ITFoM will deal with model generation using thousands of heterogeneous types of data to describe the average health status of an individual within a reference population, i.e. converting observed measurement into relevant medical information. The mapping of measured data to model parameters is widely known in engineering as system identification. The most recent advances in system identification techniques will be taken advantage to individualise the reference model using specific data of the patient such as urine metabolites, blood samples, and radiological images, etc. The estimation problem resulting in model individualisation is highly nonlinear and of a gigantic order, but not intractable. The emerging particle filtering/smoothing-based nonlinear identification techniques are in fact parallelisable with near-linear speedup and can be implemented on massive clusters of multi-core computers. Thus, the execution time of the estimation algorithm can be drastically reduced. Furthermore, some intrinsic properties of the reference model are instrumental in dealing with the curse of complexity. First, the reference model describes biological processes evolving at different time scales, which fact enables the use of motion separation techniques and reduces the dimension of the original estimation problem through the divide-and-conquer principle. Second, the model parameters, though numerous, typically belong to a priori known sets. Thus the optimisation problem of individualising the reference model to a patient is very much constrained. Deducing changes in upstream objects from information on downstream changes due to a disease and given a mathematical model is a classical inverse problem, for which new approaches will be required. Both parametrical and structural changes in the reference model can generally correspond to pathology. A model-invalidation approach circumvents the ill-posedness of this specific inverse problem by checking a number of concurrent individualised disease-specific models against the patient data and excluding the infeasible ones from further consideration. This approach is computationally sound but demands mathematical models in health and disease. This workpackage will be developed and tested in close collaboration with WPs1.6, 2 and 6.

**WP3.3 Codesign of an energy efficient HPC**

This SWP targets an optimal processing of ITFOM applications by HPC technologies with a particular emphasis on energy efficiency. Bull will focus on a 10-years program dedicated to the transformation and mapping of project applications on most up-to-date HPC solutions developed by Bull. Obviously such works will be done in close cooperation with all ITFOM workpackages and partners. In extenso, Bull will provide evolutive and scalable HPC solutions providing a large range of computing technologies including general-purpose multi/many-cores CPU, GPU, DSP and FPGA to support in turn a large variety of applications. An optimal exploitation of such systems requires a rethinking of existing applications (again inspired by the Natural ICT concept; in collaboration with WP6.6) to introduce efficient parallelisation, optimal data management in conjunction with state-of-the-art programming models but also new programming languages and new approaches to optimisation algorithms/engines. Technologies will be developed in a series of ‘working prototypes’, implementing relevant parts of the data analyses and modelling pipelines from the current analyses of use cases in the project, providing both essential infrastructure for use cases/pilots, and a robust development route to the demands of the future.

**WP3.4 HPC SW Infrastructure and Technology**

This workpackage will define, develop and deploy the main HPC software infrastructure for the ITFoM project for the medium to long term; it is positioned between the physical hardware (exascale hardware) at the bottom and the applications at the top. To the applications it presents itself as a virtualisation layer: an API (Application Programming Interface) that offers the programming abstractions to the application programmer in WP6.1-6.5 to program the underlying system without having to bother with the underlying optimisations, dealing with hardware failures, runtime scheduling, etc. The runtime system implements this virtualisation layer. It is responsible for the
optimised scheduling of tasks, guarding the energy consumption, etc. It is built using sub-modules: Communication Library, Programming Model, Resilient Work-stealing Scheduler and a Virtual Machine. The hardware simulation is needed to emulate future hardware systems and support the software/hardware co-design tasks.

**Figure 1 HPC SW Infrastructure and Technology**

**WP3.5 Dataflow Computing**

Dataflow computing offers the potential to align structured and unstructured Big Data storage and computational challenges with optimisation of power consumption. Today, a dataflow computer is ~30x smaller, faster and/or lower power than comparable high end microprocessors and we expect this advantage to increase over the next decade. Dataflow compute engines maximise overall throughput by sustaining processing through a deep pipeline at “one result per clock cycle” for an entire calculation while maximising memory utilisation by eliminating non-computational tasks such as caching, instruction decoding or branch prediction. An analogy for moving from control flow cores to dataflow cores is the Ford car manufacturing model, where expensive highly-skilled craftsman are replaced by a factory line, moving cars through a sea of single-skill workers. By 2020, an exascale dataflow computing system would require just 10-50 racks, making it practical to deploy many such machines at hospitals or other major centres in the health system. We will focus on both processing of functional genomics data and the modelling applications which are well suited to the dataflow paradigm: some of the biological processes enable calculation of part of the network for an appreciable amount of time without data exchange with other cores. Dataflow computing will have a profound effect across the execution of ITFoM computational pipelines. An early point of integration will be the development of linked data-specific decompression routines for mRNA and Proteomics data (WP4.3.3) with analysis routines in the pipelines (WP4.2-5), in which we can expect to see a 100 to 1000 fold speed up due to the elimination of a write/read cycle of explicit decompression. We will work closely (but not exclusively) with Alacris and CNAG for models and algorithms, which take billions of core hours to compute, and collaborate with Imperial College London and other universities in Maxeler’s MAX-UP programme. In particular, we will collaborate closely with ATG and DCGMSB to address the scale-up problems mentioned in the beginning of the workpackage. The gain anticipated by using Biology-driven computing and Computing-driven Biology will be examined in collaboration with WP6.

**WP3.6 Interactive 3D Visualisation**

The multidimensional interactive display of the ‘virtual patient’ will be based on the integration of high detailed anatomical 3D visualisations with different data layers (imaging, geometrical, abstract, SWP6.6) and modelling results. The aim of WP3.6 is to define the major bottlenecks for the market delivery of these innovative tools and propose adaptive R&D programs to favour end-user readiness and significant improvements in the health value chain. They include: visualisation using interactive holographic displays to follow up the data storage, management, analysis and to compare the
robustness of the models generated with different initial conditions; real-time motion-sensing Cyber-glove controls for 3D interaction on the reference model and the patient data (imaging, tumours, drugs, protein folding), augmented reality, and the use of wireless mobile devices to visualise medical data. We will build on the state of the art in layered MRI body scans which integrates numerous 2D images into a 3D interpolation of the current body. Navigation through this 3D model is performed intuitively through gestures using glove-like motion sensing controls. Segments of the 3D interpolation can be isolated for greater clarity (e.g. to show only specific organs). Ultimately, this allows for a more holistic display of the current body state, which can provide all information gathered from multiple scans into one 3D model. Future expansion of this technology for the purpose of ITFoM would be to include the ability to justify aspects of the 3D interpolation by directly displaying the “classical” 2D MRI scans responsible for the 3D output. Further integrated in a CAVE environment this technology will allow health practitioners to navigate a 3D model of the patient, generated using MRI images, which were themselves generated from MRI data. This will provide integration of data at a step further than today’s technologies, combining multiple 2D images to produce a 3D model, in order to produce a more holistic view of the patient. The Milestones will be determined closely with WP1,2,4,5,6. In collaboration with WP1.3 and WP6, the methodology will be developed to project the results of dynamic model computations into the anatomical models that are visualised through CAVE. This should enable the individualised virtual physiological human to be visible to all users in all its biochemical and functional genomic aspects. It will also enable the display of the dynamics of therapy after application of a medicinal drug or a change in diet.

**WP3.7 Data Storage**

In consideration of the data lifetime and optimal access, this task will address the topic of data distribution and optimal placement, redundancy management (e.g. duplication, data mirroring) and archiving. The resulting infrastructure (WP4) should include all features of a federated database system (e.g. hierarchical navigation, data access and composition etc.) and take into account related standards (e.g. DICOM). The access to storage data will be strongly correlated to the communication network in place to provide accessibility to heterogeneous remote data (network accessibility protocols, e.g. WP4). Special emphasis will be put on the (semantic) integration of heterogeneous data sources and the management of data quality and data evolution. This package will exploit existing infrastructures such as EUDAT (European Data Infrastructure) in the short term. In consideration of the data lifetime and optimal access, this task will address the topics of data distribution, and optimal placement, redundancy management (e.g. duplication, data mirroring) and archiving using various technologies such as Cloud and distributed data warehousing. The data requirements of ITFoM require an exascale storage scheme with high bandwidth IO to the appropriate compute resource. To achieve this we will develop a specific Cloud based infrastructure spanning the different ITFoM resources, where there is both a number of “Data Hubs” which have definitive and archival data copies and a variety of “Edge Clouds”, often sited near specific compute or expertise area (see WP6). We will also exploit existing infrastructures, such as EUDAT, VISION Cloud, and ELIXIR in the short term and extend their functionalities in compliance with the data access policies described in WP1.6 and WP4.1.1 and the security controls devised in WP3.8 (e.g. federated access). The storage will also take into account related standards (e.g. DICOM, file servers) and semantic integration of heterogeneous data sources and the management of data quality and data evolution. In collaboration with SWP6.6, data will be organised such that they are accessible to topic oriented data libraries through pointers. We will work closely with SWP6.1-6.8.

**WP3.8 Security**

This sub-workpackage will design and develop the set of security services required in the ITFoM infrastructure to ensure that ITFoM patient data are adequately protected as outlined in WP4.1.1,
WP1.6 (Infineon), and WP6. We will build upon experience of federate access control (EUDAT, UCL, EBI, IBM Haifa), security protocols for safe transfer of large datasets over the internet, data anonymisation, and privacy-preservation during data mining. We will also address data privacy issues resulting from scientists and medical staff who would want to use their own devices (e.g. iPads, tablets) for both personal and professional access to patient data; the use of Cloud technology which forces us to move away from perimeter security of an organisation; and isolation of computing environment in multi-processing, virtualisation and cloud computing. In January 2012 the World Economic Forum made Cyber Attacks their 4th Top Global Risk. We will look into resilience techniques to ensure that ITFoM infrastructure is resilient to cyber-attacks since Internet will be its main conduit. Through an interaction with Quantum computing groups (Dwave, IQC, UCL) we plan to explore Quantum cryptography, both as an option for the protection of sensitive clinical data at the later stages of the project, but also to guard against the possibility that new developments in Quantum computing might make available ‘classical’ systems insecure in the future (see Hans W).

**WP3.9 Future and Emerging Technologies**

Several innovative technologies are actually in the R&D pipelines and may be the disruptive one that will help to solve the huge challenges related to the health data digitalisation and modelling in the health value chain for a daily medical use (preventive, therapeutic and patient follow up). Among those we may cite: new mobile phone related technologies and usages, near threshold voltage, 3D stacking, new memory architectures including 3D-memory stacking, silicon photonics, quantum computing, computing with photons, zero energy computing, magnetic computing and molecular computing as well as new network infrastructures such as those developed in the FET DynaNets project and innovative software engineering. While any of these may require many more decades to mature before they are directly applicable to ITFoM, we consider them as alternative computing technologies in this proposal, since for example if we were to make a major breakthrough in Quantum Computing quickly enough; the impact on the Future of Medicine would be immense. Contributions could for example be expected in solving inverse or large scale mapping problems with a specific focus on (i) Energy demand being the ultimate limitation of any computation intensive medicine; low energy computing might therefore become key to the quality of medicine and the level of healthcare costs in the future, (ii) Network Infrastructure to consider the futuristic aspects and light paths, lambda links and schedulable network links, (iii) Software Engineering: how to design a complex system such as the ITFoM. The milestones in this SWP will be developed together with WPs 1,2,4,5 and 6.

**WP4 Data Pipelines**

ITFoM has challenging data and compute volumes both in the near term and the long term. In addition there are considerable specific detailed workflows used in genomics which do not utilise the large scale compute infrastructure due to a lack of interfacing between domain specific expertise and computational expertise. In WP4 we will develop a scalable storage and compute system, building on and where necessary extending the established compute, middleware and storage infrastructures (Partnership for Advanced Computing in Europe, PRACE; EGI.eu, EUDAT, and a variety of FP7 projects listed below). We will also bring the domain expertise from high throughput genomics and proteomics groups to utilise this infrastructure, again building and where necessary adapting the existing systems (ELIXIR, ICGC, VPH, BLUEPRINT). We describe this logical data flow as the ITFoM Data Pipeline, and it sits on top of the hardware and system provision described in WP3. This infrastructure is fed by the data generation in WP2 and it supplies either processed information for data integration for WP5 and WP6.
ITFoM has assembled some of the best-in-class researchers and service providers in this area, spanning computational high end researchers (such as IBM HPC Disk Array, Bull and Oracle), computational work flow experts (such as CERN, the UCL, University Leuven) and domain specific data experts (such as EBI, CNAG, Barcelona, KTH, ETH and UU). The combination of this expertise will allow for both the immediate delivery of working pipelines for the early project stages, drawing on the experience of pipeline delivery in 1,000 Genome project, BLUEPRINT, ICGC, VPH and ENCODE (all of which have been delivered by the partners listed) but also able to scale to the potential of every individual in Europe having a personalised healthcare plan based on their individual molecular data.

**Workpackage advisory group**

To ensure highest level expertise within the project, a high level WP4 advisory group will be established for technical advice. Members of this group will receive progress reports on the work within the workpackage minimally twice a year, will participate in regular workpackage level teleconferences, and will be invited to attend the ITFoM Meetings planned annually. There they will also advise the Steering committee of the entire ITFoM.

The Major sub-workpackages are as follows:

**Workpackage WP4.1 Development and testing of a secure, scalable, state of the art data management scheme**

**WP4.1.1 Development of a robust, federated security framework and policies**

ITFoM will mainly be handling patient level data which will have a variety of restrictions on access, from specific clinicians through to a broader access to many researchers in the consortium. The definition of the access is outlined in WP1.4, and driven by the consent the patient has given for data use, and must also comply with legal restrictions. This consent must be transformed into a robust set of data access policies which can be used in a federated system. We will build upon experience of federate access control; in particular the research into Data Placement (e.g. a certain object must not be stored outside of Europe), Federated identity management (e.g. delegated responsibility to a national system of access control for certain broad data use; see WP3.8). Early milestones will be a precise definition of access policies in a pan-European level and the development of prototype integration of federated identity management with the ITFoM storage infrastructure.

**WP4.1.2 Raw Data Management Overview**

We will develop overarching meta-data control structures to allow a coordinated multi-technology raw data associated with multiple samples from a single individual to be tracked throughout the appropriate pipelines steps in WP4.3.1 – WP4.3.5, often in different physical locations as developed in WP4.1.2. The meta-data tracking system will utilise the “data to object” component of WP3.2. The coordination of meta data, both in terms of provenance (e.g. institutional) and scientific (e.g. sample) will enable the data integration in WP6.7 and further into the modelling workpackages of WP6.1-WP6.5.

**Workpackage WP4.2 Development and testing of virtualisation infrastructure**

**WP4.2.1 The development of generic data flow components and specific modules that fit within them for ITFoM**

The development of generic data flow components and specific modules that fit within them for ITFoM.

Both the near term and long term goals of ITFoM demand a robust, practical workflow of data items flowing through the system with multiple quality control checks. Current pipelines are mixtures of
formalised workflows in a variety of batch processing environments. For this to be shared first across ITFoM and secondly more broadly we will have to virtualise many of these components; often with current entire pipelines being a single component of a “meta-pipeline” (for example, RNAseq processing, which itself is a multi-stage pipeline, will be one node in a meta pipeline of a patient data processing). We will explore emerging workflow (e.g. MapReduce) and virtualisation (e.g. OVF) paradigms to achieve this scalability.

**WP4.2.2 Development of HPC middleware**

We will develop a federated method for the selection, dispatch and tracking of compute tasks, optimised for the federate storage layout WP4.1.2. This will build upon the established middleware layers developed in PRACE, EGI.eu and DEISA, which have been extensively used by the VPH projects (VPHOP, VPH2, ContraCancrum, VPH Arch, ACTION-Grid, IMPACT). As the project matures, the new compute schemes being developed in WP3 will need to be integrated into this scheme, and ITFoM will adapt and extend these systems to manage new platforms and compute schemas.

**WP4.2.3 Deployment virtualisation testing**

In the later stages of ITFoM, the methods developed will need to migrate away from the pan-European research setting of ITFoM to a Nation state healthcare setting. However, in doing so we cannot have either a total freezing of the methodology nor a long timelines (>5 years) for updates from the research team to the practitioner. Therefore not only must these methods and pipelines be portable, (enabled by the virtualisation WP4.2.1), but also there must be a way to be confident of the deployment and quality assurance of the process in a remote setting. Therefore ITFoM will work with some of the leading federated data management groups, such as EGI.eu, to create a deployment framework of large scale virtualised pipelines in which quality assurance is integrated into the framework

**Workpackage WP4.3 Development of Domain specific pipelines inside the virtual environment.**

These workpackages will feed into the different modelling approaches in WP6.1 to WP6.5

**WP4.3.1 Development of reference based nucleic acid specific pipelines applicable for Genomics (WP2.1.1), Epigenomics (WP2.1.2), Transcriptomics (WP2.1.3)**

As each of these techniques have common components, in particular an ubiquitous step for all of these is mapping of sequence reads to a reference genome or transcriptome sequence. The choice of the best mapping method may differ and the downstream data processing steps diverge in specific analysis components. We will draw from current state of the art pipelines, such as 1000 Genomes, ENCODE, GEUVADIS, ICGC projects and BLUEPRINT and the ELIXIR infrastructure. The common alignment component will use the dataflow computing component of WP3.5.

**WP4.3.2 Development of assembly/reference free based pipelines**

The metagenomic component of data generation (WP2.1.4) requires a different front end analysis component to the alignment method, either via kmer/read analysis or assembly. We will take state of the art methods developed in METAHIT and ELIXIR and implement on the virtualisation platform of WP4.1.1.

**WP4.3.3 Development of mass-spectrometry (MS) proteomics pipelines**

We will adapt the state of the art quantitative proteomic analysis methods developed in WP 2.2.1 for discovery and systematic screening to work on the virtualisation platforms of WP5.1.1 with testing from WP5.1.2. As MS proteomics technologies are still developing extremely fast, the data management and analysis pipelines will constantly have to be updated. They also produce vast
amounts of data (e.g., LC-MS experiments produce over 1,000 CID spectra per hour), therefore robust implementation of these pipelines is necessary, if they are to be used in hospital setting. The pipeline, for instance, will include live updates of peptide identification databases to maximise the proportion of identified peptides in the sample. In first instance we will particularly concentrate on methods which are most mature for hospital use, such as detecting patterns of differentially expressed proteins (e.g., between reference samples and disease samples).

**WP4.3.4 Development of affinity proteomics pipelines**

We will adapt the data flow methods currently in place in the Human Protein Atlas project and the ones developed in WP2.2.2, to provide tissue and sub-cellular localisation readouts from tissue arrays again on the virtualisation system of WP4.1.1 and WP4.1.2. This pipeline will include live access to antibody databases containing information about antibody specificity, will implement algorithms for recombinant affinity reagents, and sophisticated image analysis algorithms for histology slides.

**WP4.3.5 Development of Metabolomics based pipeline**

We will develop the matching methods of metabolomic measurements via NMR (solution or solid state) and other spectroscopy methods (e.g., CD) to provide both matching of molecules and abundance estimates. The pipelines currently used in research setting include steps for metabolite spectra pre-processing and feature extraction (such as data denoising using wavelet-based transformations, baseline correction, spectral peak identification, binning and normalisation to reference spectrum). These are usually followed by multivariate analysis to identify and quantify metabolites in the sample, or comparing metabolic profiles of several samples. Although due to the rapid development of the technology there are still no standard ways of analysing these data, the partners of this consortium are in a position to develop most up-date workflows and implementing them in robust pipelines deployable in hospitals.

**Workpackage WP4.4 Data specific compression**

**WP4.4.1 DNA**

Recent work on compressing DNA sequence has shown that a domain specific compression framework can produce 100 to 1000 fold better compression than straightforward application of generic compression. This is mainly achieved by transforming the data to make explicit the implicit high redundancy of biological data, with only a small amount of variation between individuals. We will extend this compression scheme to utilise the growing understanding of structural variation and complex index structure in the human genome, and to adapt to new technologies, such as very long read structure. We will carefully integrate this with WP5.5 (Dataflow computing) to ensure that there can be an efficient co-decompression and analysis on data flow style CPUs.

**WP4.4.2 Proteins**

Currently there is no data type specific raw proteomics spectra compression. Drawing on our experience of DNA sequence compression in which there is a data transformation to expose the implicit redundancy in a manner which standard compression schemes can be applied to we will develop specific data compression routines for Mass spectroscopy data. Similar to WP4.4.1, we will coordinate the decompression routine with WP5.5 for data flow computing.
WP5 Computational methodologies: statistics, machine-learning, modelling and data-integration algorithms

WP5 has the following main goals:

- Develop new machine learning ICT methodologies for ITFoM that greatly empower data integration through modelling
- Develop the new statistics that are needed in the arena of many times n=1 statistics for truly individualised medicine

Statistical Machine Learning (SML) is acknowledged as the most important driver of the technology underpinning the success of companies such as Yahoo, Google, Amazon, Microsoft, and Facebook. Furthermore, SML continues to play an essential role in the genomic revolution, being at the heart of the sequencing data analysis algorithms currently employed, as well as being at the core of analytical computational methods applied to all other –omics data (transcriptome, proteome, metabolome). Efficient and robust algorithms make many of these methods effective. We see the future of healthcare, as described in ITFoM, as no different: it will be fundamentally reliant on SML and on efficient and robust algorithms. Methodology, algorithms and software for SML tasks in the modelling challenges will be developed in WP5, while the actual modelling of biomedical data will be done in WP6.

Workpackage advisory group

To ensure highest level expertise within the project, we propose to establish a high level WP advisory group of leading international stature. This group will advise the WP leader on technical issues. It will be comprised of leading individuals in both Computational Statistics and Machine Learning. Members of this group will receive progress reports on the work within the workpackage minimally twice a year, will participate in regular workpackage level teleconferences, and will be invited to attend the ITFoM Meetings planned annually. There they will also advise the Steering committee of the entire ITFoM.

WP5.1 SML-based Diagnostics

Taking a SML approach to diagnostics can be a powerful complementary approach to mechanistic modelling. Such models often have strong predictive power that is directly relevant clinically, even if their explanatory and extrapolatory powers are in general lower. Furthermore, such models can be used as an initial analysis to identify variables and parameters that should be included in heavier first-principle computational efforts, and to identify variables that could be missing from mechanistic models or that need to be used to integrate existing mechanistic models into higher-level models. Conversely, knowledge from biological insights and mechanistic models can be used to boost the predictive power of SML models. There are several well-established test cases where SML-based techniques were successfully applied to refine conventional rough diagnostics (e.g., for breast cancer and for a large range of rare genetic disorders) with direct practical implications and products now available on the market. In general, from an SML perspective, diagnostics can be modelled statistically based on a number of predictor variables coming from diverse sources. The challenge is to develop models that are highly predictive and robust over multiple patient cohorts, while staying maximally interpretable. Given the enormous heterogeneity across patient (and non-diseased) individuals, this is a hugely challenging task and many previous studies based on genetic indicators, metabolites or proteins, as well as physiological indicators have met with limited success. These shortcomings will need to be convincingly addressed in a formal framework that allows bounds on errors to be characterised. Much of this work will rely on the ability to properly digest and analyse large-scale real world evidence data.
The extensive molecular analysis and measurement techniques of WP2 give a unique opportunity for developing the next generation of SML-based diagnostics. The additional challenge is how to combine the various data sources coming from the different analytical techniques and even from different organisms and cell lines. Modern SML-based data-driven integration methods combined with as much prior information as possible from mechanistic models will help reduce the massive diversity to a smaller set of latent variables usable as potential biomarkers. This will require both the development of innovative SML techniques for data integration / data fusion and for prior knowledge integration, and their adaptation and prototyping in real-world applications.

**WP5.2 SML-based Prognosis**

Similarly to WP5.1, taking a SML approach to prognosis holds great potential for personalised medicine. Specifically, predicting patient individual risk and/or drug-response can be modelled and approached by advanced SML-based techniques that examine data emerging from diverse sources. The ability to learn from clinical data and predict for disease complications has been established in various disease areas. An example is the ability to predict for development of diabetes or congestive heart failure in metabolic syndrome patients. Most of current studies either predict based on clinical status or based on association studies that use solely genes/SNPs. Proper fusion of these highly heterogeneous (clinical and ‘omics’) data sources represents a formidable challenge that will be gradually addressed within this WP. In addition, developing prediction techniques that are robust for various sub-populations is also a hugely challenging task and many previous studies based on genetic indicators, metabolites or proteins, as well as physiological indicators have met with limited success. These shortcomings will need to be convincingly addressed in a formal framework that allows bounds on errors to be characterised. Computational efficiency would become another limiting factor as data from the analytical labs (WP2) and from external sources accumulate and the project progresses, and effective approximations and heuristics will be developed.

**WP5.3 Statistical Inference Methods for Mechanistic Modelling**

The adoption of mechanistic models of health/disease subsystems (e.g. deregulation of signalling control) necessitates the formal identification of model parameters such as kinetic rate constants (which may or may not have biochemical meaning) and the associated uncertainty in these estimates. A full distribution over the model parameter values, conditional on available experimental evidence, is essential in being able to propagate model and data uncertainty when simulating models to make predictions. Rational evidence-based predictions based on a decision theoretic framework require the characterisation of natural posterior variability in model predictions. To reason about potential effects of drugs or phenotypic changes without carefully considering strength of evidence in data is dangerous in such a clinical setting. Therefore Statistical Inference methods for large-scale mechanistic models will need to be developed that can scale to the models being suggested in ITFoM as well as the numbers in a population (all of Europe) and the heterogeneity across individual models. This will require a step-change in the capabilities of SIM methods for mechanistic models. The major way that this will be achieved is by exploiting the recent advances in the theory of Markov chain Monte Carlo (MCMC) methods employing particle MCMC and the underlying Riemannian geometry of the statistical models. These models will be directly linked to the mechanistic descriptions developed in WP6 via the systems of nonlinear differential equations describing the biochemical pathway kinetics.

In the first 2.5 years this theory and methodology will be developed to accommodate the more complex systems of differential equations describing disease pathway dynamics. This will require intense and ground-breaking advances in the theory of statistical manifold and how they translate to efficient MCMC methods. The following years to Year 5 will exploit this theory and the statistical
efficiencies in developing advanced algorithms that will approach linear complexity and geometric convergence. Furthermore advances in processor and language design will be exploited in developing such MCMC schemes jointly with WP5.4 to support the inference needed for the statistical models of disease and health state employing both mechanistic and data driven representations. At the 10 year milestone we anticipate fully geometric convergence methods of linear scaling with inherently parallel algorithmic structures.

**WP5.4 Monte Carlo Methods**

Monte Carlo Simulation (MCS) of measure of interest is of importance in the inference methods described in WP5.1-3. Whilst naive MCS can be deployed for simple chemical kinetic models, the complexity of models being considered in ITFoM where slow and fast reactions will be nested in complex non-equilibrium states will ensure that such methods will grind to a halt. Fast exact and approximate methods of simulation will have to be developed to support further the model-based methods of drug-target effect exploration. For example, the Chemical Master equation describing the stochastic kinetics of a cellular system is largely intractable and as such diffusion approximations can be employed. However, the Linear Noise Approximation can be developed in this regard as can variational approximations, as defined by collaborators at XEROX research Centre Europe. In addition, process-calculi based methods developed by Microsoft COSBI will be enhanced and adapted to the ITFoM pipeline. We anticipate that at 2.5 years linear noise approximation Monte Carlo schemes will be available; variational, deterministic and stochastic approximations, detailed in year 5, with the capability of arbitrary Monte Carlo simulation of large scale systems exhibiting slow and fast dynamics being available at year 10.

**WP5.5 Probabilistic Model Integration and Interoperability**

The diversity of approaches to modelling in ITFoM can be brought into a coherent inferential framework where models can be combined in a weighted manner based on their individual probabilities of modelling data (or representing available evidence) - so mechanistic models (WP7.1), statistical models (WP5.1), abstract computational models (WP5.2), or detailed metabolic style models (WP7.2) - can all be combined to make predictions in such a manner that the levels of uncertainty in model specific predictions - and tendency to 'overfit' - can be guarded against. The combined models represent novel insights regarding different diseases that will be translated to knowledge that can serve the physicians, e.g. under the form of treatment recommendations.

This WP will further enrich “–omics” data with phenotypic and other contextual information about the patients. Large-scale latent variable and probabilistic graphical models would typically be suitable candidates, but they are currently unable to scale to billions of observations, as targeted by ITFoM, especially if many multiple data sources need to be integrated. However, such approaches would enable medical doctors to obtain a ranked list of expected results for each type of potential treatments, thus increasing its downstream applicability. In addition to further scaling up current large-scale probabilistic models, this WP will also develop the necessary data aggregation and text summarisation techniques as well as semantic interoperability between models to facilitate information access and avoid information overload by building on large-scale text mining and recent advances in statistical natural language processing.

**WP5.6 Large Scale Data Mining**

The ITFoM vision is tightly related to the explosion of available clinical data and '–omics' data. These data types arise in different contexts and forms, and range from highly structured data to completely unstructured data, for example, in the form of free-text data provided by clinicians within patient discharge letters. Large scale data mining techniques can extract invaluable insights out of such real
world evidence data, highlighting trends in treatment allocation, revealing healthcare organisation best practices, and more.

In recent years different platforms have been developed for efficiently handling analysis of extensive large data. This includes Hadoop, an Apache project that provides an open-source implementation of frameworks for reliable, scalable, distributed computing and data storage. Various SML algorithms are available nowadays for applying large scale analysis. This WP will develop the necessary mining techniques that will be exploited by other WPs as appropriate and will leverage techniques developed in other WPs to make them applicable in a setting beyond first demonstration in order to be applied to abundant data.

Within the ITFoM framework, we will develop methods that seek for patterns in genomic, proteomic, transcriptomic, metabolomic and epigenomic data generated by WP2 (analytics) and construct abstract (non-mechanistic) models from large scale patient data. Correlations of these patterns with patient phenotypes will be sought. In the first two years two or three different disease settings will be studied in accordance with disease data available within ITFoM or publicly. Later, image and sensor data and tissue-specific patterns of the patients will be mined as well. By year 5 a system that works on Terabyte of data will be available demonstrating the value of the learning for case were data from many individuals is exploited, and data from each individual contains large size of –omics data. Next, an effort will be devoted to make the system usable at point of care in accordance to the infrastructure design of WP3.

Another direction of this SWP will be development and incorporation of advanced methods for feature/pattern selection. In order to handle large data volumes, brute force methods do not suffice and efficient techniques that find the most informative features/patterns must be utilised.

**WP6 Integration platform**

WP6 has six main goals:

- New-ICT mediated integration of maps, data, models and tools
- ICT-modelling based integration of all everything required for the virtual twin and individualised medicine
- ICT-catapulted modelling of health and multifactorial disease
- ICT-mediated integration of all ITFoM activities
- Through the above integral virtual twin models from molecule to pathophysiology and back
- Through the above the putting in place of individualised medicine through revolutionary ICT

Using biology-driven ICT strategies, WP6 will integrate biomedical and clinical data and queries from WP1 with new analytical data from WP2 and with stored data from WP4 to generate individualised and clinically relevant models of the human body in health and disease. WP3 supports the development and execution of the advanced hardware and software ICT required. WP5 provides extensive support in computational techniques such as machine learning and statistical analyses. In WPs 6.6-6.9, WP6 will integrate maps, data, models and tools at a ‘market’ with four ICT integration booths. dynamic integration of information into ICT models of individual humans at this market will use the four modelling strategies as the integrators (in WPs6.1-6.4). Interaction between the modelling strategies will produce mixed strategies (WP6.5) that combine and enhance the strengths of each strategy. In this manner, WP6 generates ICT models that address a variety of biomedical problems, *i.e.* the use cases defined in WP1.
Workpackage advisory group

To ensure highest level expertise within the project, a high level WP6 advisory group will be established for technical advice, comprised of leading individuals in both ICT integration and ICT modelling such as Denis Noble and Hiroaki Kitano. Members of this group will receive progress reports on the work within the workpackage minimally twice a year, will participate in regular workpackage level teleconferences, and will be invited to attend the ITFoM Meetings planned annually. There they will also advise the Steering committee of the entire ITFoM.

WP6.1 The Watchmaker’s Strategy: mechanistic bottom-up modelling

The watchmaker strategy will develop extensive bottom-up approaches for modelling interacting molecules, interacting pathways, interacting cells and interacting tissues. Human molecular biochemistry, as determined by all its molecular processes, drives this modelling strategy. We will deal with computational implications of tissue-specific, patient-specific genome-expression patterns for fluxes and concentrations at macromolecular, intracellular, paracellular and whole body levels, and gene expression. This will range from the dynamic modelling of the interactions between proteins/nucleic acids and their low molecular weight ligands [molecular dynamics] and will model upwards to interacting organs (mechanistic physiology; collaboration with WP6.3). We will also deal with model individualisation (incorporating single-nucleotide polymorphisms, specific drugs and nutrition). We define the relevant human network (e.g. the metabolic network in hepatocyte) in terms of its components, and their interactions, as mapped in WP6.6. Communication parameters between biological macromolecules can be inferred from analytical data, and the strategy further characterises the chemical or transport reactions carried out by these individual macromolecules experimentally (involving WP2 and WP4). Analytical data describes the molecular interactions in the form of equations, obtained from WP6.9. Equations will include, among others, enzyme kinetics, equilibrium binding, Petri equations, Boolean expressions, Master and Monte Carlo equations. Equations are precise in terms of rates, mechanism, and parameter values, enabling maximum representation of the effects of drug molecules (e.g. competitive inhibitors, DNA intercalators), SNPs and nutritional biochemistry that impinge directly on the macromolecular activity. By feeding the information for each molecular communication into a computer model of balance and rate equations, resulting differential equations are integrated in time or solved for steady state (using software and hardware solutions optimised by WP3). We will deal with spatial aspects by compartmentation of space into the actual biological compartments (organelles, cells, tissues), or where necessary by using partial differential equations. This strategy deals with low molecule numbers by using mechanistic stochastic differential equations or Monte Carlo methods, with WP6.8.
and WP5 supports. As the interactions between the molecules, enabled by their biological organisation, bring about the functional properties that determine cell behaviour and organism function in health and disease, functional genomics helps make this approach comprehensive. We will also engage pharmacokinetics / pharmacodynamics. New model checking methodologies will be used to examine model consistency, and model validation will then be done on the basis of databased (WP4), new experimental (WP2) information and consistency with the expertise of medical and patient expert panels (WP1.1).

**WP6.2 The Engineer’s Strategy: interacting molecular Monte Carlo modelling**

Many diseases have predominantly molecular causes and will have to be treated by molecular entities (drugs), which affect molecular networks. Molecular processes taking place in different tissues and cell types do however interact. This sub-workpackage will explore and implement a strategy (Interacting Molecular Monte Carlo Models, IMMCM) currently implemented in PyBios, an object oriented modelling environment, in which the individual is modelled as interacting molecular models representing different cells or tissues. Individual models, in turn, consist of interacting compartments with the different objects, which closely represent the biological objects, the complex biological networks acting in the cell (e.g. one object could represent the ras protein, with one set of functions, while the object representing the mutant Ras protein has correspondingly changed functions, see WP3.2). This object layer can automatically map into one or more computation layers, generating e.g. systems of differential equations or Petri nets, which are then solved numerically. To make this possible, in spite of the lack of knowledge on e.g. the values of rate constants inside the cell, we will use a Monte Carlo strategy, with parameters drawn randomly from probability distributions reflecting any knowledge we have. Multiple parameter vectors are then used to model all the different states to be compared (different cell types of the patient with/without different drugs, with/without different growth factors etc.), until statistically valid predictions can be made (with WP5).

To ensure computational efficiency while preserving the detail relevant for different disease areas, abstraction mechanisms will be developed to generate sub-models optimised to address specific clinical problems automatically, using different levels of abstraction in different parts of the network, and in the representation of different tissues of the body of the patient. Sub-models can then be validated by comparing the results of a small number of modelling runs of the full model with those of relevant sub-models.

To generate reference models, sets of tissue samples provided by WP1 will be analysed in WP2, to provide a first set of molecular models of different tissues, which will then be improved and validated on data on individuals analysed as part of PGP (www.personalgenomes.org, HMS). To generate ‘virtual patient/virtual individual’ models for patients from internal use cases, or outside data sets, we will integrate all available data (WP2) and converted into object descriptions (WP3.2, WP6.7), by combining downstream propagation of effects (e.g. mutations in a gene will propagate to the corresponding transcript and protein) and reverse engineering (e.g. individualisation of a liver model from urine metabolome measurements, WP3.2), which can then be used to predict effects and side effects of different therapies on the individual patient, and can then be validated in patient derived cell lines, xenografts, and ultimately the patient.

During the ramp-up phase, we will, in particular, focus on the analysis and modelling of tumour patients, and rare and metabolic diseases, based both on internal and externally provided data sets (e.g. ICGC data). In addition, we also plan to generate ‘virtual patient’ models from externally provided data sets from the high priority use cases (e.g. ICGC data, www.icgc.org), as well as on well characterised cohorts in disease or phenotype areas, for which high quality genome/exome data are available, but for which causative mechanisms allowing the establishment of mechanistic models are
not yet known (e.g. neurodegeneration, sepsis, healthy ageing, immune diseases etc.). In particularly, the rapidly increasing volumes of external data sets available through many sources will pose rapidly increasing computational challenges, inherent in analysing increasingly complex models on a rapidly increasing number of individuals, addressed in close interaction with all SWPs of WP3, with the ultimate goal to contribute to an integrated infrastructure scalable to cover all individuals in the healthcare systems in Europe.

**WP6.3 The Mechanic’s Strategy: physical and physiological modelling**

This strategy encompasses three key concepts: (i) there is no exclusive space-time scale at the basis of all pathophysiology; (ii) mechanistic knowledge is preferable to phenomenology; and (iii) the best way to capture knowledge on observable processes at different scales and to compose them into a representation of the systemic interactions is through predictive models. These *physiome* concepts have already found a concrete technological incarnation in the Virtual Physiological Human (VPH), a framework of methods and technologies mostly starting from the organ-tissue levels and accounting for physical processes involving the transmission of forces, the deformation of tissues, the propagation of electrical charges, the movement of the solid and fluid tissues, and the generation and propagation of thermal energy. In this sub-workpackage, VPH will expand its work at the organ and whole human levels, yet reinforce its interest in working ‘downward’ and integrate knowledge from the molecular level (involving WP6.6 and 6.7). We will develop predictive models that encapsulate the mechanistic knowledge on each observable process, at various scales from molecule (involving WP6.1) to tissue, organ, organism and environment. This sub-workpackage will capture quantitative information about patient physiology (medical imaging; SWP6.6), laboratory exams (WP1), environmental determinants related to physical activity, nutrition, pollution, etc.) and make this information inform personalised integrative ICT models of the patient, which are then used for diagnosis, prognosis, and treatment planning (in collaboration with WP1). We will enable reutilisation of component models into integrative models different from those for which the component was designed originally (with WP6.8). We will develop computational infrastructures (involving WP3) that enable efficient execution of these highly complex computational models for each patient in a time frame that is compatible with clinical organisation. We will further improve modelling technologies for stochastic representation to model tissue-cells interactions. With WP6.5, we will also instigate integrative models that describe how parts interact to produce pathophysiological mechanisms and are initialised both by genomics and with physiological data (such as from medical imaging, lab exams, clinical assessments, lifestyle, nutrition, thereby enabling individualised medicine; WP6.7 and WP4). This involves establishing workflows that accept digital images as input, and effect a model-assisted interpretation and prediction that is again projected into an image (involving WP1 and WP3).

**WP6.4 The Learner’s Strategy: modelling through computational statistics and machine learning**

The success of machine learning approaches in many disciplines, including for instance predicting protein structure and the weather, inspires ITFoM to develop this modelling strategy for systems medicine targeting clinical applications. The challenge of high-throughput measurement of kinetic constants of all molecular processes may be avoided by implementing an alternative strategy of machine learning from massive amounts of other (dynamic) data. This brings new challenges such as statistical model inference from data, statistical approaches to deal with underdetermined models containing hundreds to thousands of variables at multiple scales, and inference of intermediate values from end states. This strategy uses sets of phenomenological equations to characterise processes that themselves may be conglomerates of molecular or higher level processes. Kinetic
parameters of components determined experimentally in vitro are not required in this method. Rather, the model predictions are compared with experimental findings at the functional level and parameters and equations are fitted automatically so as to produce the best correspondence between model predictions and experimental result; models learn to behave properly by comparing themselves with reality. In an extreme form the model is a neural network such that there is not even a possible relationship between a parameter value in the learned model and a parameter with a meaning in physical reality. We will implement machine learning and statistical strategies for molecule-based models of human physiology and pathology. On the one hand, this involves the development and implementation of supervised learning workflows with prediction, classification and regression (with WP5). On the other hand, we will work towards unsupervised learning workflow involving for instance cluster analysis (with WP5). We will develop and implement probabilistic graphical models and Bayesian networks, and seek consistent methodologies for feature selection and extraction (WP6.8 and 6.9).

**WP6.5 Hybrid and new strategies**

This sub-workpackage will scout for strategies that are not yet in the ITFoM envelope and aims to associate the more promising ones to ITFoM. These include modelling strategies contributed by industry. It will also identify the limitations of the four core strategies of ITFoM (WP6.1 to 6.4) and develop improved mixed strategies in order to allow the simulation of systems by exploiting efficiency and problem size advantages of macroscopic simulations while accounting for microscopic detail only in critical subsections of the problem. Part of the mixed strategies will start in the Virtual Liver Network program in Germany, where mixed approaches are already performed to understand liver function in relation with liver molecules. This SWP will also ensure that all the valuable models remain available for the community, by involvement of SWP6.8.

In this SWP we shall also build modelling bridges to the medical user community and patients, creating mixed models that are optimally understandable to these groups and that encourage the communities to give input and validation material for the models in WP6. Here integration of dynamic network models with visualisation models will be brought about. Similarly there will be integrations with small medical gadgets for individuals keeping track of their health real time.

**WP6.6 Network maps: network integration**

Available genome-wide reference networks are being developed into a consensus representation. Tissue-specific representations of human metabolism expression data will be augmented and expanded to include signalling and transcription factor networks, providing a reference resource from which individual, condition-specific molecular models can be generated (as in WP6.1 and 6.2). Jamboree and crowdsourcing through open interfaces mechanisms will be entertained. Existing data resources, with coverage of metabolism, signalling and transcription factor networks, will be exploited here. A corresponding activity will be set up for consensus anatomical and functional maps. Maps will describe the human body in terms of tissues, and will describe how a tissue is spatially organised in terms of various cell types, etc., all the way down to the spatial distribution of intracellular components (compartments, enzymes and molecules). This SWP will also develop the ontologies for these mappings. This sub-workpackage will also (i) identify whether more map categories are needed for ITFoM; (ii) identify existing mapping activities for each category and ensure that these are brought to consensus and compatibility; (iii) ensure that the maps have optimal formats for use in ITFoM; (iv) instigate a web-mediated community process for the making of maps; (v) develop an editing and quality-control mechanism; (vi) implement versioning control of the continually evolving network maps; and (vii) integrate the various types of maps, such as in
combining anatomical with biochemical maps, showing that certain pathways are only active in certain organs and cells.

This SWP will also define ‘virtual ontologies’ that aid the conceptualisation & visualisation required to revolutionise the reciprocal interaction with stakeholders including patients, physicians, ICT experts and biomedical researchers; each will be approached ‘in their own language’ which will be developed here in close collaboration with WP1.1, 2, 3, 4, 5 and 6.7.

**WP6.7 Data integration**

WP6.7 and 6.8 together will ensure the effective integration of (i) the reference type-1 (see WP1.4) data sets of WP1.2 with (ii) type-2 data collected in WP2 and retrieved from WP4 with (iii) the models being developed in WP6.1-6.5 into ‘data models’ (type-3 data). These data models are also integrated further with epidemiological and economical data into type-4 data models, useful for HTA and policy makers (see WP1.6). WP6.7’s data integration into these data models will be biology-driven, i.e. integrate individual, tissue, cell type, gene, mRNA, protein, metabolites that belong to each other functionally, rather than for instance all mRNAs across the human metagenome (which will be done in WP4): WP6.7 will develop a new a comprehensive data integration system for the heterogeneous data types being generated, where the data are always also associated with models of WP6.1-6.5). Together with WP6.8 and 6.9 and 3 and 5, common-denominator data-models will be made web accessible such that the ICT/biomedical community can see their data ‘come to life’ in the sense of a dynamic display of the implications of the data for human function. Mechanisms will be developed for facilitating quality control, data tracking, ownership attribution and specific service layers defined for the access of data within the collections that this sub-workpackage will index. The data itself will be annotated “anatomically” using appropriate ontologies where available. A key part of this infrastructure will be the automatic integration of data with models to facilitate both model validation and refinement (with WP6.8). We will oversee the adoption of data integration standards within the consortium, pre-existing where fit for purpose or instigating the development of new standards where none exist. WP6.7 will also ensure integration with existing and future ESFRIs such as EuroBioimaging, BBMRI (biobanking), ELIXIR (Bioinformatics), and ISBE (Systems Biology), as well as with other European and large national data programs (FP7, SysMO-db, VLN). This sub-workpackage will also define the processes by which data is passed between WPs 1, 2, 4 and 6, and the development of methodologies for the integration of biochemical network data and imaging data.

**WP6.8 Model integration and management**

This sub-workpackage will deal with the challenge that molecule-based medicine involves much of the entire human body, and can be modelled at different scales. Molecules that operate at timescales of milliseconds and spatial scales of nanometres need to bear on disease processes that evolve over days and meters. Models for different time and length scales must be connected to each other. This sub-workpackage will address methodologies for model reduction of the fine-grained models without losing the connection to the molecular world. In addition to different scales, ITFoM must integrate models of different types (mathematical ODE/Bayesian, models of information flow, etc.). These integrated models must be constructed seamlessly, and automatically where possible, with the user keeping control of quality. This will be implemented using expertise and methodologies developed in WP5. Relationships between models of different scales will be described in Multiscale Modelling Language (MML), being developed in the FP7 MAPPER infrastructure project. This will facilitate the deployment of models onto distributed computational resources of suitable scale for the particular model (from high-powered workstations to the largest supercomputers available internationally), in collaboration with WP3 and WP4. Additionally this sub-workpackage will deliver
model repositories for ITFoM, for which we will incorporate versioning and provenance methods and tools to facilitate modularisation of models. Together with WPs6.8, 6.9, 3, 5, common-denominator data-models will be made web accessible, for use in assessment and validation by the entire ICT and biomedical community. Together with WP2.2, 3.6 and WP3.8, this SWP will also create a facility whereby a human individual can download the common denominator data model to her/his home, disconnect from the web, insert her/his individual data, calculate the implications of the data for her/his health and disease, and perhaps export these implications (and not the personal data) to her/his physician. This sub-workpackage will also define the process by which models are passed between WP1 for medical use, WP2 for experimental design and positioning of results (involving WP3 for test bedding biology-driven ICT hardware and software solutions) and back to WP6 for iterative model improvement.

**WP6.9 Tool integration**

In this sub-workpackage, modelling and data management tools (libraries and application software) will be made accessible to all partners (and after further proofing of quality to the community). This sub-workpackage will be responsible for coordinating release cycles of tools ensuring that the ITFoM toolset remains interoperable and backwardly compatible. Many tools will be newly developed or optimised within WPs 3-5 and 6.6-6.8.

**WP6.10 Integrative modelling of use cases**

WP6.10 is an extensive workpackage in which a large number of use cases of ITFoM are modelled, each by a number of the ICT modelling strategies of WP6.1-6.5. The use cases worked on by ITFoM refer to concrete inroads into truly personalised care of health and disease. The use cases are defined and listed in WP1.7 in collaboration with the medical community.

WP6.10 will have one sub-subWP for each use case, in which a group of modellers of WPs6.1-6.5 and use-case specialists (from WP1 and its connections in the medical and patient field) together integrate the biomedical data to yield corresponding virtual patient models. In each case both (i) a common denominator type-3 data-model will be made, on the basis of type-1 (medical patient record) and type-2 (molecular and physiological assays) data and the modelling methodologies developed in WP6.8, and (ii) individualisations of this common denominator model (by overlaying the individualised type-1 and type-2 data). Type-3 and type-4 data models (which include epidemiological and policy aspects) will be produced and referred back to WP1.7 and WP1.1 for communication with the medical community, back to WP2 for experimental design; to WP3 to drive hardware and software development; to WP4 to address data storage requirements; to WP5 to request improved ICT methodologies, algorithm development and computational statistics; and to WP7 for training purposes and communication with the public and policy makers.

Most use cases will also integrate other use cases and other areas of systems medicine. For instance, the overlap of signalling networks between cancer and metabolic syndrome, as also suggested by cross-effects of drugs, will be expanded upon. And the role of the commensalisms between the human host and its microbial guests will be addressed, both in the context of pathologies of infectious disease as well as in the context of the effects of the microbial flora of the intestine, skin and vagina on human well-being. This is of great interest to pharmaceutical, food and cosmetic industries, particularly in ICT-facilitated handling of large data flows from metagenomes and functional metagenomics. This sub-workpackage will also address the differential effects of drugs on diseased and healthy tissues (differential-network-based drug design), the dynamics of drug uptake by the body, and the effects of drug molecules on their network targets (physiology-based PK/PD).
We will address another cross cutting use case, the health and aging in the sense of the gliding platform from what is perceived as health to disease. Osteoporosis, and obesity are suspected examples that are not included in the use cases already specified.

In the ITFoM vision every patient and disease is different from any other. This implies that all diseases are collections of rare diseases. In this sense all uses cases will fully take into account the rare diseases in their area.

**WP6.11 WP6 coordination**

The complexity of WP6 requires close monitoring and coordination, which will be organised through a separate sub-workpackage. This sub-workpackage will be coordinated by WP6 coordination and managed by a full-time WP6 manager in close coordination with WP7. WP6 will associate all research programs that are willing to contribute to any of the above integration sub-workpackages, with the proviso that they subscribe to synergise as indicated above. This means that research programs focusing on data integration are welcome provided that they are willing to go into the specific requirements of the WP6 modelling strategies and use cases. Similarly, research programs that focus on use cases are welcome provided that they participate intensively in the sub-workpackages that address the ICT integration challenges and in all four modelling strategy sub-workpackages. Research groups should associate with research programmes.

**WP7 Coordination and Management**

This workpackage will be responsible for management and coordination of the overall project, and carry out essential tasks. It comprises the administrative/financial as well as the scientific coordination. The work in this workpackage will be both strategic in providing the overall vision and guidance to the project, and also organisational with a strong focus on communication. This workpackage will monitor and ensure that the common strategy will be followed overall in the project. Together with the Steering Committee the role of each partner will be defined and key performance indicators will be monitored and if necessary redefined. This workpackage will also be responsible for positioning in and integration into the international context of other on-going research directions or programmes. Relationships and synergies with potential and possibly relevant complementary initiatives in other blocks, in particular outside the EU, will be established and organised. In addition, it includes discussion of the technology areas, formulation and if necessary revising of research and development plans addressing all topics. This WP will also be instrumental in setting up and supporting the Governance Structure. Dissemination (incl. IP strategies) and Communication internally as well as Public Relations will be implemented and coordinated through this WP.

The following sub-workpackages will be addressed:

**WG 7.1 Project Management**

WG 7.1 coordinates and supervises all processes of the project. It is responsible for all reporting to the EU, and coordinates and supervises all contract negotiations. A stringent coordination of the two project management streams (WP 7.1.1 and WP 7.1.2) is implemented and guaranteed through the direct leadership of the Steering Committee’s Executive Subcommittee.

**WP 7.1.1 Scientific project management**

WP7.1.1 is responsible for the smooth interaction of all scientific and technological parts of the project, the planning and monitoring, and the assurance of the workflow. In addition, it is responsible
for establishing and maintaining efficient communication between the individual bodies of the project and supervises the progress in the workpackages according to their work plan and the application of agreed upon metrics.

**WP 7.1.2 Financial/administrative project management**

A project of this size and magnitude needs the implementation of a specific financial and administrative project management allowing a smooth controlling of expenses and audits. WP 7.1.2 will collect the information from the project partners and prepare the reports for the European Commission as requested, including an ethics section, and pertinent documents (approval by the local Ethics Committees involved in the project and Consent Forms to be used for the collection of samples from the donors and subsequent uses).

**WG 7.2 Mission and overall strategy**

It is of highest importance that the project follows strictly its vision and mission. A 10 year time frame on the other hand needs room for flexibility to react to new developments as well as to adjust to programmes and initiatives outside of ITFoM and to the occurrence of changing societal and medical needs. To be able to fit these two tasks a Steering Committee will be established for internal coordination and a Programme Integration Committee for external alignment.

**Tasks**

**Organisation of the Steering Committee (SC) and Programme Integration Committee (PIC) meetings.** WP7.2 will be responsible for the organisation of the meetings. The Steering Committee meetings may also be held as conference calls at least bi-weekly. Permanent members of the SC are the 7 WP leader and the chair of the Executive Subcommittee (ESC). For specific expertise additional people will be co-opted, e.g. to represent Industry, Patient organisations, special scientific disciplines not fully covered by the WP leaders like Epidemiology, Public Health/Regulatory Policy, Bioinformatics, or others.

**Coordination with other European and global initiatives.** The Project Coordinator also chairs the PIC of external projects. This committee is established to develop synergies with other related on-going research activities in the field of systems biology and computational modelling to avoid duplication of efforts. Furthermore, the concept of ITFoM will be harmonised with development needs of the Innovative Medicines Initiative. To monitor the progress made in external projects and to fine tune the activities with the various projects and programmes, scientific and/or funding representatives of the external projects and programmes will participate in the PIC.

**WG 7.3 Governance, Funding and Financing (GA4H, DCGMSB)**

WG 7.3 will set up and maintain a sound governance structure for the project taking into account the needs of funders, participants and legal instruments for implementation. During the CP-CSA phase the typical governing instruments like Governance Board, Steering Committee, Executive Subcommittee, Scientific & Ethical Advisory Board Stakeholder Forum will be established. During Horizon 2020 the current deliberations are to establish an independent legal entity for the execution of the programme, e.g. an EEIG. The Governance Structure of such an instrument differs from a project-oriented one but will keep the main decision and advisory bodies as during FP7 (Details under chapter 3. Governing Structure). The responsibility for the Stakeholder Forum lies within WP1. This WP will develop a concept for funding and financing of the full operation phase during Horizon 2020, considering the full spectrum of national, European and private funding schemes as well as financing through industrial co-operations. The concept will be elaborated in a cooperative effort
actively involving Ministries as well as national and other funding organisations from several European Member States and outside the EU.

Tasks

Establishment of the Executive Subcommittee (ESC). It will consist of the 3 founding member Hans Lehrach (chair of the ESC), Hans Westerhoff and Kurt Zatloukal, the Deputy-Coordinator of WP7 and the Programme Manager (PM). An executive management office will be established.

Organisation of meetings. PM will be responsible for the organisation of several meetings including annual Governance Council meetings (General Assembly), at least yearly Advisory Board meetings. For these meetings, the programme and agenda, information packages, working documents and meeting reports will be prepared by the PM.

Development of a concept for sustained funding and financing. The concept will follow different strategies for short-, mid- and long-term funding.

Contracting. The PM in Cooperation with the ESC will be in charge of drafting and negotiating the contracts with ministries, funding agencies, other funding sources, and industry. All funding and financing-related issues will be performed in close interaction with WP7.2. The drafting of contracts will be outsourced to lawyers.

WP 7.4 Dissemination and Media

The ITFoM project, if it is to reach its goal of transforming medicine from stochastic/empirical to inference and knowledge driven will have generated a truly extraordinary amount of material for commercial exploitation. The industry partners in this proposal will obviously generate new data storage, encryption, compute architecture, visualisation methods, as well as new technologies, and systems which will each have a new ICT driven biomedical market. The Analytical industry will develop new devices and biotech and Pharma will use the computer modelling as a new mechanism for drug development. Dissemination of results is therefore a critical objective. In addition to scientific publications and publications of the project partners and the international conference, WP7.4 will generate flyers and brochures for experts and lay public to inform about the importance of ITFoM. Furthermore, a website will be set up for internal as well as external communication. The content of the documents will be prepared in close collaboration with the SC. The design of the documents and hosting of the website will be outsourced to professional providers. Also new media (incl. social media) will be employed for the communication to reach a broader public, but they will also be helpful for internal communication within the consortium partners. A whole spectrum of instruments will be used to involve regional, national and international media (e.g. newspapers) to support the dissemination strategy beyond scientific communities The education challenges for physicians and other sector representatives will be covered in WP 7.6. WP 7.4 will set up a small policy advisory group (3-4 people) which will help to develop policy action in the implementation of R&D activities into other policy areas like health policy, regulatory and alike.

WP 7.5 IP

Establishment of a Working Group for technology transfer and IP issues. The working group will collect information from all participating institutions and provide a compilation of support and transfer measures and activities. The working group members will also evaluate the existing measures and identify gaps that need to be addressed (e.g. opportunities for public-private partnerships). The committee provides feedback on potential IP and possible routes of exploitation to the GC and SC. In case of foreground knowledge generated by more than one partner the committee can give advice and support if needed.
**WP 7.6 Education and training**

Within the research programme of ITFoM the consortium will also implement an ITFoM Academy with a specific education and training programme (E&TP) dedicated to different target groups. The implementation of the E&TP will be stepwise and according to the development of the ITFoM research programme. Besides the organisation of courses and training programmes the Academy will also set up and build a knowledge repository where course material, information, eLearning and hands-on material is collected and made available for the whole community. There will be 4 pillars of E&T: (1) Internal Training, (2) Masters programme, (3) PhD programme and (4) activities for user communities. ITFoM will develop a modular programme in line with the IMI (Innovative Medicines Initiative) EMTRAIN (European Medicines Research Training Network) and uses existing initiatives and infrastructures connected to ITFoM to optimise the efficiency of the outreach and the spectrum of topics.

**Tasks**

**Establishment of an E&T WP.** This WP will consist of one representative from each participant. This WP will form the core coordination platform and ITFoM Academy for all E&T activities. The chair of this WP will be nominated by the SC.

- Development of an E&T Policy and programme
- Development and implementation of an (International) Masters Programme
- Development and implementation of an (International) PhD Programme
- Development and implementation of training courses for different stakeholder and future user

**2.4 Availability of resources to implement the roadmap**

Europe stands before some of the greatest challenges in its peace-time history: a health budget increasing by 80 Billion euros per year, and a mortality by multifactorial diseases increasing to more than 2 million per year, an exodus of life science industries and a relative absence of innovative industries in the ICT sector. And this in the context of highly successful academic science vis-a-vis both medicine and ICT.

ITFoM formulates a vision that, if taken seriously, would make Europe the centre of a multitrillion € industry delivering much improved healthcare. It would engage science, medicine and ICT in a journey through a wilderness of unprecedented complexity and diversity, to a very concrete ICT solution to the above challenges.

But can Europe pull this off? And, can ITFoM be the mechanism to make it do so?

We have built a 10-year plan for ITFoM that exemplifies how Europe can indeed succeed in this, by getting everything on board of this flagship that is essential for it to sail, and indeed for it to sail towards its final goal. We did not yet unleash the full potential of Europe; many more partners could be and shall be involved. After all, ITFoM will be the integration motor of what will be a multibillion flux of ICT-driven medical activities. But we have at least one partner on board or in close sight, for each essential activity. And indeed, we have assembled the optimal combination of partners assembled in this 10-year plan, optimal in the sense of eagerness to underpin the vision with their internationally most excellent expertise.

ITFoM recognises that curing disease has been ineffective yet expensive because we have been with shooting hail at an elephant: hurting and perhaps killing but not very specifically so. Both functional genomics and physiology have shown that any two unrelated human individuals are highly different.
For a start, most of their proteins are different and, they have not eaten the same. In addition, often the diseases they have might be superficially similar, but are very different at a molecular level, and therefore also in their response to therapy. Obviously then we should treat each patient differently from every other patient.

This may be obvious, but it is also impossible and in fact until recently unperceivable: even if we determine the difference between two individuals at the level of their complete DNA sequence, that sequence only indirectly determines their difference in functioning and the differences between their diseases. First, function is determined by many other things than genome sequence, such as nutrition and behaviour, but also different mutations in cancer, or exposure to different pathogenic agents in infectious or immune diseases. And second, even without this, the relationship between gene sequence and function is highly complex. The prediction of disease and of the success of therapy is a problem of unprecedented complexity, requiring the integration of enormous amounts of highly heterogeneous information into a highly nonlinear information processing system simulating the human body, which itself apparently processes the same information effortlessly.

What we need therefore is computing and data storage of unprecedented capacity and software that greatly streamlines the computations. We connect to main supercomputer centres of Europe (CERN, IBM Europe, CEA, UCL, INTEL, BULL, IMEC, EXASCALE). Members of ITFoM are involved in international and European HPC/Exascale initiatives like IESP (International Exascale Software Project), the EESI-2 project, the new European technology platform for HPC (ETP4HPC), the PRACE centres, the Integrated Sustainable Pan-European Infrastructure for Researchers in Europe (Inspire), the European Grid Initiative (EGI), and initiatives like CARRIOCAS, HELIX NEBULA, GEANT, ITER etc. We also connect to the most innovative new chip design by IBM, INTEL, Maxeler. We have the top of the computation architects on board (IMEC, INTEL, BULL, Maxeler) able to provide the most efficient computation on the basis of the massive amounts of data that emerge from the medical information. We avoid a bill of impossible energy consumption we can avoid by enlisting BULL’s HPC codesign activities. We connect with world’s experts on interactive visualisation tools to display the visual patients (IMPACT, CEA, UCK, Philips, Siemens and Uppsala University). We can address security issues in collaboration with BULL, CEA, Infineon, UCL and IBM Haifa. As to the enormous requirement of data storage we have leading centres such as EBI, UCL, IBM-Haifa, and the Technical University of Vienna. The enormous data integration challenge we can front with the international expertise of UNIMAN, EBI and UICE in biology minded data integration, management and standardisation with their consensus generating mechanisms that have generated an unprecedented degree of unification in traditional cottage industry of the life sciences. We have the top experts on data to object and objects to model on board (Maxeler, CEA, UCL, DCGMSB, UU). They provide the new ICT to propel the top experts in the four modelling strategies at dynamic integration of all these heterogeneous data into full models from molecule to patient (and back). Here we profit both from the advanced molecular systems biology (Berlin’s MPIMG/DCGMSB, and Amsterdam, Manchester, Groningen and Heidelberg Universities) and the physiome communities; we have the VPH consortium on board. The standardisation we can put in place through the various networks that we found eager to do this together with UNIMAN, VUA and EBI. This will revolutionise the interoperability of hundreds of European ICT and medical centres, also though the Systems Biology Infrastructure ISBE. Indeed ITFoM will interact with and profit from many of the large European Infrastructures established under the ESFRI (European Strategy Forum on Research Infrastructures) program (BBMRI, ELIXIR, EATRIS, ECRIN, ISBE for the Life Sciences, and e.g. PRACE and EGI for the ICT context), and will benefit from a number of IMI (e.g. EHR4CR, OncoTrack) and other large EU (e.g. BLUEPRINT) and international (IHEC, 1000genomes, ICGC) consortia it has associated. Outside Europe we can take advantage of close interactions to world leading research centres (Institute of Systems Biology, Seattle, Wireless Health Institute, Harvard Medical School, and Institute of Quantum Computing) contributing key insights and technology to our effort. We also connect with international experts in
public health policy making and indeed to policy makers at the European fora of Bad Gastein and Alpbach where we also meet with patient organisations. We have the epidemiological expertise on board to show the status quo of healthcare efficiency in Europe. And we have vivid contacts with some patient organisations.

All of this rocket ICT is not without fuel: an unprecedented amount of medical data that was begging to be integrated will now be integrated because of the virtual patient vehicle, with highly motivated medical doctors to see this happen in our WP3. They realise that all their tremendous expertise with disease is lost with every new generation because it is not captured in a predictive model of disease that learns (through top machine learning expertise we have on board at UCL and Tel Aviv) from their experience to better predict what happens with their next patient. We also connect to the patients who are now confronted with cryptic diagnoses of their disease, where we engage the ICT that translates the diagnosis into a movie of the patient in the future. And then we connect to the incredible data generators of today, ranging from deep sequencing (Sanger Centre, MPG, CNAG) and high resolution imaging (Philips, Siemens) to function (VPH).

With this, ITFoM will make more than one dream come true: it will also put in place a major mechanism through which the earth shattering progress in the life sciences of the last three decades will finally obtain the impact we always anticipated, through equally fundamental ICT breakthroughs: life science for life, through ICT, at last. The participation of big organisations for applied research in ITFoM (including organisations such as France’s 3.3 G€/y CEA, The Netherlands’ 0.5 G€/y TNO and the German Fraunhofer 1.8 G€/y) testifies to this. Also major organisations for fundamental science respond to the challenge, such as Physics’ 0.8 G€/y CERN and Molecular Biology’s 0.18 G€/y EMBL.

But above all perhaps we have a strong and coherent vision, a vision that motivates the entire crew tremendously, because it is so simple: treat each patient optimally by determining testing potential therapies first on a computer model of the patient, driven by the very best European (and international) ICT. A flag (truly individualised medicine) on a flagship (ITFoM) of world experts setting sail from Europa, powered the wind of the best genomics technologies, caught by ICT sails. Europe’s gift to the world.

2.5 Metrics proposed to monitor progress in the Flagship

The progress in the Flagship will be assessed by a sound monitoring system.

The monitoring will be done in accordance to the vision and the ultimate goal to develop a full virtual human and consider time, cost, resources, scope, quality, and activities that are executed in the frame of the program. Due to the long term of the Flagship a key element in the monitoring process is to keep flexibility and the ability to adapt to technological, societal, legislative and regulatory developments.

The milestones will be measured against

- Their accordance with the vision
- Scientific excellence
- Industry requirements
- Technological developments
- Medical needs
- Ethical standards
• Regulatory requirements
• The consortium performance and development as a whole
• Existing standards in different areas, e.g. in the analytical methodology, medical treatments, or patient sample handling and processing.

Medical needs will be defined in WP1.1, ethical standards and regulatory requirements are continuously monitored and assessed in foresight studies performed in WP1.5.

Further development of technologies and reagents for molecular analyses in clinical settings will be monitored in WP2. Future and emerging technologies in the ICT field will be monitored in WP3.9. This information will feed into the metrics.

All products and models emerging from the integration workpackage will be assessed against their fit with current regulatory requirements, in addition.

In general ITFoM also needs metrics to ensure scientific excellence throughout the program. These metrics will be the number of publication in peer reviewed journals, presentations in international conferences, awards, participation in prestigious boards of associations and committees, position as editor of journals, or as advisor for governmental, societal, or political institutions.

This will be ensured by the involvement of the external Advisory Boards (WP advisory groups, SEAB, Stakeholder Forum), which will include medical professionals, stakeholder representatives and industry observers beside expert scientists.

2.6 Coordination of activities and research communities

Europe has made a major effort to create excellent infrastructures and fundamental initiatives to foster new developments and innovation in the European Research Area. ITFoM is placed in this landscape and will benefit from these European, but also national and international initiatives to create an impact on whole Europe. Due to the large scale of ITFoM, the project will incorporate and cooperate with a variety of such initiatives to make use of existing efforts and combine them with the innovative approaches of the Flagship project. Several of the members of ITFoM are engaged in these research activities providing a strong basis of cooperation and ensuring communication needed for an optimal partnership.

The linking to and cooperation with other research initiatives and relevant infrastructures across Europe and beyond is essential for the success of ITFoM. The consortium will implement a strategic instrument in the governance structure to optimise the alignment with other relevant initiatives. This will be achieved via the Programme Coordination Committee (PIC) that is a separate entity within the governance structure (see 3.1). The highly visible positioning as an additional board within the governance structure is reflecting the importance that the ITFoM consortium is giving to the cooperation strategy. The main tasks of the PIC are the integration of relevant initiatives and the development of further joint activities to complement the expertise and the infrastructure of ITFoM and create synergies between all relevant players in the field. The PIC is also an instrument that will contribute to the further development of the ERA and make it a world leader in the ITFM relevant fields.
3) Implementation

3.1 Governance and scientific leadership

ITFoM merges cutting edge research from several disciplines assembled in a strong interdisciplinary team of leading experts who have developed the ITFoM vision and have the capacity to follow through in making this ambitious project a success. ITFoM has assembled the very best of Europe’s scientists in the areas of medicine/healthcare, analytical techniques, infrastructure, hard- and software, data pipelines, computational and algorithmic methodologies and data integration/modelling. The WP and SWP leaders have a track record of demonstrated leadership in large scale projects and can keep the Flagship on track in short term while forging new horizons for the future direction by incorporating new technologies and paradigms as they will emerge. To maximise the synergy through most productive interactions and knowledge transfer we are fostering a climate of open communication and organic governance with decisions made by the Steering Committee / Governance Council and the interaction between the various parties from academia, research and industry mediated by the Coordinator as explained in detail here and in WP7. The project is particularly unusual in the high fraction of partners working in a highly interdisciplinary fashion, simplifying the interactions across very distant fields.

Nevertheless the management of the intended ultimate project will be a challenge by itself. If all project components were to interact directly with all other components, then diversity in approaches rather than properly standardised interfacing would be the result. Therefore the underlying concept of organisation is that of ‘organic modularity’. The ‘organic’ implies that the modularity follows the natural organisation of the information flows in the modern Life Sciences. This ‘organic modularity’ concept is applied iteratively, such that the overall organisation structure emerges.

To most effectively make use of the current resources available across the ITFoM Consortium we will start by building on use cases reflecting particular medical problem areas of interest which already have proven results and integrate these at a level from which we can contribute a significant impact /breakthrough in the respective area by the ITFoM ICT-driven approach.

The participants in this proposal cover the entire range of expertise required to generate and engage the future ICT to generate the revolutionary changes in medicine, analytical techniques and computing that will transform the health sector from a centralised, entrenched innovation roadblock which disempower the patient, into an individualised, flexible ecosystem fostering innovation and empowering the individual to take responsibility for their own health.

To be able to handle the long term development components of the project appropriately, an appropriate governance structure will be implemented that answers to the size and scope of the initiative.

For the ramp-up phase, funded as CP-CSA within the EC FP7 programme, we anticipate to build a structure based on defined core centres covering the expertise needed for the first 30 months of ITFoM. The governance model will be described and agreed in the Consortium Agreement (including appointment of members, meeting and voting procedures).

For the 2nd funding period in Horizon 2020 the centre-of-centres structure will form the core of the project with the possibility to integrate new centres, groups, institutions and companies to collaborate for the achievement of the long-term goals of ITFoM. In this phase we intend to establish a special legal entity for implementation, namely a European Economic Interest Grouping (EEIG) unless the EC has provided a new special legal entity for FET flagships. Local centres (academic institutions as well as other legal entities) will form one local entity (very likely a limited company),

58
which will serve as the shareholder of the EEIG. This procedure will reduce the number of formally participating entities to a manageable number, will allow the involvement of local partners which otherwise could not be included as full participants due to liability problems.

For Horizon 2020 ITFoM proposes the following Principles for the Governing Structure. They are based on the desire of an effective management structure that honours the achievement of such structures in industry whilst ensuring that the original vision of the flagship is also honoured:

The core beneficiaries and new funding partners will form a Partner Group which will allow overall coordination with Programmes outside of the flagship project (similar to the PIC of the CP-CSA phase). For execution of the project, unless a new legal instrument is foreseen by the European Commission, we are thinking of establishing a legal European based entity, an European Economic Interest Grouping (EEIG). The shareholders of this company are all beneficiaries of the EU funding through local entities (most likely limited companies) as well as additional funders for the time of funding. This structure allows a flexibility to give funders a way to interact and decide on a global level. The Shareholder Assembly equals the Governance Council. In case a Funding Agency wants to bring in a programme or fund ITFoM it will receive shares of the EEIG for the time of funding.

The Scientific agenda within the EEIG is developed and then supervised by a Steering Committee (SC) similar to that during CP-CSA. Instead of a Programme Manager and the Executive Subcommittee (ESC) an executive structure in line with the legal framework will be established (e.g. CEO, COO, CFO). The details of such a structure will be developed during the ramp-up phase.

During both phases (FP7 and Horizon 2020) 20-25% of the EU funding will be reserved for new partners. There are 2 ways foreseen to bring in new projects and new partners but both will be implemented through open calls: (a) new interesting projects can be submitted to the SC, which will decide on a full application. Once a positive review is done (for the review PTJ will serve as the organisational unit), the new project will be integrated in the running programme and the new legal entities can join the Consortium through an Amendment of the EU grant. (b) for certain software and technology developments the instrument of subcontracting will be used. In such a case the results can be incorporated into the project without changing the EC Grant Agreement. Such a procurement procedure will follow EU and National Member State jurisdiction.

ERA-NETs and other regional or national programmes can be aligned to ITFoM through the PIC and the Partner Group. A MoU/contract will detail how the programmes complement each other and how resources can be used.

During the ramp-up phase a classical project-oriented governance structure will be set-up based on existing and well-working model structures, which will consist of a Governance Council with two supporting Committees (Scientific and Ethical Advisory Board and the Stakeholder Forum); a Steering Committee and associated a Programme Integration Committee; and an Executive Subcommittee.
For the CP-CSA phase the following Governance Structure will be implemented:

![Governance Structure Diagram](image)

**Figure 3 Governance structure for the CP-CSA phase**

### 3.2 Management plan

For ITFoM the consortium will set up an appropriate management that is in accordance with the scope, the magnitude and the ambitions of the programme and ensures the integration of all ITFoM stakeholder groups (including e.g. industry, representatives of the healthcare system, regulators and policy makers) to allow an optimal transition along the line of the value chain until implementation in the clinical practice.

The management structure and measures need to be stringent and robust, but with a flexibility to respond adequately to the scientific, technological, societal and political changes expected during a 10 year research programme in general and in highly dynamic research markets targeted in ITFoM in particular.

For the overall ITFoM work plan the PM supported by WP7.1.1 Scientific Project Management and in cooperation with the WP leaders will develop a detailed milestone plan engulfing the entire value chain (WP1 to WP6) including definition of interfaces between WP1 to WP6 and the respective time lines. As ITFoM is a highly complex and long running programme comprising a series of different activities within the workpackages that are linked to and depend on each other it will require a careful planning of all activities in timelines and milestones. PERT diagrams will be developed based on the work plan and milestones to identify the activities on the critical path requiring particular monitoring initially in the CP-CSA and later in the full flagship programme.
3.3 Expertise of individual participants and quality of core consortium as a whole

The ITFoM Consortium was assembled with the aim of bringing together in a synergetic cluster the most competitive and respected European scientists leading cutting edge world class future and emerging technologies research in ICT infrastructures, high performance computing, data integration and modelling, the application of large scale analytical techniques (genomics, transcriptomics, epigenomics, proteomics, metabolomics, imaging, sensors etc.) as well as many aspects of medicine, ranging from medical samples (BBMRI) to public health/health economy/personalised medicine (iHealth) experts. The consortium unites groups currently carrying out large scale data production in biology/medicine with groups in informatics/machine learning/systems biology and medicine, groups carrying on fundamental research in institutes and university research centres with SMEs and the major industrial corporations which currently dominate many of the relevant ICT and biomedical areas commercially.

Since effective collaboration cannot be enforced but requires all parties wanting to work together, it is worth noting that a proportion of the ITFoM partners already have a proven collaboration track record from previous and on-going EU and otherwise funded projects. A small number of selected large-scale projects merit extra mention in this context. ITFoM participants collaborate in large international projects like the 1000 Genomes Project (Sanger, EMBL-EBI, MPG), ICGC - the International Cancer Genome Consortium (OICR, MPG, EMBL-EBI, CNAG, Sanger, CEA, DFKZ, UIO, CCCC, KI), IHEC - the International Human Epigenome Project (UCL, CNIO, EMBL, Sanger, MPG, CNAG, ONT, INSERM, UEDIN, UNICO, UNIGE, BSC), in various Innovative Medicine Projects like EHR4CR - Electronic Health Records Systems for Clinical Research (EMBL, INSERM, UCL, UNIGE, UNIMAN, WWUM, UNIGL, UEDIN, BAY, SA, and Janssen), Oncotrack - Methods for systematic next generation oncology biomarker development (MPG, BAY, MUG, UU, UCL, IPRI, ATG, CCCC, Janssen, Roche), in large infrastructure projects like ESFRI, BBMRI (MUG, IPRI, MPG, KI, INSERM, UNIMAN, EGCUT, TIGEM, IBMT, CU, AAU, IMER, UMCG), ELIXIR - Life Science Infrastructure for Biological information (Sanger, EMBL-EBI, CNIO), EATRIS (UNICO, DFKZ, ISS, UIO), ISBE - Infrastructure for Systems Biology Europe (UNIMAN, UvA, DFKZ, ICL, MDC, VUA, UNILJ, EMBL-EBI), ESGI – European Sequencing and Genotyping Infrastructure (MPG, Sanger, EMBL, CEA, INSERM, UU, MUG), READNA - Revolutionary Approaches and Devices for Nucleic Acid analysis (CNAG, MPG, ONT, UU, ULEIC), BioMedBridges (EmBL-EBI, KI, VUA, INSERM, UNICO, MUG), PRACE - Partnership for Advanced Computing in Europe (VSb, BSC, CEC, PSNC, UNILJ, ETH). In addition to PRACE, various members are already involved with the most important international and European HPC/Exascale initiatives like IESP, the EESI-2 project, the new European technology platform for HPC (ETP4HPC), the European Grid Initiative (EGI), and the US Government Exascale activities. These initiatives have all identified life sciences and health as key applications for exascale. ITFoM will take a lead in the definition of application requirements and the implementation of solutions for the exascale era. In particular ITFoM will address many of the topics lined out in the EESI report on life sciences and health.

These already established connections can be expected to bloom into even larger interdisciplinary synergies and spread to encompass international collaborations. ITFoM intends to be as inclusive as possible with respect to associated partners and also envisions additional full partners to join the consortium during the lifetime of the project. In order to facilitate this, ITFoM will keep approx. 20% of the overall budget as non-allocated reserve. In addition to this we will set aside 7% of the overall budget for including new expertise by subcontracting (see 3.4).

The WP leaders are most reputed scientists in their respective areas, as follows:

Hans Lehrach, the coordinator of ITFoM, has spent most of his time well in front of the fields he worked in. He was among the first to develop automated systems to map and sequence genomes, to
carry out functional genomics on a large scale, and to use Systems Biology to generate predictions from genomic data. In this area he is used to coordinate large-scale collaborative projects. He is director at the DCGMSB and the MPG. He is partner in two International Cancer Genome Consortium (ICGC) projects and member of the ICGC Technologies Working Group. They coordinate the EU I3 European Sequencing and Genotyping Initiative. The department participates in the 1000 Genomes Project and is the leader of the managing entity of the IMI project OncoTrack, a private-public partnership on oncology biomarker development, and coordinates the FP7-project ADAMS. Hans Lehrach is partner in several FP7 projects (READNA, GEUVADIS, APO-SYS, Trireme, and BBMRI). His group has a strong expertise in next generation sequencing, systems biology as well as in the analysis and modelling of cellular processes with respect to development and diseases. The research is driven by the integration of diverse ‘-omics data, as generated by current (high-throughput) technologies.

Kurt Zatloukal, M.D. is professor of Pathology at the Medical University of Graz and coordinated the preparatory phase of a European biobanking and biomolecular research infrastructure (BBMRI) within the 7th EU framework programme. He also leads in the FP7-funded large integrated project SPIDIA the development of new European standards and norms for pre-analytical processing of tissue samples. In the context of the Austrian Genome Programme and several nationally funded projects the group of K. Zatloukal has developed innovative solutions for extraction of data from medical records, data integration, data clean-up and visual analytics. He was member of the OECD task force on biological resource centres and the Roadmap Working Group of the European Strategy Forum on Research Infrastructures. Moreover, he contributed to the OECD best practice guidelines for biological resource centres, the regulations for genetic testing of the Austrian Gene Technology Law, and was member of Bioethics Commission at the Austrian Federal Chancellery.

Nora Benhabiles manages the business development for the health sector and is responsible for the health program in the CEA List institute in Saclay (FR). Among former positions in Europe, she was head of the bioinformatics and IT department in a biotech developing drugs for cardiovascular related diseases (FR), Visiting Professor of Biomedicine in the Katholieke Universiteit of Leuven (BE). She is expert for the EU commission (FP7 Health) and for the French National Research Agency (ANR). She is vice-president of the ANR Emergence committee (president health sub-section, 2012) devoted to emerging innovative R&D programs.

Oskar Mencer, CEO Maxeler Technologies and Head of the Computer Architecture Research Group, Imperial College London, Consulting Professor in Geophysics at Stanford University. Prior to founding Maxeler, he was at the Computing Sciences Center at Bell Labs in Murray Hill. His work received two Best Paper Awards, an Imperial College Research Excellence Award in 2007, first place in the American Finance Technology Awards in 2011, and a Special Award from Com.sult in 2012 for "revolutionising the world of computers”.

Ewan Birney is Associate Director of the EMBL-EBI. His background was in biochemistry but has been a leading member of the bioinformatics community since 2000. He has developed a number of algorithms in widespread use in molecular biology (e.g., GeneWise, Exonerate and Velvet) and founded the bioinformatics resources of Ensembl and Reactome. He has participated in a number of large genomics consortia including the Human, Mouse, Rat, Anopheles and Chicken genome projects. In particular he lead the ENCODE Pilot project (published in 2007) and leads the ENCODE scale up project which is building a comprehensive catalogue of functional elements in the Human genome. Via the leadership role of EMBL-EBI he is heavily involved in the ELIXIR ESFRI project to develop a sustainable bioinformatics infrastructure in Europe, and the BioMedBridges project to link the different life science ESFRIs together. He has won a number of awards, including the Crick Lecture (Royal Society, 2003), the Overington Prize (ICSB, 2005) and was elected as an EMBO Fellow in 2012.

Mark Girolami is a world leader in the development and deployment of advanced statistical methodology theory and practice. He is Professor of Statistics in the Department of Statistical
Science, holds a professorial post in the Department of Computer Science and is Director of the Centre for Computational Statistics and Machine Learning (CSML) at UCL, a large multi-faculty centre of world leading research excellence in statistical science, statistical learning theory, and statistical machine learning. His research, and that of his group, addresses the theory, methodology and application of Computational Statistics and has very strong multidisciplinary interactions with the life, clinical, physical and engineering sciences. Girolami is one of a small handful of Computing Scientists and Statisticians that has contributed at the interface of the Life and Medical Sciences and Mathematical and Predictive modelling and his Machine Learning books are considered the authoritative textbooks in several universities around the world.

Hans Westerhoff has been a Systems Biologist avant la lettre. He is AstraZeneca Professor for Systems Biology at the University of Manchester and Professor of Microbial Physiology at VUA and directs both the MCISB and the DTC-ISBML. He is also effective co-leader of the Systems Biology ESFRI Infrastructure for Systems Biology Europe (ISBE). Westerhoff is coordinating a Systems Biology project that includes five European nations (MOSES) and involved in 6 other such transnational research projects. Westerhoff has been chair of the Advisory Board of the multimillion German (BMBF) Hepatosys research program, as well as of the committee advising on the creation of its successor 'The virtual Liver' project. He is also one of the four scientific advisory board members of the €100M Luxemburg systems biomedicine program, as well as Director of the biennial FEBS Advanced Lecture Course on Systems Biology.

Other key participants bring very complementary skills and expertise to the consortium.

Peter Boyle is President of the International Prevention Research Institute, Lyon, France and is Founder and President of the “World Prevention Alliance”, a non-governmental organisation for prevention research and actions in lower income countries. He was the editor of the World Cancer Report 2008 highlighting the growing global cancer crisis. He holds honorary or adjunct professorships at Oxford University, Yale University, Glasgow University and Dundee University.

Markus Pasterk is a highly experienced research manager, currently holding the post of a COO and VP/Science at IPRI. He previously was the Scientific Coordinator of IARC, Lyon, and acting Director General for Research in the Austrian Federal Ministry for Science and Research.

Angela Brand, Founder and Full Professor of the Institute for Public Health Genomics (IPHG) at Maastricht University, and Adjunct Professor at the Manipal Life Sciences Centre of Manipal University, India. Director of the European Centre for Public Health Genomics (ECPHG), Coordinator of the Public Health Genomics European Network (PHGEN), Executive Director of the Public Health Genomics international network GRaPH-Int.

Ruth Chadwick is Distinguished Research Professor and Director of the ESRC (Economic and Social Sciences Research Council) Centre for Economic and Social Aspects of Genomics (Cesagen), Cardiff University, UK. She is Chair of the Human Genome Organisation Committee on Ethics, Law and Society.

Tim Hubbard, Head of Informatics at the Wellcome Trust Sanger Institute, co-founder of the Ensembl genome database project and of the Structural Classification of Proteins (SCOP) database, PI of the GENCODE human genome annotation project of the ENCODE consortium and co-PI of the Genome Reference Consortium, Chair of advisory board of UK Pubmed Central and member of the UK Expert Advisory Group on Data Access.

Since January 2010, Ivo Gut, PhD is the founding director of the Spanish National Genome Analysis Center (CNAG, www.cnag.eu) in Barcelona. CNAG is one of the largest nucleic acid sequence analysis centres in Europe with sequencing capacity of more than 600 Gbases a day. It has matching data analysis capability using a 1000 core supercomputer and 1.2 pbytes of storage. In addition it benefits
from direct connections to the Barcelona Supercomputing Center. In the 11 years prior to CNAG, he was associate director and in charge of technology development at the Centre National de Genotypage in France.

**Rudi Balling** coordinated EATRIS, the European Initiative for Advanced Translational Infrastructure. He was also the coordinator of the Bill and Melinda Gates Grand Challenge Projects in Global Public Health (GC4) and one of the coordinators of the German Human Genome Project.

**Leroy Hood** is one of the world’s leading scientists in genomics, molecular biotechnology and systems medicine. He is President of the Institute for Systems Biology in Seattle, United States, since 1999.

**George Church** is the American molecular geneticist. He is Professor of Genetics, Harvard Medical School, Director of the Center for Computational Genetics and developed the first genomic sequencing method (1984). He has been honoured for his work on several occasions and has been acting as advisor for more than 20 companies and founded/co-founded several biotech-start-ups.

**Peter V Coveney** is Director of the Centre for Computational Science, an Honorary Professor in Computer Science (both at UCL) and Professor Adjunct at Yale School of Medicine. He exploits state of the art HPC, steering and visualisation methods and develops data warehouses to manage medical data. He is leading the VPH Network of Excellence within the EU FP7 VPH initiative. He is Director of the UCL Computational Life and Medical Science Network, a member of the UK Government's E-Leadership Council and advisor to the Prime Minister's Council on Science and Technology.

**Mihaela Ulieru** is an expert in ICT-enabled innovation and President of the IMPACT Institute for the Digital Economy which she spun from her internationally praised work as the Canada Research Chair in Adaptive Information Infrastructures for the eSociety. She founded the ICST IT Revolutions Forum and Conference series and the IEEE Industrial Informatics Community. She had numerous appointments on a wide range of scientific boards and councils, among which the Science Technology and Innovation Council of Canada, the Singapore ASTAR and UEFISCDI - the Scientific Research Financing Council of Romania, as well as on several EC FET Consortia among which PERADA (Pervasive Adaptation) and IPROMS (Innovative Production Machines and Systems).

**Ron Shamir** is the Raymond and Beverly Sackler professor of Bioinformatics at the School of Computer Science and head of the Edmond J. Safra Center for Bioinformatics, Tel Aviv University. He develops novel algorithms to advance biology. He specialises in gene expression analysis, gene regulation, systems biology and disease bioinformatics. He is on the editorial board of eleven scientific journals and series, and has published over 220 scientific papers.

With one of the first PhDs in Systems Biology (from VUA) and coming from a combined experimental and computational background, **Frank Bruggeman** now leads the junior investigator group at the Netherlands Institute for Systems Biology in Amsterdam. He is one of the world’s experts on mechanistic bottom-up systems biology, and on noise and timing in transcription regulation and deputy leader of WP6.

**Marco Viceconti** is Professor of Biomechanics at the Department of Mechanical Engineering at the University of Sheffield and Scientific Director of the Insigneo Research Institute. Marco Viceconti is one of the key figures in the emerging Virtual Physiological Human (VPH) community: co-author of the first white paper on VPH, he is also currently chairing the Board of Directors of the VPH Institute

**Patrick Soon-Shiong**, Member of ITFoM Scientific and Ethical Advisory Board, is the Executive Chairman and Chief Executive Officer of Abraxis Health. He was recently appointed Executive Director of the UCLA Wireless Health Institute, and is Professor of Microbiology, Immunology, and Molecular Genetics Professor of Bioengineering at UCLA. Soon-Shiong’s research has been recognised by national and international awards such as the Association for Academic Surgery Award
for Research, the American College of Surgeons Schering Scholar, the Royal College Physicians and Surgeons Research Award, the Peter Kiewit Distinguished Membership in Medicine Award, and the International J.W. Hyatt Award for Service to Mankind. Soon-Shiong received the 2006 Gilda Club Award for the advancement of cancer medicine and is a recipient of a 2007 Ellis Island Medal of Honor as well as the St. Mary Medical Center Life Achievement Award in 2007 and the St. John’s Health Center Caritas Award in 2007. In 2008 he received the Medical Visionary Award from the Pancreatic Cancer Action Network for his work in pancreatic cancer. In 2009 he was appointed to the President’s Council at RAND Corporation, Chairman of the Steering Committee of Life Sciences of the X-Prize Foundation and Founding Board member to Dossia Foundation.

Jonathan Knowles, Member of ITFoM Scientific and Ethical Advisory Board, is Professor of Translational Medicine at EPFL in Switzerland, holds a Distinguished Professorship in Personalised Healthcare at Finnish Institute for Molecular Medicine (FIMM), and has been appointed to a Visiting chair at the University of Oxford. He is a Member of the European Molecular Biology Organisation and serves on the scientific advisory boards of a number of publically funded initiatives.

Barbara Prainsack, Member of ITFoM Scientific and Ethical Advisory Board, is Professor of Sociology and Politics of Bioscience at Brunel University, and Honorary Senior Research Fellow at the Department of Twin Research and Genetic Epidemiology at St Thomas’ Hospital, King’s College London. She chairs the Scientific Committee of the ESF Forward Look on Personalised Medicine for the European Citizen and is a member of the Austrian National Bioethics Commission advising the federal government in Vienna.

Michael Morgan, Member of ITFoM Scientific and Ethical Advisory Board, was director of Research Partnerships & Ventures at the Wellcome Trust. He played a major role in the international coordination of the Human Genome Project and was also responsible (as Chief Executive) for developing the Wellcome Trust Genome Campus. In 2006 he was appointed Chief Scientific Officer of Genome Canada. He retired in 2009.

3.4 Approach to flexibility of partnership and involvement of relevant actors

ICT has developed exponentially over many decades, resulting in continuous, dramatic improvements in performance/cost ratios. With our new ICT we expect to initiate a new strand of ICT development. Because we appreciate the tremendous complexity of the ‘virtual individual/virtual patient’ models to be built, ITFoM will put in place an institution that can support the ITFoM mission over many decades. The phase described in this report should be seen as a start that develops ITFoM into a self-funding program. The VPH institute will inspire ITFoM.

Partnership management needs to deal with a number of paradoxes. The first is that on the one hand, to remain at the cutting edge of development, mechanisms able to attract and incorporate new partners are required. Conversely, in long term projects with a significant development component, longer term funding mechanisms for critical core groups will be needed, to ensure that the original flagship vision materialises. Only then long term commitment by the key groups driving the process is feasible, in a project that can then focus on ultimate patient benefits rather than rapid publications. A second paradox is that to obtain the best possible virtual human models, ICT technology and information from a wide variety of sources needs to be integrated; a community effort is required that is orthogonal to an interaction model that best protects IP. ITFoM therefore aims at much activity in a precompetitive arena, with commercialisation occurring in one-ones between industries and academic partners. Here some financial return on investment into the ITFoM core project is expected on the basis of the principle of fairness. ITFoM governance council will
instate an independent fairness committee to which ITFoM members should declare authority. A third paradox is that some but not all of the core funding comes from the FP7 and H2020 programs, in ways that still need to be established, whereas much is to come from national funders, who will then also wish to have a say in the partnership composition of ITFoM. These paradoxes and the long time frame require flexible partnerships.

**CP-CSA phase**

During the CP-CSA funding period of FP7, 4 mechanisms will guarantee inclusion or alignment of new partners, projects or programs.

(1) A special Committee, the Programme Integration Committee (PIC) is responsible for alignment and integration of outside funded projects and programmes. The ERA-NET Plus will be coordinated through this PIC. In addition in some Member States similar initiatives are running or in a planning phase. We intend to invite the programme owners to discuss common strategies, joint activities and will be able to co-fund core activities. The PIC will consist of an integration of the ITFoM Steering Committee (SC) and SCs of the associated research programs.

(2) 13% of the EU core funding will be reserved to bring in new partners based on an open call for new ideas, concepts and technologies relevant to ITFoM. First phase proposals will be first reviewed by the steering committee to ensure relevance to the goals of the project. Full proposals will then be reviewed externally by a panel of external referees to ensure technical expertise, and ultimately voted on by the steering committee, either directly, or after an evaluation session with the applicants. Partner PTJ will be responsible for the implementation of this process. PTJ in coordination with the Steering Committee will be commissioned to organise the call and its evaluation process. Such new projects will be integrated into the existing Workpackages and the legal entities will join the consortium based on amendments to the Grant agreement.

In addition, in some cases, we anticipate the inclusion of national partners, to be able to provide ITFoM infrastructures in additional countries. Such national partners will be selected in close interaction between ITFoM and national funding bodies, to select groups co-funded by both bodies in an appropriate ratio.

This funding stream will also be used to stimulate the sharing of research and competitive practices between academia and industry, to discuss R&D opportunities, to identify new development, and to identify new partners through the funding of meetings.

(3) 7% of the EU funding is reserved for subcontracting of research work. Specific software and/or technology development can in such a way be bought in. This will allow new or not existing expertise to be integrated into the project. Such calls for tender will follow EU and Member State regulations and will be formulated by the SC.

4) Research programmes outside of ITFoM will be invited to joint ventures, in which cases their SCs will join up in the PIC. Core funding of both programs will then be used on an overall parity basis to facilitate cross fertilisation, such as in the sense of postdocs and technicians appointed in both programmes. In addition to these more formal mechanisms, the overall structure of ITFoM will be very open for scientific interactions at all levels. Already now we have many more than 100 associated partners: institutions with groups, which can participate in the intellectual challenge of discussing strategies, participate in ITFoM meetings, and propose projects to include in the work plan.

Well defined interfaces to the data sets generated on individual patients will allow groups both within, but also outside of the consortium, to take advantage of the open infrastructures, standards and interfaces allowing both academic groups and industry to interact with ITFoM data sets, resources and pipelines in their own developments, based on open standards and well defined
interfaces, allowing a large degree of ‘plug-and–play’ interactions with the ITFoM structure. Indeed, all external communities will be invited to participate in the multiple ICT standardisation and sharing activities of WPs6.6-6.9. ITFoM will pay for travel and subsistence. In the CP-CSA phase therefore 20% of the core funding will be open to partners outside the consortium.
4) Impact

4.1. Impact on Science and technology

Medicine

The medicine of today is practiced in a ‘one size fits all’ manner, inherently statistical, with therapies selected on the basis of clinical trials performed on large, often heterogeneous groups, often resulting in the selection of therapies, which are inappropriate or even detrimental for specific patients belonging to this group. As a next step, stratified medicine based e.g. on the use of biomarkers aims to reduce group heterogeneity, and increase response rates to therapies. However this is still far removed from truly individualised treatment.

ITFoM takes a frontal, radical approach to personalised medicine aiming on one side to address the specifics of each individual by modelling precisely his/her condition and on the other side offers an unprecedented tool, the ‘virtual I’ holding every person’s key to their own health and wellbeing, a tool empowering each of us to access essential information about our current condition, to monitor the evolution of our health and predict eventual illness and how it will unfold under various alternative therapies. This will transform medicine in a fundamental manner by transferring the accountability to each individual for their own health and by this will significantly unload the medical system from its current burden since everyone will be in a position to prevent eventual illnesses and take timely, immediate action to cure incipient conditions and maintain an optimal status in case of unavoidable conditions, through continuous remote monitoring via integrated wireless sensor technologies, and remote contact with the nurse or physician e.g. via Skype.

Drug Discovery and Development

ITFoM will have a strong impact on pharmaceutical research and development model from empirical to knowledge and inference based. The current pharmaceutical model i.e. blockbuster drugs based on common targets e.g. COX2 and compounds identified through high throughput screening was quite successful in the second half of the twentieth century and made the industry one of the most profitable industries of all time. However, these common targets have mostly been discovered, and the elucidation of genome level information for each person makes knowledge and inference based drug discovery and development, not only possible, but necessary. The only requirement for this approach to be successful is that the inference models are effective and there is now wide acceptance in the field that they are becoming more and more so.

Clinical trials will become much cheaper and the resulting drugs much more effective and safe for their targeted populations. ITFoM will work with regulators such as the EMA (European Medicines Agency), national HTA agencies and pharmaceutical industry organisations like EFPIA to ensure that important information about the drugs that are developed are made available to clinicians and healthcare providers so they can more intelligently prescribe the necessary drugs and drug combinations. Some of the data necessary are already generated by pharmaceutical companies as part of their development, such as inhibition profiles for the targets of their drugs. Others would need to be added as standard practice, such as X-Ray or NMR structures of their drugs bound with the appropriate CYP450s and a full determination of the metabolism of their drugs.

The radical transformation which ITFoM will bring is intelligent target selection ranging from stratified to truly individualised approaches. With the genomes of vast numbers of consenting patients available, any putative target can be examined carefully for its penetrance in the population as well as the presence of other genetic polymorphisms that enhance or reduce the effectiveness of the particular target for a particular indication. Once a target has been selected and potential hits
identified, either empirically through screening or computationally through modelling, the properties of the lead can be intelligently optimised to increase its efficacy and reduce side effects. With this everyone will be treated for the particular manifestation of a disease on their own particular condition as such avoiding the inefficacies and inconveniences, sometimes fatal, typical for the current ‘trial-and-error’ style of today’s therapies. ITFoM will ensure that everyone will receive what works best for them at the right time and for the right condition.

Diagnostics

For the ‘right’ condition to be diagnosed correctly, ITFoM will bring a revolution in medical understanding of the complexity of each individual. The empirical and stochastic approaches to biomedical science have reached their limits but a human being is an incredibly complex organism and to move biomedical science to a knowledge and inference based science is a challenging task. We currently deal with this complexity by understanding ourselves at multiple levels without knowing the detailed interfaces between the levels. For instance, we measure blood pressure, body temperature, weight etc. to get an understanding of global health, we may then monitor organ health by measuring blood chemistry surrogates etc., and possibly at the cellular level we may genomics or transcriptomics.

The links between these levels are currently determined empirically and stochastically, for instance, it had been noticed that different people react differently (i.e. some bleed unacceptably putting them at risk for stroke) to the same dosage of a blood thinning drug. ITFoM will enable immediate access to the right knowledge about the patient’s protein structure which will enable the doctor to prescribe the appropriate blood thinner which is effective for this particular patient, avoiding the dangerous, even fatal side effects of the wrong blood thinner.

The challenge here is presented by today’s entrenched medical procedures and regulations rather than technical, since every aspect of this future scenario could be accomplished using today’s computational technologies, computational models, and clinical and biomolecular analysis technologies. What is needed however to put this available technologies in action is a quantum leap kind of revolution both in attitude and in ICT capabilities.

Computational technologies and paradigms

Currently, the ICT research and clinical computing worlds are completely separate, with very different computational paradigms. Companies like Complete Genomics (www.completegenomics.com) or Structural Bioinformatics (www.strubix.com) live in the research world and dedicate many millions of dollars to computer systems for solving research problems that are of the same type and order of magnitude that will be faced routinely in the clinic within the next decade. However personal genomics products which offer access to one’s individual genome on their iPhone are already available and this trend will accelerate the revolution in the clinics by making such essential information easily accessible to the clinicians who will be pressured to use it in their practices. ITFoM will have a tremendous impact on and stimulate the development of novel ICTs even beyond what is available today in the lab but also in industrial settings. The computing requirements will remain a source that will push such novel technologies to be further developed.

We anticipate that the level of computational need in the clinic will be so great that there will be opportunities for partners in this proposal like Intel, IBM and Maxeler to develop special purpose chips (e.g. FPGA’s) for certain types of simulations, for partners like IBM, Microsoft, and HPC-Europa2 to develop easy to use high performance compute clusters optimised for biomolecular medical modelling, for partners like NVIDIA and HealthSolve to develop visualisation systems and user interaction paradigms for new (mobile) devices for healthcare systems providers like proposal partners Philips and Siemens and for policy makers using new models such as the LAL model to build
integrated assay/analysis/modelling systems, perhaps using arrays of FPGA’s combined with more standard highly parallel general purpose CPU’s.

**Storage technologies and methodologies**

Each patient in the new world of molecular diagnostics will have multiple terabytes of molecular and high resolution imaging data associated with their clinical records. This data will need to be kept securely as well as made available rapidly when necessary. ITFoM will push the development of revolutionary solutions *e.g.* by a combination of new technologies, for instance holographic storage, biologically intelligent encryption algorithms (*e.g.* genome data could be stored as the differences between the patient’s genome and a reference genome), and new data architectures which intelligently separate and handle raw molecular data (*necessary sometimes*) and various levels of abstraction for post processed data. A number of ITFoM partners have deep interest in developing the next generation of these technologies, in particular Bull, IBM, Siemens, Microsoft, and Intel. Additionally, all of these types of data will need to be integrated seamlessly into e-Health records centred on the individual who will need to be re-architected to support these large datasets, an area of keen interest for partners IBM, HealthSolve, and Microsoft.

### 4.2 Impact on economy and society

There is a huge need in global healthcare for novel and innovative approaches for clinical research. In the present post-genomic era fundamental new diagnostic tools and treatments, sometimes based on very recent breakthroughs in fundamental research need to be implemented and made available for the clinics. Especially the developments in systems biology and more recently first results in the field of synthetic biology are coming close to potential clinical applications. All these new technologies share the character of tremendous interdisciplinarity in their basics but also in potential applications. They also share an imperative need for IT support to a unique extent. ITFoM will contribute to these challenges by combining excellent IT science with a unique expertise in the field of post genomic molecular biology and systems biology. Its financial basis and its scientific technological set-up will allow to building a critical mass of basic science in wet lab and IT as well as to creating excellent opportunities for clinical development. Thus, ITFoM will represent a real “flagship” for IT in medical surrounding. It will significantly contribute to the development of the European Research Area by improving the coherence and coordination of research in IT and medicine across Europe. In addition, ITFoM aims at empowering the patients to take accountability for their own health, via direct access to essential information and knowledge. By this proactive preventative approach enabled by ITFoM, ITFoM will reduce the current burden from the medical system through a significant reduction of many ailments as well as through most effective treatments when these are necessary.

**Potential contribution to EU competitiveness**

ITFoM, based on the unique and highly innovative LAL model, can be seen as a raw model or best practice for all industries in Europe and beyond, since it will accelerate and increase not only the likelihood of successful market introduction of personalised health interventions, but also of their application in the healthcare system as a whole. While challenging existing European frameworks for assessing effectiveness of healthcare interventions, it will provide solutions meeting the Europe 2020 goals of growth, innovation and social inclusion, as well as contributing to the already mentioned Innovation Union and also to the Horizon 2020 goals. ITFoM promotes and supports
- research that is cross-sectoral and interdisciplinary, involving both technological and social innovation
- research that acknowledges the ‘right to health’ and European values in health such as equity, universality and access
- innovation stimuli which takes into account the special needs of product development, including long lead times, high intensity of investments and high attrition rates.

With its focus and ultimate objective to deliver “good health”, ITFoM will be paramount to achieving the objectives of Europe 2020, the Innovation Union and Horizon 2020 with which it is perfectly aligned by delivering (1) research for health, (2) innovation for health and health equity, and (3) significant contributions to the global research and innovation system.

1. Research for health: ITFoM combines ICT with disciplines “from cell to society” to understand the impact on health of policies, programmes, processes, actions or events originating in any sector – including, but not limited to the health sector itself. It assists in developing interventions to help prevent or mitigate that impact, and contributes to the achievement of health equity and better health for all.

2. Innovation for health and health equity: ITFoM promotes ICT driven disruptive novel ideas, inventions and processes and applies them to achieving improved health and greater health equity. This is achieved for European’s citizen by better access to information for individual self-management. A combination of social and technological innovation is essential to respond to the Europe 2020 agenda.

3. The global health research and innovation system: ITFoM encompasses a comprehensive systems perspective to guide efforts to achieve greater effectiveness and efficiency. ITFoM will include health impact assessments and the development of new indicators to take into account the impact of individualised medicine on health. In addition, ITFoM aims to create standardised methods for the assessment of health status and for data interpretation and comparison on individual as well as societal level.

4. Similar to the biological networks within the project ITFoM will also establish a network model of regulatory, reimbursement, societal and governmental needs as well as individual needs. By this it will lead to ICT developments, which will not only strengthen the European position, but may also lead the global way.

Thus, ITFoM is of high European added value, since the project results are expected to:

- strengthening the EU’s economy through support for a key but challenged area of Europe an innovation
- providing competitive advantage for European industry and research
- improving health for Europeans and globally, with consequent positive effects on health systems
- sustaining EU credibility with regard to commitments made across a range of EU policies, including those on health, economic growth, social inclusion and development
- promoting European goals and values in health
- facilitating global health, knowledge sharing and common solutions to problems.

By this, ITFoM is establishing a blueprint for a pan-European personalised medicine and healthcare paradigm:
To gain acceptance at a governmental and regulatory level, the benefit to society must outweigh the increased initial healthcare costs. In reimbursement the additional benefit to the individual is of utmost importance. The current stratified medicine approaches offer only limited benefits at high costs as they still treat only subgroups of patient so non-responders will be still included.

ITFoM will be the first to develop truly personalised medicine. By treating each patient individually he receives a true personal benefit. Side effects go down and less wrong-targeted therapies happen. By this a cost-benefit assessment will show further benefits of ITFoM technologies.

**Impact on EU Health Policy Development and Policy Implementation**

As a European large-scale scientific initiative, ITFoM is in line with the Horizon 2020 strategy for research and innovation and the Europe 2020 vision for an Innovation Union and will contribute actively to the implementation of the Horizon 2020 strategy:

- ITFoM is an ambitious project at the crossing of two major scientific and innovation fields: Information and Communication Technologies and medical research, bringing key enabling technologies into the health and well-being sector

- ITFoM brings together the most dynamic European and international labs worldwide with top leading companies, and will therefore avoid fragmentation and duplication of the research and innovation effort to reach the goal of personalised medicine, for which reaching critical mass is a requirement

- ITFoM is a powerful driver for the development of niche markets where SMEs play an important role, such as the medical device market

ITFoM will have a high impact in Europe at different levels:

Socially, with high impact on health, well-being, ageing population and social inclusion, by providing grounds to propose personalised medicine for all European citizens

Economically, by supporting European companies (large companies, SMEs and start-ups) to maintain and take leading positions in the health and IT market of the future with the implementation of new value chains in key industrial sectors such as IT, pharmacy, and medical devices

Scientifically, with high impact expected in biology, thanks to an increased capacity to understand the human body, and in IT, with the development of new hardware and software architectures to answer to the complexity of the human body.

ITFoM contributes to the implementation and completion of the European Research Area by its very design: It is a consortium consisting of academic institutions and industry partners from a plurality of countries collaborating to generate knowledge which circulates freely. ITFoM, thanks to its inclusive governance, materialises the ideals of the European Research Area and deepens its design. It will act as catalyst of the current and future European research activities linked either to biology infrastructure or diseases oriented activities. It provides an opportunity to make full use of new scientific knowledge and an opportunity to translate it into market opportunities.

The strong medical and health focus inherent in ITFoM is a specific asset with regards to tackling the great challenges of health and demographic change addressed by the Innovation Union. ITFoM’s “virtual patient” concept will revolutionise the way medicine is conducted and will subsequently contribute to the long-term reduction of healthcare cost. The ICT component of ITFoM is totally in line with the Digital Agenda for Europe.

72
ITFoM generates a stream of innovation relying on a few key enabling technologies like micro and nanoelectronics, biotechnology, wireless ICT, secure use of health data bringing a new dynamics to a patient centric health market (diagnostics tools and new targeted drugs). Partnership with industry leaders, global companies as well as SMEs, guarantees a rapid scale up and a fast dissemination of new technologies and the establishment of new value chains thanks to open innovation.

Finally, the long-term approach offered by the European Flagship initiatives will allow to combine a 3 pillars bridge approaches to move from knowledge to market:

- strong coordination of European research clusters;
- pilot activities through user cases during the first years of the project
- increased ability to transfer knowledge to industry leaders (pharma, diagnostic tools providers and IT) for rapid market take up, therefore bridging over the ‘valley of death’.

**Fostering Innovation**

ITFoM with its research efforts along the entire value chain of modern health research will not only contribute to develop new methods and technologies, it will also help to quickly turn them into new innovative products and supply services. Its set-up will allow shortening “time to market” for innovation made in its scientific context and also shortening the time from “bench to bedside” on the clinical side.

**Setting Standards**

The enormous impact expected by ITFoM offers a unique opportunity to develop common standards in the field of IT for medicine. These standards will comprise procedures and quality level in wet laboratories and on the IT side, but also on the application of the results in clinical settings.

**Training**

The ITFoM initiative will put special emphasis on the training of young scientists across disciplines and specialisations, because new ideas tend to emerge at the juncture of specialisations. Through the training activities, scientists will be trained across faculties in biology, mathematics, computer science and physics. Scientists will be in permanent contact with a network of basic research, medical application, SME, pharmaceutical industry and economic growth.

### 4.3 Use of results and dissemination of knowledge

The results obtained within this planned project are of different nature. First of all, scientific data will be collected. These data are of very different nature, from “pure bulk data” e.g. from screening and sequencing approaches to highly delicate patient information. It is imperative to carefully consider the nature of data to choose the appropriate method of dissemination (or confidentiality).

The natural way to disseminate scientific data is to write scientific reports in peer reviewed journal and to give talks at scientific conferences. The impact of the work will make it feasible to continuously publish articles in high ranking journals and it should be possible to achieve a “branding” for ITFoM in this respect. Individual and more detailed publication strategies will be discussed at consortium meetings (see 2.1). Publications might be prepared either by the programme manager (see 2.1) or by a responsible partner of the consortium. Over the development of the initiative it will be possible to receive IPR on the results of ITFoM prior to a commercialisation.

In parallel it is intended to extensively report on ITFoM results on scientific events and meetings. In addition, the potential of modern ways of internet communication and dissemination e.g. Facebook etc. will be explored and applied where appropriate.
Beyond this, a dedicated dissemination strategy for ITFoM will be developed which aims to create visibility of the flagship initiative at its beginning, changing gradually to a reporting of results and scientific achievements and develops during the duration of the project towards a communication of general achievements and outcomes. In order to achieve a fully professional communication with the public it is considered to outsource certain activities.

Since ITFoM will use ICT technologies to transform the way medicine is practiced, drugs are discovered and developed; spreading excellence, exploiting results, and disseminating knowledge are critical elements vitally necessary to the success of this project.

Given that the project proposers comprise internationally recognised leaders in the biomolecular sciences, information and communication technologies, computational modelling and systems biology, data visualisation, and the clinical sciences as well as public health the primary focus of our dissemination strategy is to use the key members of each of our workpackages to communicate the strategy and vision of the initiative to the respective national and pan-European bodies, including outreach to international bodies.

Outreach and dissemination efforts will be structured for each ITFoM platform such that platform leaders in will be responsible for outreach and dissemination based on the focus and expertise of the workpackage and its team. Based on this model, the leaders for the medical platform will be responsible for communication and dialog with the appropriate medical societies as well as national (FDA, & others) and European regulatory agencies for medicine (EMA). Platform leaders in analytical techniques will be responsible for outreach to the appropriate standards committees such as NIST in the US and the JRC IRMM in Europe, research institutions, and national funding agencies. The computer Science and Statistics platform leaders will take the lead in outreach to national supercomputing organisations like IDRIS-CNRS in France, FZI and RZG in Germany, CINECA in Italy, EPCC and ECMWF in the UK, CSC in Finland, SARA in the Netherlands and NCSA and NERSC in the US as well as to industry trade organisations and motivated major IT centric corporations. The integration platform leaders will be responsible for working with standards committees for data formats such as SBML.org as well as with content providers like the European Bioinformatics Institute (EBI, one of the projects proposers) and the National Library of Medicine and the National Center for Bioinformatics (NCBI) in the US, as well as other national and European health agencies, leading modellers from research institutions and national funding agencies. The leaders of the coordination platform will be responsible for keeping overall track of dissemination efforts and ensuring that the projects overall goals are met.

Internal Dissemination

A critical element in the success of ITFoM will be the dissemination of information and technology among the members of the consortium and, subsequently, communication of the results to academia, industry and patient and medical groups. Bearing in mind that this project will generate exploitable data and methods a structure has been established to ensure that publication and distribution of data is appropriately regulated.

To facilitate communication and dissemination, ITFoM will hold an annual Scientific Meeting that will be used to share information, not only within the consortium, but also with invited academics, industry partners and representatives from regulatory bodies. It is also intended that, once potential IP issues have been resolved, findings will be presented at major scientific congresses relevant to ICT technologies, computational modelling, systems biology and medicine. All members of the consortium will be encouraged to submit individual abstracts or presentations. SOPs developed within the project will be published at the website and/or deposited in a methods database developed at UU – MolMeth (www.molmeth.org). In addition, however, the consortium will
encourage members to regularly present interesting findings to the interested scientific public and members of complementary international research groups. The focus for such presentations will be meetings in Europe or with significant European participation such as:

- The annual meetings of the ISC European Supercomputer conference
- The annual meetings of the European Grid Infrastructure (EGI) initiative, formerly known as EGEE
- European IEEE conferences focused on advances in Information Technology, Networking, and Storage Technologies
- EMBO and ISMB conference in systems biology and computational modelling
- Gordon and Keystone Research Conferences in appropriate disease areas
- The joint meetings of the European Organisation for Research and Treatment of Cancer (EORTC) with the National Cancer Institute (NCI), American Society of Clinical Oncology (ASCO) and the American Association for Cancer Research (AACR)
- Meetings of the German National Genome Research Network (NGFN)
- Meetings of the International Society of Oncology and Biomarkers (ISOBM)
- Clinical conferences held in the various medical areas such as Metabolic, Neurodegenerative, Infertility, Regenerative Medicine, Kidney, Liver, Heart, Lung and other relevant disease areas
- Meetings of European Policy Stakeholders such as European Health Forum Gastein (EHFG)

Engaging regulatory and governmental agencies

Whilst the outcome of our proposal, however, as outlined previously, promises a significant societal and individual health benefit, such an effect can only be realised if the regulatory authorities and governmental bodies are prepared to accept and recommend such testing procedures. For example, OncoType Dx has developed a 21 gene assay which for instance a Canadian patient with breast cancer must pay privately CAD3700 to discover if she can avoid chemotherapy - and in doing so save her government ca. CAD15000 for unnecessary treatment! Thus if ITFoM ultimately provides access to novel molecular methods for stratification and/or therapy monitoring in various diseases, the full benefit to the patient can only be realised if the relevant agencies are in a position to evaluate and sanction such procedures. Of equal importance, the research and development activities for novel diagnostics must be tailored to ensure that the formal requirements of the regulatory authorities can be met in a timely fashion. In Europe, incorporation of pharmacogenomic information in the regulatory process is under discussion. With this regulatory background in mind, the consortium will establish a dialogue with the regulatory authorities and healthcare providers so as to best pave the way for the ready implementation/uptake of novel molecular diagnostic and prognostic procedures. It is therefore our intention to ensure communication with EMA, for example by inviting participation of their IMI Liaison Officer in the SEAB and in scientific workshops.

Exploitation of results

The ITFoM project if it is to reach its goal of transforming medicine from stochastic/empirical to inference and knowledge driven will have generated a truly extraordinary amount of material for commercial exploitation. The industry IT partners in this proposal will obviously generate new data storage, encryption, compute architecture, visualisation methods, as well as new technologies, and systems which will each have a new ICT driven biomedical market. Industry partners like Siemens and
others will have the opportunity to build value added reseller systems based on a combination of IT innovations and methods developed by the computational and modelling partners in this proposal.

Furthermore, at the moment there are no models or tools that facilitate the implementation of health innovations such as individualised diagnostics or drugs into the healthcare systems in an effective, efficient and timely manner. Tools that are currently available either support the industry side (technology transfer principles) or the health policy side (health assessment tools), which means there is a lack of tools that combine both ‘worlds’ and form a bridge between industry and health policy-related stakeholders. However, recently one of the ITFoM partners developed the unique LAL (Learning-Adapting-Levelling) model which combines Public Health Assessment Tools (PHAT) with principles from Technology Transfer (TT). This model will be of considerable importance for ITFoM in order to bridge the gap between these stakeholders. In the end it will highly contribute to increase the likelihood of profit of the industry to introduce valuable innovations to the end-user, the patient.

These systems and tools will form the backbone of the Medicine of the Future and will, if we are successful, have an opportunity to transform the entire clinical marketplace. Additionally, these same systems, in configuration more amenable to primary research will form the backbone of the drug discovery and development systems of any pharmaceutical companies that are interested in investing in the new individualised Medicine of the Future.

Management of IPR

As ITFoM will revolutionise the modern healthcare systems the consortium expects to generate innovation in many different aspects and areas following the aim of the Innovation Union concept “to turn ideas into more jobs, improved lives and better society”.

In many European economies, health is the largest industry and is a growing industry. Europe thus has a unique opportunity to take a more prominent role in moving health research and innovation forward. Support for commercialisation to turn research from industry as well as technological and social innovations on the healthcare level, supported by systems and innovative distribution models, will greatly benefit Europe and beyond.

Objectives of the IP policy and technology transfer guidelines

The overall aim of this IP and technology transfer policy is to foster technology transfer from academia (universities and public research institutes) to industry (large groups, SMEs, start-up). The present IP and technology transfer policy does not have as a goal to supersede or replace the IP rules already developed around national and/or European initiatives for supporting collaborative R&D, but rather to develop new and complementary approaches in IP serving the above mentioned objectives with a focus on rights and obligations for all parties involved.

The IP and technology transfer guidelines, support organisation and support measures, that will serve the objectives of the ITFoM and support its long term strategy are detailed below.

IP and technology transfer guidelines

IP is considered by the ITFoM partners as a tool to raise European competitiveness in its relevant fields by:

- ensuring adapted protection (intellectual property rights and know how) for the various innovations developed within ITFoM and made during the entire R&D processes, from concept to exploitation by raising the level of professionalism in IP and contract management throughout the community;
• fostering the mobility of researchers while encouraging them to pursue the development of their innovations by following several steps in the R&D process while identifying and structuring IP portfolios in an "open innovation environment" benefiting the partners involved;

• disseminating knowledge for R&D purposes between ITFoM partners and towards all other potential contributors who could serve the objectives of the parties, through efficient access to IP, be they background or IPs covering new R&D results (including large information on the R&D results themselves after their protection and/or under confidentiality agreements);

• offering efficient access to IP portfolios on the themes covered by ITFoM to industrial companies, be they large groups, SMEs or start-ups, or agreed third parties, with fair and adapted remuneration schemes to the owners: this will ease the investment decisions requested for engineering and manufacturing development in order to finally result in new products and/or services adapted to development of new markets;

• supporting business models for technology transfer adapted to ITFoM new technologies and scientific breakthroughs, while bringing fair remuneration and reward schemes to IPs’ inventors and owners who have been part of the R&D process: IP commercialisation plans will be worked out by the parties from the beginning and adapted during all the R&D process to be available when the technology is ready for industrialisation.

The above IP guidelines will be implemented by an IP and technology transfer support organisation which will propose adapted IP processes to provide flexibility, allowing ITFoM to offer the most relevant IP framework to serve its strategy.

**IP for technology transfer support organisation and support measures**

Support organisation - constitution of high level working groups

ITFoM will provide an efficient way to interlink the IP structures of its partners, by constituting working groups composed of high level IP for technology transfer professionals within the community, which will take in charge the following missions:

• Identifying the different IP and technology transfer strategies to address the challenges of ITFoM

• Identifying on an on-going basis best practices in IP to design training sessions to raise the professionalism and level of understanding of IP and contract issues of the ITFoM partners;

• Working on potential "new" IP processes

• Implementing a survey and mapping IP portfolios within ITFoM on its key thematics

• Structuring IP portfolios, if relevant and on a voluntary basis, in view of fostering technology transfer

• Facilitating quick resolution of IP issues linked to R&D projects

• Acting as an escalation board (dispute resolution) when needed.

• Defining, collecting and reviewing key IP performance indicators

Based on the good experiences in large research programs e.g. in Germany (like DHGP, GABI) the consortium will implement an innovation officer dedicated to the IPR and technology transfer issues. Prior to any publication or patent filing the text must be submitted to the innovation officer to check and give advice

• for conflict of interest
• for potential (other) innovation not identified before
• for potential routes of exploitation

The committee provides feedback on potential IP and possible routes of exploitation.

Support measures

In addition, the following specific support measures fostering IP for technology transfer will be implemented:

Patent co-funding fund

Cost of patents is an important matter and could hinder ambitious worldwide patenting strategies for ITFoM partners lacking appropriate resources (such as some SMEs or universities). ITFoM will therefore fund a part of IP protection costs as initiation grants on operations such as extension of patents. Once revenues on IP exploitation arising, the benefactors of those grants will reimburse ITFoM according to their capabilities.

Business Intelligence fund

For new identified technologies and markets, it can be of the utmost importance to perform business intelligence studies, including IP intelligence studies, for example studies of freedom to operate and of prior art and with special focus on identifying and structuring IP portfolios. ITFoM will take in charge the costs of software tools and manpower to carry on those studies.

Equity fund (SME licensing context)

A well-known weakness in technology transfer from public research to SMEs is the financial difficulty that SMEs have to implement the outputs of technology transfer into products (industrialisation and marketing costs). ITFoM will provide appropriate support to SMEs through a dedicated equity fund.

4.4 Education and training at European level

Within the work plan ITFoM intends to have an entire workpackage dedicated to training activities. The interdisciplinarity of ITFoM offers unique opportunities to have cross discipline training activities firstly for young researchers, but also on a senior level. These interdisciplinary training efforts are key factors for a long term success of all methods and technologies to be developed under ITFoM. Within the workpackage “Training” a comprehensive set of tasks will facilitate a large variety of training efforts open to scientists from all ITFoM partners and, where appropriate, also for scientists from “non-partner organisations”. It is intended to have

i.) “physical” workshops on upcoming topics in the field,
ii.) web-based seminars for ad-hoc information of young scientists on relevant issues,
iii.) newsgroups on the intra-net for a continuous information and teaching
iv.) summer schools for dedicated education of the researchers in the initiative
v.) lab exchanges in order to learn new techniques and protocols
vi.) “cross faculty” exchanges in order to better understand the “ways of thinking” in other research disciplines.

Within the research program of ITFoM the consortium will also implement an ITFoM Academy with a specific education and training program dedicated to different target groups. The implementation of the education and training program will be stepwise and according to the development of the ITFoM research program. The ITFoM partner institutions will appoint a contact person for their institution
responsible for education and training issues. Among these contact persons the partners will choose one central responsible person for the consortium as a whole to co-ordinate all education and training activities within the Academy. Besides the organisation of courses and training programs the Academy will also set up and build a knowledge repository where course material, information, eLearning and hands-on material is collected and made available for the whole community.

**Internal training**

An “internal” training program will be set up as a first element for all members of ITFoM. The ITFoM program will require the close collaboration between different scientific disciplines that have not been integrated and interlinked before on such a scale. To ensure optimal interaction and communication between all members the consortium will implement an internal training program providing a basic overall knowledge across all disciplines involved in ITFoM. The initial training will be compulsory for all scientists of the ITFoM program. The topics of the courses will be chosen and agreed upon by the consortium members. Over the years the course program will be refined and updated according to the needs of the participants. As all participating members of the ITFoM consortium are experts in education and training and have implemented already successful training programs it will allow a fast start to build upon these existing structures. New additional training modules will be developed according to the needs of the consortium. In addition to seminars and hands-on courses the Academy structure will also organise internships. The participation of academic institutions and companies in the ITFoM consortium provides an excellent chance to use this partnership for training of young scientists both in academic and industrial and policy environment and thereby foster interaction and co-operation between these three areas. Via the internships and the more lively interactions between academia, industry and policy bodies we expect that “open innovation” processes will be leading to the generation of innovative products and services and the development of IP and timely implementation in healthcare systems.

**PhD program**

The ambitious ITFoM research program needs young researchers who have training to work at the interface between medical science, translational science and ICT or respectively lab technology development and ICT. To train specialists in these integrated scientific disciplines ITFoM will establish a joint PhD program that will allow specification of the graduate students in one of the interface areas. The PhD program will also be part of the ITFoM Academy. As the aim of the new PhD program is the special training at scientific interfaces with ICT all PhD students will have as at least two direct supervisors; one from the ICT field and one from the other two respective science areas. A third supervisor will be chosen in addition as a mentor for the career development; this third supervisor will also be responsible if conflict situations occur. This system will assure that the PhD students will receive an optimal supervision and support for all aspects of their PhD work. As also the international co-operation will be an important issue one of the supervisors should be from another participating institution apart from the hosting institute. Once a year a joint retreat will be organised for all PhD students to support community building and foster communication among the students. Anticipated start of the first PhD students will be in year 2 of the ITFoM program.

**Master programs**

In the second phase of the ITFoM program the consortium will engage also in Master programs and offer training elements for existing programs especially addressing the training at the interface between ICT and other science areas. The engagement in training and education will ensure the availability of young scientists trained to work in the highly ambitious and innovative science field of ITFoM. Anticipated start of activities towards education and training will be in year 4 of the program.
Activities for user community

An additional target group for education and training activities will be the potential users of the ITFoM patient model. Based on a broad-scale dissemination and information initiative that will start already in year 1 the consortium of ITFoM will offer seminars and workshops in the EU Member States to inform interested representatives and stakeholders about the new approach and the system. We will offer courses, talks and seminars for initiatives and existing programs addressing practitioners, nurses, and hospital staff as well as other healthcare providers from the policy field (e.g. HTA agencies). The activities for the user community – giving the possibility to also integrate stakeholders – are very valuable in two aspects: The user community has open access to information early on and can follow the development over time. On the other hand ITFoM benefits from the user perspective and integrates the feedback information to optimise the system and make it user friendly. This is extremely important for the creation of a user interface that should be self-explanatory and easy to handle.

Scholarships and specific grant system

In case that some European regions are not participating with expert groups in the ITFoM program the consortium will provide a special support system targeted especially to these countries to foster the integration and participation of these countries in the Academy activities, especially the PhD and Master programs. ITFoM will reserve places for study courses and offer specific support and scholarships to students from these European regions.

4.5 Potential ethical and legal implications

The aim of the ITFoM project is nothing less than to contribute to a process of ‘revolutionising our healthcare system’ by creating a breakthrough in individualised medicine. Thus, the explicit goal of ITFoM is a radical and fundamental alteration of the existing structures of the biomedical-scientific complex and of the healthcare system. Such fundamental changes mean that both the material conditions and the social norms that govern the conduct of a domain are altered in ways that will constitute deep challenges for the existing ways of understanding and operating medical research and healthcare with extensive implications for researchers, scientists, medical personnel, patients, the pharmaceutical industry and society at large. Typically, revolutions are carried out outside the existing legal forms, outside of the traditional material basis, and outside the existing cultural normative system. That’s why revolutions are rare. But if they happen they have deep impact and their central pre-condition and condition sine qua non is a good deal of ‘outside the box thinking’ as traditional concepts and methodologies limit the potentials for apprehending the transformations that are occurring.

To develop a new, data-rich, individualised medicine, information availability is essential to develop the necessary knowledge. Traditionally, significant information availability clashes with the right of privacy and data protection. The ITFoM project presents the chance to apply some ‘Janusian thinking’ to the development of the relevant regulatory framework. Janus, the Greek god of doors and gates and beginning and endings, was most often depicted as a man with two heads, each facing in opposite directions. The benefit of such dual perspective (and the underlying power of Janusian thinking) is that it provides the ability to consider multiple perspectives simultaneously. Instead of selecting between the opposites it allows exploring surprising ways of uniting opposing theories (AND instead of OR). From the legal point of view, consistently with the project’s visionary approach, we will apply some Janusian thinking and look for solutions that combine availability of information, on the one side, AND patient rights, privacy, data protection, and data security, on the other side. Nevertheless, in carrying out such revolution ITFoM will consider the existing applicable regulatory
framework and suggest changes to it only where necessary and after having carefully considered and balanced the interests of all the parties involved.

ITFoM, thus, will have deep impact in a number of areas that will have to rethink well-established concepts, ways of doing, and modes of interaction, and this applies no less in the social-ethical-legal fields. There are several challenges to be addressed.

**The challenge of the ICT ethics/biomedical ethics interface**

To date biomedical ethics and ICT ethics have developed separately. The ethics of biomedicine has expanded, from an initial focus on the clinical encounter and the health-professional relationship, to take in issues that are beyond the clinic, involving the impact of scientific and technological developments, in the postgenome era, for example, and a concern for public health. ICT Ethics is a younger field of specialism, in ethical terms, than the ethics of biomedicine, and it is already producing challenging issues to think about, e.g., in the field of ‘smart’ implants: how they make us conceptualise enhancement and what it means to be human. This point makes it clear that there are deep philosophical questions to be addressed, in addition to ethical framings. Some of the emerging issues have been discussed in the 2011 EU publication *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*, but there is a need for on-going monitoring of developments.

**The Big Data Challenge**

One of the key specific ethical-regulatory challenges of ITFoM at the ICT/biomedical interface will come from the fact that large amounts of data related to healthy individuals and patients will be collected and stored. This brings up a number of ethical-legal issues. Contemporary and future biomedicine projects such as ITFoM, however, are increasingly concerned with collection, storage, interpretation and use, of ever larger quantities of data, and it is here that there is a clear interaction with ICT ethics. Key ethical concepts invoked include privacy, and rights to know and not to know, in addition to issues of autonomy, justice and equity. It has become clear from the genomics context, however, that ethical concepts co-evolve with science. It is at the interface of ICT and biomedicine that we can expect further co-evolution, in the light of the potential implications for patients, professionals and society of ‘big data’ on the one hand, and the increasing ‘personalisation’ that the data make possible, on the other. There is both a broadening and narrowing of focus. Also, important legal issues come up here due to massive data communication and transfer to subjects established inside and outside the European Economic Area. Moreover, the data that will be processed may fall under the definition of sensitive data’ according to the present data protection regulation or, more precisely, ‘data concerning health’, ‘biometric data’ and potentially ‘genetic data’ according to the EU Commission Proposal for a General Data Protection Regulation. More stringent rules apply to the processing of such categories of data both form the data protection and from the data security point of view. An evolution of strongly ICT-based solutions to ensure privacy, data protection, data security and patient’s rights are desirable. Identifying them represent a stimulating challenge for the project. Moreover, other legal implications of the ITFoM model for access to social services, such as healthcare and life insurance will have to be carefully dealt with.

**The challenge of ‘personalisation’ What is “personalised medicine”? For whom? In which healthcare system?**

A further challenge comes not only from the social and ethical implications of personalisation, but from how it is conceptualised. The promise of ITFoM is the potential for a genuinely personalised medicine in the sense of providing for the first time a mechanism for accessing the level of detailed data about an individual that could make tailored medicine a reality. While some may see the provision of such personal information as empowering, some may be concerned, perhaps because a version of the ‘right not to know’ argument might suggest that the volume of information could be
overwhelming, if not appropriately handled. For the growing number of Europeans subscribing to complementary medicine (with a different understanding of personalisation) and its various treatment options even the very nature and value of the medical data produced will be questionable. In other words, the very conceptualisations in, as well as the implementation of, ITFoM, are a subject for ethical and social debate.

While the arguments for personalised medicine are very strong, in light of the challenges currently faced by society, another concern expressed about the turn towards personalised health advice based on genomic information, for example, has been that it represents a ‘boutique’ model of healthcare appropriate only for societies demonstrating an individualistic culture – and with sufficient resources to afford it. Issues of equity of access therefore need to be addressed. There is also the question to be considered as to what extent European societies will subscribe to the new world of personalised medicine and its routines and will be willing to cooperate. These considerations go to the heart of the objectives of personalisation rather than its implementation.

Furthermore, providing individuals with information may be promoted in the name of informed choice, but there is a question about the implications of a presupposition that there is a ‘right’ choice, leading to being held responsible for wrong choices, in particular. Personalisation may also be understood in terms of responsibilisation: thus how it is being conceptualised needs on-going monitoring. From the data protection point of view, there are a number of legal questions that will need to be carefully addressed, especially related to data minimisation and to the principle of proportionality (how to assure that only relevant data will be processed, avoiding excessive and unlawful data processing?), data quality (how to assure that data are correct and constantly kept up to date?), and, last but not least, data security (how to effectively protect such amount of ‘sensitive data’ against risk of loss or unauthorised access, destruction, use, modification or disclosure?)

The challenge for professional practice

The ethical question about the extent to which the volume of information will be meaningful, particularly in the early stages of development, arises not only for future patients but also for health professionals who will be delivering care in the future. How might their professional ethics and practice be affected by the requirements to inform? What level of detail would count as adequate information for the purposes of informed consent? Are there also grounds for thinking that the privacy concerns are increased in light of the increased volume of data and what it might reveal about an individual? What would need to be put in place in terms of training of professionals and safeguards against error and misuse? ITFoM will thus have a substantial impact on the daily performance and practice of medical professionals in the healthcare system: What will it take to achieve compliance and cooperation? What are the potentials for public resistance and lack of willingness to accept the new world of personalised medicine? In the legal fields important issues come up here such as on the physician-patient relationship, clinical trials, off-label prescriptions and the standard of care and associated liabilities

The challenge of ethical change

Ethics has to deal with differences not only in place, in the global context, but also in time. For example, in the early nineties the issues surrounding disclosure and rights to know and not to know tended to fall into two main types: conflicts of interest in the clinical setting, where one person’s reluctance to disclose information might have implications for another’s right to know it; and situations where an individual had to deal with predictive information, which they would prefer not to have knowledge of. So discussion turned on the circumstances in which it could be right to disclose; and choice in relation to tests. The ethical considerations at stake included, primarily, individual autonomy, confidentiality, and beneficence. Since that time, with an increasing emphasis on population research and public health, there has been a move towards solidarity and equity as
important principles to consider. Organisations such as the Human Genome Organisation and the Nuffield Council on Bioethics have addressed the greater prominence of solidarity in society and in healthcare, and the implications of this move away from a focus of ethics on the individual (Prainsack and Buyx, 2011).

A good example of the change in ethical thinking is the trend towards rethinking the concept of privacy. Even if we could rely on public institutions to adhere completely to regulations on data protection, changing as they are, there is still a need to look at the issues in different ways. In the genomics context, there has been a suggestion that emphasis on privacy should be replaced by the concept of open consent. The concept of open consent is used in the context of the Personal Genome Project, which aims to build a framework for the development and evaluation of personal genome technologies for public good. The advent of ‘celebrity genomes’ where key figures willingly publish their genomes to illustrate a different attitude towards privacy and disclosure has caused some concern (Check, 2007). We also can expect the emergence of IT tools (Apps), which might “enhance” ELSI rights of research and data subjects. These applications are being designed to serve a dual purpose. They enable participants to concretely exercise/enforce their ethical-legal rights. But they also allow for self-reporting by participants, thus contributing substantially to projects like ITFoM. In this respect, it will be worth exploring and developing further the principle of ‘privacy by choice’; in combination with the principles of ‘privacy by design’, ‘privacy by default’, and ‘data portability’ recently codified in the EU Commission Proposal for a General Data Protection Regulation).

Central themes to consider

So how is the debate, e.g. on the right to know and the right not to know, likely to develop? The pace of development of the technologies is very fast: ethical thinking is faced with the move towards greater ‘personalisation’, on the one hand, and the requirements to think about the context, including the economic and global context, on the other. The discussion of individual rights to know and not to know has to be continually renegotiated in the light of interests in tension. There is potential, however, for added value and further theoretical development from the interaction of ethical traditions from the communities of ICT and biomedicine. The public acceptance of ITFoM, concerns and anxieties connected with large medical data collection and transformations in the operation of the healthcare system remain issues to be aware of. Ethics is sometimes understood as a mechanism of control, and indeed its role in governance (if not control as such) is an important one. However, ethics as a multidisciplinary field of study has an equally important role in its contribution to monitoring how emerging issues are conceptualised in public, academic and policy arenas. From a legal perspective, ITFoM raises a range of new questions from the physician-patient relationship to clinical trials, standard of cares and associated liabilities and reimbursement. The ITFoM approach might also have considerable implications for access to social services, such as healthcare and life insurance. The WP 1.5 platform will deal with these different societal, legal, and ethical challenges by providing the ethical and legal groundwork to deal with the ethical and legal governance, developing a broad repertoire of methods with the goal to create governance challenges embedded in ITFoM and the social mechanisms to work on the IT-society interface.

Ethical matrix

Ethical reflection will be conducted in collaboration with social science. This will be an evidence-based approach to ethics. Below is an example of an ethical tool, the ethical matrix developed by Ben Mepham. The figure below, which is illustrative only, shows a practical approach to ethical principles being considered in relation to different stakeholder groups, providing an opening for their input. The boxes in this illustration give examples of considerations relating to the different groups with reference to the principle in question. Mepham’s version however, has been criticised for what is left out, and competes with other approaches; thus it needs to be supplemented with reference to
solidarity and other emerging principles. The list of stakeholders will also be expanded to include other groups such as the media. The ethical work will therefore develop a project-specific matrix in the light of philosophical, legal and social science work. It is to be expected that through the course of ITFoM other considerations will emerge, as they have done through the experience of the ELSI-genomic era. Thus ethics cannot be a one-time tick-box activity but is dynamic and has to be conducted in collaboration with social science.

Table 1 Social Science/Public Engagement Methods Matrix

<table>
<thead>
<tr>
<th></th>
<th>Autonomy</th>
<th>Well-being</th>
<th>Justice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientists</td>
<td>Freedom to research</td>
<td>Facilities and funding</td>
<td>Appropriate reward e.g. IP</td>
</tr>
<tr>
<td>Patients</td>
<td>Right to know or not to know</td>
<td>Improved treatment options</td>
<td>Access to resources</td>
</tr>
<tr>
<td>Vulnerable groups</td>
<td>Right to be heard</td>
<td>Alleviation of disadvantage</td>
<td>Equality</td>
</tr>
<tr>
<td>Professional groups</td>
<td>Professional judgment</td>
<td>Increased burden?</td>
<td>Implications for practice</td>
</tr>
</tbody>
</table>

Table 2 Focus Groups and Surveys

<table>
<thead>
<tr>
<th></th>
<th>Focus Groups Surveys</th>
<th>Scenario Workshops</th>
<th>Interviews</th>
<th>Public Laboratories</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Public</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Patients</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Medical Personnel</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Industry</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societal Experts</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on this social science research, we will establish a link between science and society, namely by setting up focus groups and by conducting a number of flash Eurobarometer surveys in Europe. We will use this mixed-methods design to develop a precise and valid understanding of the ethical and societal issues of ITFoM as they emerge in public perception. We will conduct the surveys for the EU 27, and the focus group in selected countries that represent a wide range of the European opinion spectrum when it comes to issues of Information Technology, medicine and society in different European countries.

Interview Work

The approaches and solutions of ITFoM need to be tightly integrated with a broad variety of stakeholders, not only the general public, but also medical doctors, insurance companies, industry, and policy-makers. It is planned to develop a more precise understanding of the perspectives of key stakeholders by conducting systematically in-depth interviews with representatives of organisations such as pharmaceutical companies and policy-makers who might not be interested or willing to engage in other forms of dealing with issues of ITFoM, such as scenario workshops and public laboratories. These interviews should be conducted in different EU countries and we anticipate the need for 60 in-depth qualitative interviews.

For both focus groups and interviews, in-depth interpretive analysis will be performed with the aim of revealing structures of perception and discourse to deepen our knowledge and interpretations beyond the manifest content Scenario Workshops.
It is the aim of ITFoM to “harness the vast potential of ICT to revolutionise human healthcare” to lead “the way towards truly personalised healthcare” (ITFoM 2011). Hence, the project’s vision addresses many important disputed aspects of healthcare – from the use of information technology in healthcare to unjust access to these new insights and developments. The intention of ITFoM is to translate the innovations addressed by the project timely into health policies and evidence-based health practice. This raises the question of how to communicate the project’s promises at all and in particular timely to society in general, patients, physicians, insurance companies, politicians etc.?

A method to ameliorate interaction between research and science with society and politics is to apply scenario workshops in the context of ITFoM – which, since 1994, are utilised as so-called ‘European Awareness Scenario Workshops’ (EASW) by the European Commission to “explore possible new actions and social experiments for the promotion of environment favouring innovation in Europe” (CORDIS 1997): “The Scenario Workshop is a development of the ‘Future Workshop’ and basically it follows the same three phases for criticism, vision, and fantasy. However, the Scenario Workshop is based on a presentation of possible future developments in the area. These so-called Scenarios have been formulated in advance. The criticism of the Scenarios by the participants together with their own experiences form the basis for visions and action plans” (CIPAST n.y.). Scenario workshops are utilised to discuss and – in the best case – to solve problems by connecting research and socio-political needs (cf. people and participation.net 2011). Scenario workshops combine different forms of participation, facilitating knowledge transfer and exchange between science, research and lived experience (cf. people and participation.net 2011). In doing so, scenario workshops strengthen not only decision-making processes via participation and interaction, but also initiate and facilitate discussions about challenges which might impede the ITFoM project’s path from vision to reality: The ITFoM project’s aim to “change the future of medicine” (ITFoM n.y.) can be best addressed by constantly exchanging knowledge which is helpful and important to connect current experiences (positive and negative, such as fears, reservations and hopes) with the fundamental changes in medicine and healthcare proposed by ITFoM.

Public Laboratories

How different stakeholders will react to attempts to assemble large, networked science projects is hard to predict. If to unleash the powers of science is what is promised to citizens, it will be important to create spaces in which it is possible to reflect on the possible consequences in an inclusive and sometimes experimental fashion, in order to grasp what envisioned changes may mean to different constituencies. Attempts to demonstrate how scientific projects are assembled, what they do in order to produce facts and useful devices and in how far they address general concerns that matter outside seminar rooms and laboratories, are indispensable in this respect.

Public participation, public dialogue and public deliberation have become catchwords in governance research and especially in relation to questions of S&T governance. But the science and life science governance discourse often takes on an ‘object avoiding’ tendency. Our societies are increasingly built around separations between specialists and ‘lay persons’. ‘Ordinary’ people have delegated the task of producing knowledge and machines to engineers and scientists, who tell them what it actually is that they need. But these specialists are habituating a world that appears to be very distant since these (hyper) modern tribe is in a way locked up in artificial eco-systems called ‘laboratory’.

In order to overcome these deficiencies this WP will curate an interactive ITFoM exhibition that will put the things, the technologies, the practices and the knowledge that constitute the Consortium center stage. It is going to invite all already identified stakeholders and other interested groups and organised interests to make their own experiences with the practices and technologies that are to be connected by this consortium. Together with the other WP we will expose the technical artefacts, the knowledge trajectories and the assertions that are assembled by the project and create a hybrid forum. What is at stake here isn’t a large scale PR event, but an innovative political laboratory and an
experimental exercise in technological democracy that is supposed to trigger a series of lively debates in different formats, which potentially reveal new and surprising perspectives and concerns relating to the project in different ways and possibly uttered by unanticipated stakeholders.

Since publics are not simply ‘out there’ waiting for someone to ask them questions or urge them to present their opinions the exhibition will in a sense be an experiment in making ITFoM public without deciding who is to be concerned or what has to be considered beforehand. In reality publics are often assembled by common matters, concerns or issues and emerge around common problems. That is why this exhibition will be a place to assemble publics and induce public debates around the project’s aims, visions, assumptions, and - equally important - its technological components and the knowledge bases on which they rest. It will be a platform for discussing and identifying different concerns, practical problems but also hopes and anxieties with a large variety of known and emerging stakeholders. Last but not least it will also be governance tool for integrating the multidisciplinary consortium partners by offering an arena in which common issues can be presented to each other in innovative and interactive settings.
Supplementary Material (in annex)

Annex II: Overview participants and associated partners
Annex III: Glossary
Annex IV: Multimedia Material
Flagship Final Report
IT Future of Medicine

Annex II
Overview participants and associate partners
<table>
<thead>
<tr>
<th>Participant No</th>
<th>Participant organisation name</th>
<th>Part. short name</th>
<th>Country</th>
<th>Sector</th>
<th>Principal Investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Max Planck Institute for Molecular Genetics</td>
<td>MPG</td>
<td>DE</td>
<td>A</td>
<td>Hans Lehrach</td>
</tr>
<tr>
<td>2</td>
<td>Medical University of Graz</td>
<td>MUG</td>
<td>AT</td>
<td>A</td>
<td>Kurt Zatloukal</td>
</tr>
<tr>
<td>3</td>
<td>University College London</td>
<td>UCL</td>
<td>UK</td>
<td>A</td>
<td>Mark Girolami, Stephan Beck</td>
</tr>
<tr>
<td>4</td>
<td>University Amsterdam</td>
<td>VUA</td>
<td>NL</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>University of Manchester</td>
<td>UNIMAN</td>
<td>UK</td>
<td>A</td>
<td>Hans Westerhoff</td>
</tr>
<tr>
<td>6</td>
<td>European Molecular Biology Laboratory - European Bioinformatics Institute</td>
<td>EMBL-EBI</td>
<td>UK</td>
<td>A</td>
<td>Janet Thornton</td>
</tr>
<tr>
<td>7</td>
<td>Wellcome Trust Sanger Institute</td>
<td>Sanger</td>
<td>UK</td>
<td>A</td>
<td>Tim Hubbard</td>
</tr>
<tr>
<td>8</td>
<td>Kungliga Tekniska högskolan</td>
<td>KTH</td>
<td>SE</td>
<td>A</td>
<td>Mathias Uhlen</td>
</tr>
<tr>
<td>9</td>
<td>Imperial College London</td>
<td>ICL</td>
<td>UK</td>
<td>A</td>
<td>Jeremy Nicholson</td>
</tr>
<tr>
<td>10a</td>
<td>Maastricht University</td>
<td>UM</td>
<td>NL</td>
<td>A</td>
<td>Jos Kleinjans</td>
</tr>
<tr>
<td>10b</td>
<td>Maastricht University - Institute for Public Health Genomics</td>
<td>UM-IPHG</td>
<td>NL</td>
<td>A</td>
<td>Angela Brand</td>
</tr>
<tr>
<td>11</td>
<td>Consorzio Interuniversitario Risonanze Magnetiche di Metalloproteine Paramagnetiche</td>
<td>CIRMMP</td>
<td>IT</td>
<td>A</td>
<td>Ivano Bertini</td>
</tr>
<tr>
<td>12</td>
<td>International Prevention Research Institute</td>
<td>IPRI</td>
<td>FR</td>
<td>I</td>
<td>Peter Boyle, Markus Pasterk</td>
</tr>
<tr>
<td>13</td>
<td>Uppsala University</td>
<td>UU</td>
<td>SE</td>
<td>A</td>
<td>Ulf Landegren</td>
</tr>
<tr>
<td>14</td>
<td>University of Luxembourg - Luxembourg Centre for Systems Biomedicine</td>
<td>LCSB</td>
<td>LU</td>
<td>A</td>
<td>Rudi Balling</td>
</tr>
<tr>
<td>15</td>
<td>University of Leicester</td>
<td>ULEIC</td>
<td>UK</td>
<td>A</td>
<td>Anthony Brookes</td>
</tr>
<tr>
<td>16</td>
<td>ISC Intelligence in science</td>
<td>ISC</td>
<td>BE</td>
<td>A</td>
<td>Declan Kirrane</td>
</tr>
<tr>
<td>17</td>
<td>University of Auckland</td>
<td>UOA</td>
<td>NZ</td>
<td>A</td>
<td>Peter Hunter</td>
</tr>
<tr>
<td>18</td>
<td>Universite de Geneve</td>
<td>UNIGE</td>
<td>CH</td>
<td>A</td>
<td>Denis Hochstrasser</td>
</tr>
<tr>
<td>19</td>
<td>Centro Nacional de Análisis Genómico</td>
<td>CNAG</td>
<td>ES</td>
<td>A</td>
<td>Ivo Gut</td>
</tr>
<tr>
<td>20</td>
<td>Siemens Aktiengesellschaft Oesterreich</td>
<td>SAGOE</td>
<td>AT</td>
<td>I</td>
<td>Paul Pavetich, Achim Plum</td>
</tr>
<tr>
<td>21</td>
<td>Alacris Theranostics GmbH</td>
<td>ATG</td>
<td>DE</td>
<td>I</td>
<td>Bodo Lange</td>
</tr>
<tr>
<td>22</td>
<td>Charite Universitätsmedizin Berlin - Charité Comprehensive Cancer Center</td>
<td>CCCC</td>
<td>DE</td>
<td>A</td>
<td>Ulrich Keilholz</td>
</tr>
<tr>
<td>23</td>
<td>Illumina</td>
<td>Illumina</td>
<td>UK</td>
<td>I</td>
<td>David Bentley</td>
</tr>
<tr>
<td>24</td>
<td>Commissariat a l'énergie atomique et aux energies alternatives</td>
<td>CEA</td>
<td>FR</td>
<td>A</td>
<td>Nora Benhabiles, Jean-Pierre Meyer</td>
</tr>
</tbody>
</table>

Associate partners (Letters of support downloadable at https://ws.molgen.mpg.de/ws/505322/LoS_all.pdf)
<table>
<thead>
<tr>
<th>Participant No</th>
<th>Participant organisation name</th>
<th>Part. short name</th>
<th>Country</th>
<th>Sector</th>
<th>Principal Investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>Ontario Institute for Cancer Research</td>
<td>OICR</td>
<td>CA</td>
<td>A</td>
<td>Tom Hudson</td>
</tr>
<tr>
<td>103</td>
<td>International Foundation for Computer Assisted Radiology and Surgery</td>
<td>IFCARS</td>
<td>DE</td>
<td>A</td>
<td>Heinz Lemke</td>
</tr>
<tr>
<td>104</td>
<td>German Cancer Research Center</td>
<td>DKFZ</td>
<td>DE</td>
<td>A</td>
<td>Stefan Wiemann</td>
</tr>
<tr>
<td>105</td>
<td>Microsoft Cooperation</td>
<td>COSBI</td>
<td>IT</td>
<td>I</td>
<td>Corrado Priami</td>
</tr>
<tr>
<td>106</td>
<td>IBM Research</td>
<td>IBM</td>
<td></td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>106a</td>
<td>IBM Research Zurich</td>
<td>IBM Zurich</td>
<td>CH</td>
<td>I</td>
<td>Michal Rosen-Zvi</td>
</tr>
<tr>
<td>106b</td>
<td>IBM Research Haifa</td>
<td>IBM Haifa</td>
<td>IL</td>
<td>I</td>
<td>Ban Kawas</td>
</tr>
<tr>
<td>107</td>
<td>Roche Diagnostics</td>
<td>Roche</td>
<td>DE</td>
<td>I</td>
<td>Paul Schaffer</td>
</tr>
<tr>
<td>108</td>
<td>Agilent Technology</td>
<td>Agilent</td>
<td>UK</td>
<td>I</td>
<td>Nigel Skinner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Maurice Sanciaume</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zohar Yakhini</td>
</tr>
<tr>
<td>109</td>
<td>Life Technologies/Visigene</td>
<td>LTEK</td>
<td></td>
<td>I</td>
<td>Kai Seehaus</td>
</tr>
<tr>
<td>110</td>
<td>Bruker Daltonics</td>
<td>BDAL</td>
<td>DE</td>
<td>I</td>
<td>Arnd Ingendoh</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Detlev Suckau</td>
</tr>
<tr>
<td>111</td>
<td>AB SCIEX</td>
<td>ABSCIEX</td>
<td>DE</td>
<td>I</td>
<td>Volker Kruft</td>
</tr>
<tr>
<td>112</td>
<td>Intel</td>
<td>Intel</td>
<td>DE</td>
<td>I</td>
<td>Karl Solchenbach</td>
</tr>
<tr>
<td>113</td>
<td>Oxford Nanopore Technologies</td>
<td>ONT</td>
<td>UK</td>
<td>I</td>
<td>Gordon Shanghera</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spike Willcocks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zoe McDougall</td>
</tr>
<tr>
<td>114</td>
<td>Bayer Healthcare</td>
<td>BAY</td>
<td>DE</td>
<td>I</td>
<td>David Henderson</td>
</tr>
<tr>
<td>115</td>
<td>Janssen Pharmaceutica</td>
<td>Janssen</td>
<td>BE</td>
<td>I</td>
<td>Hans Winkler</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Carola Hawinkel</td>
</tr>
<tr>
<td>116</td>
<td>Institute for the Digital Economy</td>
<td>IMPACT</td>
<td>CA</td>
<td>A</td>
<td>Mihaela Ulieriu</td>
</tr>
<tr>
<td>117</td>
<td>Harvard Medical School</td>
<td>HMS</td>
<td>US</td>
<td>A</td>
<td>George Church</td>
</tr>
<tr>
<td>118</td>
<td>Dahlem Centre for Genome Research and Medical Systems Biology</td>
<td>DCGMSB</td>
<td>DE</td>
<td>A</td>
<td>Hans Lehrach</td>
</tr>
<tr>
<td>119</td>
<td>Tel Aviv University</td>
<td>TAU</td>
<td>IL</td>
<td>A</td>
<td>Ron Shamir</td>
</tr>
<tr>
<td>120</td>
<td>XEROX Research Centre Europe</td>
<td>Xerox</td>
<td>FR</td>
<td>I</td>
<td>Cedric Archambeau</td>
</tr>
<tr>
<td>121</td>
<td>Oracle Corporation</td>
<td>ORACLE</td>
<td>DE</td>
<td>I</td>
<td>Till Brügelmann</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wilfried Stüttgen</td>
</tr>
<tr>
<td>122</td>
<td>University of Genoa - Department of Biophysical and Electronic Engineering</td>
<td>DIBE</td>
<td>IT</td>
<td>A</td>
<td>Davide Anguita</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Alessandro Ghio</td>
</tr>
<tr>
<td>123a</td>
<td>University of Tartu - Estonian Genome Center</td>
<td>UT (EGCUT)</td>
<td>EE</td>
<td>A</td>
<td>Prof. Andres Metspalu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Erkki Leego</td>
</tr>
<tr>
<td>123b</td>
<td>University of Tartu - The Bioinformatics, Algorithmics, and Data Mining group</td>
<td>UT (BIIT)</td>
<td>EE</td>
<td>A</td>
<td>Jaak Vilo</td>
</tr>
<tr>
<td>124</td>
<td>Healthsolve</td>
<td>Healthsolve</td>
<td>UK</td>
<td>I</td>
<td>Alan Payne</td>
</tr>
<tr>
<td>125</td>
<td>Institut Pasteur</td>
<td>IP</td>
<td>FR</td>
<td>A</td>
<td>Matthew Albert</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Darragh Duffy</td>
</tr>
<tr>
<td>126</td>
<td>MPI of Molecular Plant</td>
<td>MPIMPP</td>
<td>DE</td>
<td>A</td>
<td>Lothar Willmitzter</td>
</tr>
<tr>
<td>Participant No</td>
<td>Participant organisation name</td>
<td>Part. short name</td>
<td>Country</td>
<td>Sector (A=Academia, I=Industry, O=Other)</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------------------------------------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>127</td>
<td>Physiology Vienna University of Technology</td>
<td>TUVienna</td>
<td>AT</td>
<td>A</td>
<td>Schahram Dustdar</td>
</tr>
<tr>
<td>128</td>
<td>VSB -Technical University Ostrava</td>
<td>VSB</td>
<td>CZ</td>
<td>A</td>
<td>Ivo Vondrak</td>
</tr>
<tr>
<td>129</td>
<td>Westfälische Wilhelms-Universität Münster</td>
<td>WWUM</td>
<td>DE</td>
<td>A</td>
<td>Martin Dugas</td>
</tr>
<tr>
<td>130</td>
<td>Metanomics</td>
<td>MTNX</td>
<td>DE</td>
<td>A</td>
<td>Patricia Ruiz-Noppinger</td>
</tr>
<tr>
<td>131</td>
<td>CertiCon a.s</td>
<td>CertiCon</td>
<td>CZ</td>
<td>I</td>
<td>Karel Galuska</td>
</tr>
<tr>
<td>132</td>
<td>Czech Technical University Prague</td>
<td>CVUT</td>
<td>CZ</td>
<td>A</td>
<td>Vladimír Marík Lenka Lhotská</td>
</tr>
<tr>
<td>133</td>
<td>Dublin City University</td>
<td>DCU</td>
<td>IR</td>
<td>A</td>
<td>Martin Crane</td>
</tr>
<tr>
<td>134</td>
<td>Helsinki Institute for Information Technology/Aalto University</td>
<td>HIIT</td>
<td>FI</td>
<td>A</td>
<td>Samuel Kaski</td>
</tr>
<tr>
<td>135</td>
<td>Athens University of Economics and Business</td>
<td>AUEB</td>
<td>GR</td>
<td>A</td>
<td>Petros Dellaportas</td>
</tr>
<tr>
<td>136</td>
<td>SIVECO Romania</td>
<td>Siveco</td>
<td>RO</td>
<td></td>
<td>Traian Ionescu Monica Florea</td>
</tr>
<tr>
<td>137</td>
<td>Universität Klagenfurt</td>
<td>AAU</td>
<td>AT</td>
<td>A</td>
<td>Johann Eder</td>
</tr>
<tr>
<td>138</td>
<td>University of Strathclyde</td>
<td>UNISTR</td>
<td>UK</td>
<td>A</td>
<td>David Littlejohn</td>
</tr>
<tr>
<td>139</td>
<td>UCD Systems Biology Ireland</td>
<td>UCD-SBI</td>
<td>IR</td>
<td>A</td>
<td>Walter Kolch</td>
</tr>
<tr>
<td>140</td>
<td>University of Copenhagen</td>
<td>UNICO</td>
<td>DK</td>
<td>A</td>
<td>Anders Krogh</td>
</tr>
<tr>
<td>141</td>
<td>University of Edinburgh</td>
<td>UEDIN</td>
<td>UK</td>
<td>A</td>
<td>Jane Hillston</td>
</tr>
<tr>
<td>142</td>
<td>University of Waterloo-Institute for Quantum Computing</td>
<td>UW-IQC</td>
<td>CA</td>
<td>A</td>
<td>David Cory</td>
</tr>
<tr>
<td>143</td>
<td>University of Rijeka</td>
<td>UNIRI</td>
<td>CR</td>
<td>A</td>
<td>Jasna Peter-Katalinic</td>
</tr>
<tr>
<td>144</td>
<td>IST</td>
<td>IST</td>
<td>IT</td>
<td></td>
<td>Ulrich Pfeffer</td>
</tr>
<tr>
<td>145</td>
<td>TNO</td>
<td>TNO</td>
<td>NL</td>
<td></td>
<td>Ben van Ommen</td>
</tr>
<tr>
<td>146</td>
<td>KU Leuven</td>
<td>KUL</td>
<td>BE</td>
<td>A</td>
<td>Svetla Nikova</td>
</tr>
<tr>
<td>147</td>
<td>Donders Center for Neuroscience</td>
<td>DCN</td>
<td>NL</td>
<td>A</td>
<td>Bert Kappen</td>
</tr>
<tr>
<td>148</td>
<td>Genome Sciences Centre, BC Cancer Agency.</td>
<td>BCGSC</td>
<td>CA</td>
<td>A</td>
<td>Steven Jones</td>
</tr>
<tr>
<td>149</td>
<td>Jozef Stefan Institute</td>
<td>JSI</td>
<td>SI</td>
<td></td>
<td>Dunja Mladenic</td>
</tr>
<tr>
<td>150</td>
<td>Fraunhofer Institute for Biomedical Engineering</td>
<td>IBMT</td>
<td>DE</td>
<td>A</td>
<td>Frank Bier</td>
</tr>
<tr>
<td>151</td>
<td>FUJITSU LABORATORIES OF EUROPE LIMITED</td>
<td>FLE</td>
<td>UK</td>
<td>I</td>
<td>Ross Nobes Lester Russell James Southern</td>
</tr>
<tr>
<td>152</td>
<td>Karolinska Institutet</td>
<td>KI</td>
<td>SE</td>
<td>A</td>
<td>Jan-Eric Litton</td>
</tr>
<tr>
<td>153</td>
<td>University of Ljubljana</td>
<td>UNILJ</td>
<td>SI</td>
<td>A</td>
<td>Blaz Zupan</td>
</tr>
<tr>
<td>154</td>
<td>University of Glasgow</td>
<td>UNIGL</td>
<td>UK</td>
<td>A</td>
<td>Muffy Calder</td>
</tr>
<tr>
<td>155</td>
<td>University of Twente-CTIT</td>
<td>CTIT</td>
<td>NL</td>
<td>A</td>
<td>Pieter Hartel</td>
</tr>
<tr>
<td>156</td>
<td>European Molecular Biology Laboratory</td>
<td>EMBL</td>
<td>DE</td>
<td>A</td>
<td>Peer Bork</td>
</tr>
<tr>
<td>157</td>
<td>MPI of Colloids and Interfaces</td>
<td>MPIKG</td>
<td>DE</td>
<td>A</td>
<td>Peter Seeberger</td>
</tr>
<tr>
<td>Participant No</td>
<td>Participant organisation name</td>
<td>Part. short name</td>
<td>Country</td>
<td>Sector (A=Academia, I=Industry, O=Other)</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------</td>
<td>------------------</td>
<td>---------</td>
<td>------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>158</td>
<td>Graz University of Technology</td>
<td>TUG</td>
<td>AT</td>
<td>A</td>
<td>Dieter Schmalstieg</td>
</tr>
<tr>
<td>159</td>
<td>Imec</td>
<td>imec</td>
<td>BE</td>
<td>I</td>
<td>Wilfried Verachtert</td>
</tr>
<tr>
<td>160</td>
<td>Hussman Institute for Human Genomics</td>
<td>HIHG</td>
<td>USA</td>
<td>A</td>
<td>Margaret A. Pericak-Vance</td>
</tr>
<tr>
<td>161</td>
<td>Max Delbrück Center for Molecular Medicine</td>
<td>MDC</td>
<td>DE</td>
<td>A</td>
<td>Miguel Andrade</td>
</tr>
<tr>
<td>162</td>
<td>Bull</td>
<td>Bull</td>
<td>FR</td>
<td>I</td>
<td>Medur Sridharan</td>
</tr>
<tr>
<td>163</td>
<td>Meganexus</td>
<td>Meganexus</td>
<td>UK</td>
<td>I</td>
<td>Daniel Brown</td>
</tr>
<tr>
<td>164</td>
<td>American University of Beirut</td>
<td>AUB</td>
<td>LB</td>
<td>A</td>
<td>Ali Chehab</td>
</tr>
<tr>
<td>165</td>
<td>Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences</td>
<td>Renyi</td>
<td>HU</td>
<td>A</td>
<td>István Miklós</td>
</tr>
<tr>
<td>166</td>
<td>SAP</td>
<td>SAP</td>
<td>DE</td>
<td>I</td>
<td>Cafer Tosun Mathias Steinbrecher</td>
</tr>
<tr>
<td>167</td>
<td>Universidad Politecnica de Madrid</td>
<td>UPM</td>
<td>ES</td>
<td>A</td>
<td>Victor Maojo</td>
</tr>
<tr>
<td>168</td>
<td>Maxeler Technologies</td>
<td>MT</td>
<td>DE</td>
<td>I</td>
<td>Oskar Mencer</td>
</tr>
<tr>
<td>169</td>
<td>University of Sheffield</td>
<td>UNISH</td>
<td>UK</td>
<td>A</td>
<td>Marco Viceconti</td>
</tr>
<tr>
<td>170</td>
<td>Complete Genomics Inc</td>
<td>CG</td>
<td>USA</td>
<td>I</td>
<td>Edgar Funk</td>
</tr>
<tr>
<td>171</td>
<td>Health Level Seven International Foundation</td>
<td>HL7</td>
<td>BE</td>
<td>O</td>
<td>Charles Jaffe</td>
</tr>
<tr>
<td>172</td>
<td>prior Knowledge Inc</td>
<td>prior</td>
<td>USA</td>
<td>I</td>
<td>Vikash Mansinghka</td>
</tr>
<tr>
<td>173</td>
<td>University of Iceland</td>
<td>UICE</td>
<td>IS</td>
<td>A</td>
<td>Ines Thiele</td>
</tr>
<tr>
<td>174</td>
<td>Global Helix LLC</td>
<td>GH</td>
<td>USA</td>
<td>I</td>
<td>Richard A. Johnson</td>
</tr>
<tr>
<td>175</td>
<td>Sabanci University</td>
<td>UNISAB</td>
<td>TK</td>
<td>A</td>
<td>Ugur Secerman</td>
</tr>
<tr>
<td>176</td>
<td>Tallinn University of Technology</td>
<td>IOC</td>
<td>EE</td>
<td>A</td>
<td>Marko Vendelin Andrus Salupere</td>
</tr>
<tr>
<td>177</td>
<td>University of Vienna</td>
<td>UNIVE</td>
<td>AT</td>
<td>A</td>
<td>Herbert Gottweis</td>
</tr>
<tr>
<td>178</td>
<td>University of Florence</td>
<td>UNIFI</td>
<td>IT</td>
<td>A</td>
<td>Mario Pazzagli</td>
</tr>
<tr>
<td>179</td>
<td>Eindhoven University of Technology</td>
<td>TUE</td>
<td>NL</td>
<td>A</td>
<td>Natal van Riel</td>
</tr>
<tr>
<td>180</td>
<td>Instituto de Telecomunicacoes</td>
<td>IT</td>
<td>PT</td>
<td>O</td>
<td>Miguel Coimbra</td>
</tr>
<tr>
<td>181</td>
<td>TUBITAK - Genetics Engineering and Biotechnology Institute of Turkey</td>
<td>GEBI</td>
<td>TK</td>
<td>A</td>
<td>Carlos Salema</td>
</tr>
<tr>
<td>182</td>
<td>Centrum Wiskunde &amp; Informatica</td>
<td>CWI</td>
<td>NL</td>
<td>A</td>
<td>Oğuzhan Külekci</td>
</tr>
<tr>
<td>183</td>
<td>ARCTUR d.o.o.</td>
<td>ARC</td>
<td>SL</td>
<td>I</td>
<td>Arjen de Vries Bořut Šatej</td>
</tr>
<tr>
<td>184</td>
<td>Spanish National Cancer Research Center</td>
<td>CNIO</td>
<td>ES</td>
<td>A</td>
<td>Peter Klatt</td>
</tr>
<tr>
<td>185</td>
<td>University of Helsinki</td>
<td>FIMM</td>
<td>FI</td>
<td>A</td>
<td>Olli Kallioniemi</td>
</tr>
<tr>
<td>186</td>
<td>University of California</td>
<td>UCLA</td>
<td>US</td>
<td>A</td>
<td>Majid Sarrafzadeh</td>
</tr>
<tr>
<td>187</td>
<td>Spanish Institute of Bioinformatics</td>
<td>INB</td>
<td>ES</td>
<td>A</td>
<td>Alfonso Valencia</td>
</tr>
<tr>
<td>188</td>
<td>Cardiff University</td>
<td>CU</td>
<td>UK</td>
<td>A</td>
<td>Ruth Chadwick</td>
</tr>
<tr>
<td>Participant No</td>
<td>Participant organisation name</td>
<td>Part. short name</td>
<td>Country</td>
<td>Sector (A=Academia, I=Industry, O=Other)</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>189</td>
<td>Korea Advanced Institute of Science and Technology</td>
<td>KAIST</td>
<td>KR</td>
<td>A</td>
<td>Doheon Lee</td>
</tr>
<tr>
<td>190</td>
<td>Poznao Supercomputing and Networking Center</td>
<td>PSNC</td>
<td>PL</td>
<td>A</td>
<td>Gerard Frankowski</td>
</tr>
<tr>
<td>191</td>
<td>University of Oslo - Institute of Cancer Research Osaka University Hospital</td>
<td>UIO</td>
<td>NO</td>
<td>A</td>
<td>Anne-Lise Børresen-Dale</td>
</tr>
<tr>
<td>192</td>
<td>Institut Mérieux - bioMérieux (BMER)</td>
<td>IMER</td>
<td>FR</td>
<td>A</td>
<td>Deborah Braun</td>
</tr>
<tr>
<td>193</td>
<td>Institut Mérieux</td>
<td>IMER</td>
<td>FR</td>
<td>A</td>
<td>Deborah Braun</td>
</tr>
<tr>
<td>194</td>
<td>Telethon Institute of Genetics and Medicine</td>
<td>TIGEM</td>
<td>IT</td>
<td>A</td>
<td>Graciana Diez Roux</td>
</tr>
<tr>
<td>195</td>
<td>University of Calgary</td>
<td>UC</td>
<td>CA</td>
<td>A</td>
<td>Christoph W. Sensen</td>
</tr>
<tr>
<td>196</td>
<td>University of Melbourne</td>
<td>UMEL</td>
<td>AU</td>
<td>A</td>
<td>Fernando Martin-Sanchez</td>
</tr>
<tr>
<td>197</td>
<td>University Medical Center Groningen</td>
<td>UMCG</td>
<td>NL</td>
<td>A</td>
<td>Barbara Bakker</td>
</tr>
<tr>
<td>198</td>
<td>University of Amsterdam</td>
<td>UVA</td>
<td>NL</td>
<td>A</td>
<td>Hans Westerhoff, Willem Stiekema</td>
</tr>
<tr>
<td>199</td>
<td>CERN - European Organization for Nuclear Research</td>
<td>CERN</td>
<td>CH</td>
<td>A</td>
<td>Robert Jones</td>
</tr>
<tr>
<td>TBD</td>
<td>BIONEXT, S.A</td>
<td>BN</td>
<td>FR</td>
<td>I</td>
<td>Dr. Laurent-Philippe Albou</td>
</tr>
<tr>
<td>TBD</td>
<td>Barcelona Supercomputing Center</td>
<td>BSC</td>
<td>ES</td>
<td>A</td>
<td>Toni Cortes</td>
</tr>
<tr>
<td>TBD</td>
<td>Istanbul University - Institute of Experimental Medicine</td>
<td>DETAE</td>
<td>TK</td>
<td>A</td>
<td>Prof. Dr. Uğur Özbek- Prof. Dr. Ahmet Gül</td>
</tr>
<tr>
<td>TBD</td>
<td>German Institute of Medical Documentation and Information</td>
<td>DIMDI</td>
<td>DE</td>
<td>O</td>
<td>Hans Peter Dauben</td>
</tr>
<tr>
<td>TBD</td>
<td>European Genetics Foundation</td>
<td>Eurogene</td>
<td>IT</td>
<td>A</td>
<td>Prof. Giovanni Romeo</td>
</tr>
<tr>
<td>TBD</td>
<td>Global Action 4 Health Institute Ltd</td>
<td>GA4H</td>
<td>IE</td>
<td>O</td>
<td>LoS in preparation</td>
</tr>
<tr>
<td>TBD</td>
<td>Hacettepe University</td>
<td>HACET</td>
<td>TK</td>
<td>A</td>
<td>LoS in preparation</td>
</tr>
<tr>
<td>TBD</td>
<td>Institut national de la santé et de la recherche médicale</td>
<td>INSERM</td>
<td>FR</td>
<td>A</td>
<td>Behazine Combadiere</td>
</tr>
<tr>
<td>TBD</td>
<td>National Centre for Rare Diseases</td>
<td>ISS</td>
<td>IT</td>
<td>A</td>
<td>Domenica Taruscio</td>
</tr>
<tr>
<td>TBD</td>
<td>KOL Medical Media</td>
<td>KOLMEDIA</td>
<td>RO</td>
<td>O</td>
<td>Marius Geanta</td>
</tr>
<tr>
<td>TBD</td>
<td>Microdiscovery</td>
<td>MD</td>
<td>DE</td>
<td>I</td>
<td>Johannes Schuchhardt</td>
</tr>
<tr>
<td>TBD</td>
<td>Nhouette Technologies</td>
<td>nhumi</td>
<td>CH</td>
<td>I</td>
<td>Andre Elisseeff</td>
</tr>
<tr>
<td>TBD</td>
<td>Norwegian University of Science and Technology</td>
<td>NTNU</td>
<td>NO</td>
<td>A</td>
<td>LoS in preparation</td>
</tr>
<tr>
<td>TBD</td>
<td>Olink AB</td>
<td>olink</td>
<td>SE</td>
<td>I</td>
<td>Mats Gullberg</td>
</tr>
<tr>
<td>TBD</td>
<td>Osteba, Basque Office for HTA</td>
<td>Osteba</td>
<td>ES</td>
<td>O</td>
<td>Iñaki Gutiérrez-Ibarluzea;</td>
</tr>
<tr>
<td>TBD</td>
<td>Philips Research</td>
<td>PHR</td>
<td>NL</td>
<td>I</td>
<td>Hans Hofstraat</td>
</tr>
<tr>
<td>TBD</td>
<td>Sanofi Aventis</td>
<td>SA</td>
<td>FR</td>
<td>A</td>
<td>Jorge Rosenfeld</td>
</tr>
<tr>
<td>TBD</td>
<td>World Health Connections</td>
<td>WHC</td>
<td>NO</td>
<td>O</td>
<td>Werner Christie</td>
</tr>
<tr>
<td>TBD</td>
<td>Digital Art Forms</td>
<td>DAF</td>
<td>US</td>
<td>I</td>
<td>Jason Jerald</td>
</tr>
<tr>
<td>Participant No</td>
<td>Participant organisation name</td>
<td>Part. short name</td>
<td>Country</td>
<td>Sector (A=Academia, I=Industry, O=Other)</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>TBD</td>
<td>TUBITAK-Bilgem</td>
<td>BILGEM</td>
<td>TK</td>
<td>A</td>
<td>Odd Arild Lehne</td>
</tr>
<tr>
<td>TBD</td>
<td>Pubgene</td>
<td>Pubgene</td>
<td>NO</td>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>
Flagship Final Report
IT Future of Medicine

Annex III
Glossary
Glossary

AACR  American Association for Cancer Research
ALT  Alanine aminotransferase
Analytical data  Data generated from samples which are sent away for analysis in a lab - generally omics-type data.
API  Application Programming Interface WP3.4
ASCO  American Society of Clinical Oncology
AST  Aspartate aminotransferase
Bayesian model  Modelling approach that updates probability estimates for a testable hypothesis while additional evidence is learned. This method is used in statistics to (automatically) ensure that a certain method works as well as any alternative method, which is especially relevant to the dynamic analysis of sequences of data.
BBMRI  Biobanking and Biomolecular Resources Research Infrastructure
BioMedBriges  FP7-funded cluster project that will bring together the ESFRI Research Infrastructures in the Biological and Medical Sciences.
BLUEPRINT  Large-scale research project for BLUEPRINT of Haematopoietic Epigenomes.
Boolean expressions  Logical statement that is either true (=1) or false (=0) as expressed by its numerical value. Boolean expressions can compare data of any type as long as both parts of the expression have the same data type. One can test, for instance, if data is equal to, greater than, or less than other data.
BRIC  Brazil, Russia, India, China
CARRIOCAS  Project studied and deployed an ultra high bit rate optical network (up to 40 Gb/s per channel) able to answer the industrial needs in terms of interactive numerical simulations on remote supercomputers and of processing of huge volumes of distributed data.
CAVE  Cave Automatic Virtual Environment
CD  Circular Dichroism
CEO  Chief Executive Officer
CERN  European Organization for Nuclear Research
CFO  Chief Financial Officer
CID  Collision-Induced Dissociation
CINECA  Non profit Consortium, made up of 51 Italian universities, the National Institute of Oceanography and Experimental Geophysics - OGS, the CNR (National Research Council), and the Ministry of Education, University and Research (MIUR). Today it is the largest Italian computing centre, one of the most important worldwide.
COO  Chief Operating Officer
COPD  Chronic Obstructive Pulmonary Disease
COX  Cyclooxygenase
CP  Collaborative Project
CP-CSA  Combination of CP and CSA
CPU  Central Processing Unit
CSA  Coordination and Support Action
CSC  Center for Science
CSML  Computational Statistics and Machine Learning
CT  Computed Tomography
CUDA
Compute Unified Device Architecture
D2I
Data 2 Intelligence
DEISA
Distributed European Infrastructure for Supercomputing Applications
Deletion
Mutation (a genetic aberration) in which a part of a chromosome or a sequence of DNA is missing
DG ICT
Directorate-General for ICT
DG RTDI
Directorate-General for Research & Innovation
DG SANCO
Directorate-General for Health and Consumers
DICOM
Digital Imaging and Communications in Medicine
DNA
Deoxyribonucleic acid
DSP
Digital Signal Processor
EHR
An electronic health record is an evolving concept defined as a systematic collection of electronic health information about individual patients or populations. It is a record in digital format that is theoretically capable of being shared across different health care settings.
E&TP
Education & Training Programme
EASW
European Awareness Scenario Workshop
EATRIS
European Advanced Translational Research Infrastructure in Medicine
EC
European Commission
ECG
Electrocardiography
ECMWF
European Centre for Medium-Range Weather Forecasts
ECRIN
European Clinical Research Infrastructures Network
EDMA
European Diagnostic Manufacturers Association
EEIG
European Economic Interest Grouping
EESI-2
European Exascale Software Initiative
EFB
European Federation of Biotechnology
EFPIA
European Federation of Pharmaceutical Industries and Associations
EGI
European Grid Infrastructure
EHFG
European Health Forum Gastein
EHR4CR
Electronic Health Records for Clinical Research
EIP
European Infrastructure Project
ELIXIR
European Life Science Infrastructure for Biological information
ELSI
Ethical, Legal, and Social Implications
EMA
European Medicines Agency
EMT
Executive Management Team
EMTRAIN
European Medicines Research Training Network
ENCODEx
Encyclopedia Of DNA Elements
ENSEMBL
Joint scientific project between the European Bioinformatics Institute and the Wellcome Trust Sanger Institute, which was launched in 1999 in response to the imminent completion of the Human Genome Project.
EORTC
European Organisation for Research and Treatment of Cancer
EPCC
Edinburgh Parallel Computing Centre
ERANET
Network of national ministries or agencies that run national programmes in different areas of science and technology. This scheme was designed to promote the coordination of such national programmes in European countries in order to support the development of the European Research Area.
One of the seven flagship initiatives of the Europe 2020 strategy, set out to define the key enabling role that use of ICT will have to play if Europe wants to succeed in its ambitions for 2020.

EuroScan International Information Network on New and Emerging Health Technologies. It is a collaborative network of member agencies for the exchange of information on important emerging new drugs, devices, procedures, programmes, and settings in health care.

European cloud computing platform will support the massive IT requirements of European scientists, supported by industrial partners, and provide stable computing capacities and services that elastically meet demand.

Hadoop Apache project that provides an open-source implementation of frameworks for reliable, scalable, distributed computing and data storage.

HPC High Performance Computing

HTA Health Technology Assessment, a set of procedure that evaluate new technologies that might help improve health care.

ICT Information and Communication Technologies; i.e. the recent computation- and electronics- based technologies that revolutionize the communication between human individuals and between their computational devices and that process and integrate information through massive (as compared to 10 years ago) computations.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDRIS-CNRS</td>
<td>Institut du Développement et des Ressources en Informatique Scientifique</td>
</tr>
<tr>
<td>IEEE</td>
<td>Industrial Electronics Society</td>
</tr>
<tr>
<td>IESP</td>
<td>International Exascale Software Project</td>
</tr>
<tr>
<td>IHEC</td>
<td>International Human Epigenome Consortium</td>
</tr>
<tr>
<td>IMI</td>
<td>Innovative Medicines Initiative</td>
</tr>
<tr>
<td>IMMCM</td>
<td>Interacting Molecular Monte Carlo Models</td>
</tr>
<tr>
<td>IMO</td>
<td>Intelligent Monitoring for Health and Well-being</td>
</tr>
<tr>
<td>INAHTA</td>
<td>International Network of Agencies for Health Technology Assessment</td>
</tr>
<tr>
<td>Indel</td>
<td>Special mutation class, defined as a mutation resulting in a colocalised insertion and deletion and a net gain or loss in nucleotides.</td>
</tr>
<tr>
<td>Individualised model</td>
<td>An executable model parameterised with data collected about an individual.</td>
</tr>
<tr>
<td>IP</td>
<td>Intellectual Property</td>
</tr>
<tr>
<td>IPROMS</td>
<td>Innovative Production Machines and Systems</td>
</tr>
<tr>
<td>IPS</td>
<td>Induced Pluripotent Stem Cell</td>
</tr>
<tr>
<td>ISBE</td>
<td>Infrastructure for Systems Biology Europe</td>
</tr>
<tr>
<td>ISOBM</td>
<td>International Society of Oncology and Biomarkers</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technologies</td>
</tr>
<tr>
<td>ITFoM</td>
<td>IT Future of Medicine</td>
</tr>
<tr>
<td>ITER</td>
<td>ITER means ‘the way’ in Latin – is an international research and development project conceived to take the next major step in the development of fusion energy.</td>
</tr>
<tr>
<td>JTI</td>
<td>Join Technology Initiative</td>
</tr>
<tr>
<td>KIC</td>
<td>Knowledge and Innovation Community</td>
</tr>
<tr>
<td>LAL</td>
<td>Learning Adapting Leveling</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatography–Mass Spectrometry</td>
</tr>
<tr>
<td>LOH</td>
<td>Loss of Heterozygosity</td>
</tr>
<tr>
<td>M-health</td>
<td>Mobile health</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>MCS</td>
<td>Monte Carlo Simulation</td>
</tr>
<tr>
<td>Medical data</td>
<td>(i) Clinical data - routinely recorded data from the clinic relating to an individual. (ii) Lifestyle data - data about an individual's lifestyle. (iii) Environmental data - data about an individual's environmental exposure.</td>
</tr>
<tr>
<td>METAHIT</td>
<td>Metagenomics of the Human Intestinal Tract consortium</td>
</tr>
<tr>
<td>MML</td>
<td>Multiscale Modelling Language</td>
</tr>
<tr>
<td>Model</td>
<td>An executable representation of a biological system which has inputs and outputs.</td>
</tr>
<tr>
<td>MolPage</td>
<td>Molecular Phenotyping to Accelerate Genomic Epidemiology</td>
</tr>
<tr>
<td>MolTool</td>
<td>Advanced molecular tools for array-based analyses of genomes</td>
</tr>
<tr>
<td>MoU</td>
<td>Memorandum of Understanding</td>
</tr>
<tr>
<td>MPI</td>
<td>Message Passing Interface</td>
</tr>
<tr>
<td>MR</td>
<td>Magnetic Resonance</td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA (Ribonucleic acid)</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MSB</td>
<td>Molecular Systems Biology, which is an area of systems biology that begins from interacting molecules and until ITFoM did not really</td>
</tr>
</tbody>
</table>
extend beyond the whole-cell level.

NCBI  National Center for Bioinformatics
NCI  National Cancer Institute
NGFN  National German Genome Research Network (Nationales Genomforschungsnetzwerk)
NMR  Nuclear Magnetic Resonance
NUTS  Nomenclature of Units Territorial for Statistics
OECD  Organisation for Economic Co-operation and Development
OncoTrack  European consortium for novel genomic cancer diagnostics
OpenCL  Open Computing Language
OpenMP  Open Multi-Processing
OVF  Open Virtualization Format
PBMC  Peripheral Blood Mononuclear Cell
PC  Project Coordinator
PET  Positron Emission Tomography
PET-MS  Positron Emission Tomography-Mass Spectrometry
PGP  Personal Genome Project
PHAT  Public Health Assessment Tools
PHGEN  Public Health Genomics European Network
PIC  Programme Integration Committee
PK/PD  Pharmacokinetic/Pharmacodynamic
PM  Personalized Medicine
PMI  Personalized Medicine Initiative
PoC  Point of Care
PR  Public Relations
PRACE  Partnership for Advanced Computing in Europe
R&D  Research and Development
READNA  Revolutionary Approaches and Devices for Nucleic Acid Analysis
Reference model  A non-individualised executable model of a specific biological system constructed from the reference networks and reference data sets.
Reference network  A collection of biological data gathered from the literature and experimental results, describing the consensus relationship between entities in a biological system (e.g. metabolites and enzymes or nutrients and muscle health). This is the de-facto template that may have measured parameters added to it to produce a reference model.
RNA  Ribonucleic acid
ROI  Return on Investment
RZG  Rechenzentrum Garching
S&T  Science & Technology
SEAB  Scientific Ethical Advisory Board
SaIWe  Strategic Center for Health and Wellbeing
SC  Steering committee
SF  Stakeholder Forum
shRNA  Short hairpin RNA
SME  Small and Medium Enterprises
SML  Statistical Machine Learning
SNP  Single Nucleotide Polymorphism
SNV  Single Nucleotide Variant
SOP  Standards Operating Procedure
SPIDIA Standardisation and Improvement of generic Pre-analytical tools and Procedures for in-vitro Diagnostics
SRM Standard Reference Method
SWP Sub-workpackage
TB Terabyte
Translocation A chromosome translocation is a chromosome abnormality caused by rearrangement of parts between non homologous chromosomes.
TT Technology Transfer
US Ultrasound
USA United States of America
Use case Vertical activities that produce models of various specific diseases or systems (e.g. cancer, neurodegenerative disease).
Virtual Individual/Virtual Patient An ICT-mathematical model of the common-denominator human/
A clinically utilisable integrated set of individualised models corresponding to a patient
VLN Virtual Liver Network
VPH The Virtual Physiological Human refers to a consortium of research groups that have been and are active in physiology-based modelling of the human. The consortium is organized in the Virtual Physiological Human Institute, which currently represents over 60 institutions involved with VPH research.
WP Workpackage
Flagship Final Report
IT Future of Medicine

Annex IV
Multimedia material
ITFoM is present on the web at www.itfom.eu
Facebook: http://www.facebook.com/itfom
Twitter: http://twitter.com/ITFoM
LinkedIn: IT Future of Medicine

The ITFoM movie can be accessed at:
http://www.itfomthemovie.eu/

People can:
- Share the movie through Facebook and Twitter
- View the movie directly
- Embed the movie into their own website
- Download the movie for desktop, Ipad and mobile phone.

Further multimedia material (publications, flyer, poster, press releases, videos and audio records) is accessible on the ITFoM Website at http://www.itfom.eu/digital-media