

Network and systems medicine: Position paper of the European Collaboration on Science and Technology action on Open Multiscale Systems Medicine

Blandine Comte^{01*}, Jan Baumbach⁰², Arriel Benis⁰³, José Basílio⁰⁴, Nataša Debeljak⁰⁵, Åsmund Flobak^{06a,b}, Christian Franken^{07a,b}, Nissim Harel⁰³, Feng He⁰⁸, Martin Kuiper⁰⁹, Juan Albino Méndez Pérez¹⁰, Estelle Pujos-Guillot⁰¹, Tadeja Režen¹¹, Damjana Rozman¹¹, Johannes A. Schmid⁰⁴, Jeanesse Scerri¹², Paolo Tieri¹³, Kristel Van Steen¹⁴, Sona Vasudevan¹⁵, Steven Watterson¹⁶, Harald H.H.W. Schmidt¹⁷

⁰¹Université Clermont Auvergne, INRAE, UNH, Plateforme d'Exploration du Métabolisme, MetaboHUB Clermont, F-63000 Clermont-Ferrand, France. Tels: +33-473624687, +33-473624141; Emails: blandine.comte@inrae.fr, estelle.pujos-guillot@inrae.fr

⁰²TUM School of Life Sciences Weihenstephan (WZW), Technical University of Munich (TUM), Maximus-von-Imhof-Forum 3, 85354, Freising-Weihenstephan, Germany. Tel: +49-8161-71-2136; Email jan.baumbach@wzw.tum.de

⁰³Holon Institute of Technology, Faculty of Technology Management, Holon, Israel. Tels: +972 (0)3-5026892, +972 (0)3-5026534; Emails: arrielb@hit.ac.il, nissimh@hit.ac.il

⁰⁴Inst. of Vascular Biology and Thrombosis Research, Center for Physiology and Pharmacology, Medical Univ. of Vienna, Austria. Tels: +43-1-40160-31162, +43-1-40160-31155; Emails: jose.basilio@meduniwien.ac.at, johannes.schmid@meduniwien.ac.at

⁰⁵Medical centre for Molecular Biology, Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Slovenia. Tel: +386-1-543-7645; Email: natasa.debeljak@mf.uni-lj.si

^{06a}Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway. ^{06b}The Cancer Clinic, St Olav's University Hospital, Trondheim, Norway. Tel: +47 725 73717; Email: asmund.flobak@ntnu.no

^{07a}Digital Health Systems, 89079 Eisingen, Germany. Tel: +49 151 19603414; ^{07b}Department of Pharmacology & Personalised Medicine, Faculty of Health, Medicine & Life Science, Maastricht University, Universiteitssingel 40, 6229 ER Maastricht, the Netherlands; Email: c.franken@maastrichtuniversity.nl

⁰⁸Department of Infection and Immunity, Luxembourg Institute of Health, 29 rue Henri Koch, L-4354 Esch-sur-Alzette, Luxembourg. Tel: +352 26970-343; Email: Feng.He@lih.lu

⁰⁹Department of Biology, Faculty of Natural Sciences, Norwegian University of Science and Technology, Trondheim, Norway. Tel: +47 735 50348; Email: martin.kuiper@ntnu.no

¹⁰Department of Computer Science and Systems Engineering, Universidad de La Laguna, Tenerife, Spain. Tel: +34 9223 18265; Email: jamendez@ull.edu.es

¹¹Centre for Functional Genomics and Bio-Chips, Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Slovenia. Tels: +386-1-543-7591, +386-1-543-7592; Emails: damjana.rozman@mf.uni-lj.si, tadeja.rezen@mf.uni-lj.si

¹²Department of Physiology & Biochemistry, Faculty of Medicine & Surgery, University of Malta, Msida, Malta. Tel: +356 7945 9473; Email: jsce4@um.edu.mt

¹³CNR National Research Council, IAC Institute for Applied Computing, Via dei Taurini 19, 00185 Rome, Italy. Tel: +39 06 49 93 73 49; Email: paolo.tieri@cnr.it

¹⁴GIGA-R Medical Genomics - BIO3, University of Liège, Belgium. Tel: +32 (0)43662692; Email: kristel.vansteen@uliege.be

¹⁵Georgetown University Medical Centre, Washington D.C. USA. Tel: +1202-687-2242, Email: sv67@georgetown.edu

¹⁶Northern Ireland Centre for Stratified Medicine, Ulster University, Derry, Co Londonderry, Northern Ireland, BT47 6SB, UK. Tel: +44 28 7167 5665; Email: s.watterson@ulster.ac.uk

¹⁷Department of Pharmacology & Personalised Medicine, Faculty of Health, Medicine & Life Science, MeHNS, Maastricht University, Universiteitssingel 40, 6229 ER Maastricht, the Netherlands. Tel: +31 43-3881421; Email: h.schmidt@maastrichtuniversity.nl

Running title: Implementation of systems medicine: a review

***Corresponding author:** Blandine Comte, Université Clermont Auvergne, INRAE, UNH, Plateforme d'Exploration du Métabolisme, MetaboHUB Clermont, F-63000 Clermont-Ferrand, France. Email: blandine.comte@inrae.fr

Keywords: systems medicine, omics, big data, data integration, integrated healthcare

Authorship Confirmation Statement:

B. C. designed and directed the project. B. C. further wrote part of the manuscript, specifically the 'Abstract', 'Introduction', 'Multidimensional omics', 'FAIR', 'Embracing the challenge of data integration and validation', 'SWOT Analysis'.

J. A.M.P. wrote part of the manuscript, specifically the 'Personalized drug therapy, Drug dosing'.

Jan B. wrote part of the manuscript, specifically the 'Basic science and data for network and systems medicine: FAIR, Privacy and federated machine learning'.

A. B. wrote part of the manuscript, specifically the 'Multidimensional omics', 'Big data availability and information systems', 'Enlarging network and systems medicine applications: Improving patient engagement and treatment adherence'.

José B. wrote part of the manuscript, specifically the 'Why we need new medicine', 'Definitions and goals of network and systems medicine', 'SWOT Analysis'.

N. D. wrote part of the manuscript, specifically the 'Enlarging network and systems medicine applications: training'.

N. H. wrote part of the manuscript, specifically the 'Multidimensional omics', 'Big data availability and information systems', 'Enlarging network and systems medicine applications: Improving patient engagement and treatment adherence'.

Å. F. wrote part of the manuscript, specifically the 'Cancer pathways and personalized therapy'.

C. F. wrote part of the manuscript, specifically the 'Why we need new in medicine' and '2D-, 3D drug printing'.

F. H. wrote part of the manuscript, specifically the 'Definition and goals of network and systems medicine'.

N. H. wrote part of the manuscript, specifically the 'Multidimensional omics data generation', 'Big data, data analytics, and information systems', 'Improving patient engagement and treatment adherence'.

M. K. wrote part of the manuscript, specifically the 'Cancer pathways and personalized therapy'.

E. P-G. wrote part of the manuscript, specifically the 'Abstract', 'Introduction', 'Multidimensional omics', 'FAIR', 'Embracing the challenge of data integration and validation', 'SWOT Analysis'.

T. R. wrote part of the manuscript, specifically the 'Enlarging network and systems medicine applications: training'.

D. R. wrote and revised a part of the manuscript, particularly the 'Enlarging network and systems medicine applications: training'.

J. S. wrote part of the manuscript, specifically the 'SWOT Analysis'.

J.A. S. wrote part of the manuscript, specifically the 'Definition and goals of network and systems medicine', 'Enlarging network and systems medicine applications: training'.

P. T. wrote part of the manuscript, specifically the 'Concluding remarks and outlook'.

K. V.S. wrote part of the manuscript, specifically the 'Embracing the challenge of data integration and validation', 'Out-box of thinking', 'Validation'.

S. V. wrote part of the manuscript, specifically the 'Enlarging network and systems medicine applications: training'.

S. W. wrote part of the manuscript, specifically describing systems biology file formats and standards: 'Basic science and data for network and systems medicine: standardization'.

H.H.H.W. S. wrote part of the manuscript, specifically the 'Why we need new in medicine', 'Personalized drug therapy: 2D-, 3D drug printing, Drug repurposing'.

All authors have reviewed and approved of the manuscript prior to submission.

The manuscript has been submitted solely to the Network and Systems Medicine journal, and is not published, in press, or submitted elsewhere.

Author Disclosure Statements:

No competing financial interests exist.

Abstract:

Network and systems medicine has rapidly evolved over the last decade, thanks to computational and integrative tools, which stem in part from systems biology. However, major challenges and hurdles are still present concerning validation and translation into clinical application and decision-making for precision medicine. In this context, the Collaboration on Science and Technology (COST) action on Open Multiscale Systems Medicine (OpenMultiMed) reviewed the available advanced technologies for multidimensional data generation and integration in an open-science approach as well as key clinical applications of network and systems medicine and the main issues and opportunities for the future. The development of multi-omic approaches as well as new digital tools provides a unique opportunity to explore complex biological systems and networks at different scales. Moreover, the application of FAIR (Findable, Applicable, Interoperable and Reusable) principles and adoption of standards increases data availability and sharing for multiscale integration and interpretation. These innovations have led to first clinical applications of network and systems medicine particularly in the field of personalized therapy and drug dosing. Enlarging network and systems medicine application would now imply to increase patient engagement and healthcare providers as well as to educate the novel generations of medical doctors and biomedical researchers to shift the current organ- and symptom-based medical concepts towards network- and systems-based ones for more precise diagnoses, interventions and ideally prevention. In this dynamic, the healthcare system will also have to evolve if not revolutionize in terms of organization and management.

Introduction

Why we need new medicine

About 70% of all medical interventions pertain to the prescription of a drug. However, for several drugs that are on the market, population-based studies fail to show patient-relevant benefits.^{1,2} Research covering drug approval since the 1970s suggests only a limited number of new drugs provide real advances over existing ones; most studies place the proportion of true innovation at under 15%.³ For every person they do help, the 10 highest-grossing drugs in the USA fail to improve the conditions of most other patients, leading to so-called high 'numbers needed to treat' (NNT, Fig 1c).⁴ Since the 1950s, the efficacy of translating biomedical research into successful drug discovery is on a constant decline.^{4,5} Two key factors have contributed to this innovation roadblock. One is the irreproducibility of pre-clinical and basic research data⁶ to which, besides data manipulation/fabrication, poor study quality, such as lack of statistical power and positive publication bias by scientific journals, are the main contributors.^{7,8} The second reason is the conceptual medical knowledge gap about many of our current, in particular chronic, disease definitions. Except for infectious and rare diseases, most disease definitions are based on signs and reported symptoms, pointing to organ-centric mechanisms (see Fig. 1b), and not on causal molecular mechanisms which for many diseases are not even known. Consequently, there is a disease-based care system, where the focus is put on treating and monitoring the symptoms (see Fig. 1a), rather than true health care that treats the causes by helping to implement lifestyle decisions for a healthier life.

A drug can only be effectively developed and applied in a precise manner if the molecular disease mechanism is known. Not knowing a disease mechanism also affects basic and pre-clinical (animal) research, where often cellular or animal models that mimic symptoms of human disease are used, but neither the underlying mechanism of the animal model nor that of the human disease is known.⁹⁻¹⁴ Noteworthy exceptions to these limitations and shortcomings are again infectious or rare diseases, where a precise - often monogenetic - mechanism is known. Many common and complex clinical disease phenotypes, once they are fully endo-phenotyped and mechanistically understood, will segregate into several distinct mechanotypes.¹⁵ Many common diseases appear complex because we combine several molecular diseases under one umbrella term based on shared prominent signs and symptoms. For example, high blood pressure is in 95% of the cases diagnosed as Essential Hypertension, meaning that the blood pressure is elevated, but we do not know why. These patients are then treated with different blood vessel-dilating drugs and the clinical sign, elevated blood pressure, disappears, yet the cause is not known and remains untreated. Once mechanistically understood, chronic diseases can be cured or even prevented and no longer just treated (see Fig. 1d). The UK National Institute for Health and Care Excellence (NICE) published a list of the absolute benefits of treatments of common conditions in terms of their NNT. For example, for every thousand low-risk patients prescribed statins for primary prevention, only a single

stroke is prevented per year; one needs to treat more than a thousand patients with antihypertensives per year to prevent one death; and over 800, to prevent one heart attack; nearly 600, to prevent one stroke. In high-risk patients, the NNTs are smaller, but the problem persists.¹⁶ Thus, a move towards a more precise ideally curative therapy that works for almost every patient is of utmost importance.

Applied biomedical research and drug discovery

This efficacy problem also pertains to basic research and its translation into applications such as drug discovery. Of 25,190 articles published from 1979 to 1983 in the six basic science journals, *Nature*, *Cell*, the *Journal of Experimental Medicine*, and the *Journal of Clinical Investigation*, which had the highest impact factors in 2000, and the *Journal of Biological Chemistry*, which received the most citations, only a single claim of relevance may have led to actual application follow-up.¹⁷ Concerning the discovery of new drugs, for three decades, costs have increased exponentially and are now stable at an extremely high level. Since the 1950s, however, the efficacy of drug discovery is on a constant logarithmic decline indicating a fundamental and conceptual problem of how we define and approach disease.⁵ For drug discovery, recently, systematic drug repurposing is being increasingly explored and represents a conceptual change to a mechanism-based disease definition allowing for a mechanism-based patient stratification, which increases the precision for any subsequent mechanism-based drug intervention. This will massively de-risk drug development, yet at the downside that in the future, drugs will be developed for much smaller patient numbers.

From single targets to validated, causal networks

In the diseasome, disease clusters are mechanistically defined by several genes and proteins forming a signaling network.¹⁸ This has been most extensively hypothesized for 3 distinct networks for macular degeneration^{18,19} and cancer.²⁰ The validity of these networks is essential because both the diagnostic and therapeutic strategies reside on it. Defining the causal signaling network is not trivial and not at all obvious. It is state-of-the-art to rely on highly curated signaling pathway databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) or WikiPathways, a collection of manually drawn pathway maps representing our apparent knowledge on molecular interactions, reactions and relation networks, or review articles. KEGG, however, shows 29 cyclic GMP and 12 reactive oxygen pathways, none of which is comprehensive and all of which fail to cover a recently discovered functional and molecular link between the two,²¹ uniting both in fact to one network. Moreover, subcellular compartmentalization and transition over time also matter in defining disease modules,^{18,21} contributing to further deviation from static pathway concepts.

Mechanism-based diagnostics

Biomarkers are increasingly hypothesized as important for precision medicine,^{22–24} although the term is ambiguous and used for several applications such as screening, stratification, efficacy, differentiation, toxicity and prognosis. In place of validated causal disease mechanisms, the state-of-the-art biomarkers used for these purposes are mostly correlative surrogate and omic markers, rarely established risk factors and never a full functional analysis of a patient's activity state of a causal disease mechanism.²² The limited availability of predictive and precise biomarkers represents a key bottleneck in the progress from mechanism-based disease definition to clinical validation by mechanism-based therapeutic intervention. This causes inefficient drug therapy and clinical trials with a high failure rate (see above).

From single or combination therapy to network pharmacology

The first line of treatment for many diseases involves the administration of a single drug, assuming a single relevant target. If the therapeutic effect is insufficient, drugs are combined. Sometimes, such combination therapies can get out of control when so-called poly-pharmacy results in ten or more drugs being prescribed to a patient with unwanted drug-drug interactions and side effects. Network pharmacology may be easily confused with such combination therapies. The important difference, however, is that in combination therapy symptomatically acting drugs on unrelated targets are combined and act in an additive manner, while in network pharmacology, all drugs act on the same and causal network, and are thereby highly synergistic. This allows for a substantial reduction in the dose of each drug whilst still achieving the same therapeutic effect. This will in all likelihood reduce any side effect of each drug or possible unwanted drug-drug interactions.^{21,25}

Definition and goals of network and systems medicine

Major socio-economic innovations are not only triggered by unmet needs as above, but also by critical technological advances. Insofar, network and systems medicine would not have emerged without decades of development of its antecedent discipline, i.e., systems biology as defined by pioneers in the field.^{26,27} Therefore, it is impossible to discuss systems medicine alone without first talking about systems biology. Systems biology emphasizes analyzing interactions within complex biological systems using holistic and integrative high-throughput experimental and computational approaches. One of the hallmarks in the complex systems, such as multicellular organisms and multi-organ organisms (e.g., animals and humans) is that several components (different cell types, tissues or organs) interact with each other as a local subnetwork or global network to generate emergent effects.^{28,29} The challenges behind this hallmark

cannot be solved *per se* by the reductionist paradigm that decomposes the complex systems into smaller and simpler components and understands their functions and roles one by one. Thus, the emergence of systems biology is to tackle the essential limits of the reductionist approaches.³⁰

While systems biology focuses on basic mechanisms and principles in biology or at most translational preclinical research, systems medicine aims to directly handle the challenges related to health and diseases.^{31–33} In a way, it can also be considered as a modern advancement of physiology. So far, there is no consensus on the definition of systems medicine, an emergent and fast evolving field. Our perception about systems medicine is the application systems biology approaches to the clinical settings of individuals by the combination of large-scale multi-layer, high-throughput quantitative molecular and image measurements at different spatial scales (from molecules, through cells and tissues, to organs), over various time scales, with different types of clinical information.³⁴ The aims of systems medicine are also multi-dimensional, i.e. from the understanding of disease mechanisms to accurate diagnosis, prediction and eventual prevention using accessible biopsies, tissues and samples, to patient subgroup stratification of complex diseases, to the development of novel approaches in drug discoveries, to more precision treatment based on tailored measurements of distinct patients.³⁵ Systems medicine is based on a holistic approach to medicine in opposition to the current symptom/organ-based view. As proposed by Leroy Hood, systems medicine should eventually enable predictive, preventive, personalized and participatory (P4) medicine to improve the wellness of our society.³⁶ Network and systems medicine is at the crossroad of pure and applied sciences, wet and dry labs, life, and computer sciences. Main scientific and technological components of this new field have therefore yet to emerge and evolve into a well-established process.

Network and systems medicine is built similarly to a modern knowledge discovery flow. To implement systems medicine, as the name “systems” indicated, the first required technology should be the development of system-level multidimensional technologies. In this context, the present review on systems medicine first introduces the intended outcomes and definitions. It discusses in a first section the required technologies for multi-dimensional data generation, the current data availability, as well as the computational tools for data integration and interpretation. In a second section, it illustrates its potentiality through some clinical applications. Finally, it discusses the current remaining issues and prospects of this large domain.

Basic science and data for network and systems medicine

Network and systems medicine is built similarly to a modern knowledge discovery flow. In particular, multidimensional omics data generation and integration are key elements in the big data analytics era.

Multidimensional omics

The importance of taking into account the complexity of biological systems has been recognized as the basis of systems approaches. Indeed, it appears that in individuals, a different combination of genetic and environmental factors defines the pathology progress, which accumulates with age. We face co-occurrence of pathologies in the ever-aging population. In addition to cardiovascular complications, there is a rise in neurodegenerative pathologies and metabolic pathologies, where diabetes mellitus and non-alcoholic fatty liver disease are among key components. It is important to note that metabolic diseases (metabolic syndrome, type 2 diabetes, osteoporosis, etc.) show strong comorbidities or co-occurrence with other diseases, such as cardiovascular diseases, cancers and even neurodegenerative diseases, all major health problems of today's societies.^{37–39} There is a challenging situation where, on one hand there is a large progress in understanding the molecular players of disease stages and overlap with other diseases, while on the other hand, the inconsistencies from different studies and different populations leave the impression that we are indeed at the start. In this context, the objective of omics research within systems medicine is to study and understand regulatory mechanisms, identify corresponding specific biomarkers, and characterize their interaction within and between systems,⁴⁰ with the analysis of large sets of biological molecules, including genomics, epigenomics, proteomics, metabolomics and much more, in combination with methodologies from the computer and mathematical science.

Genome-wide association studies (GWASs) have shown their importance in the discovery of single-nucleotide polymorphisms (SNPs) as markers associated with disease-specific clinical phenotypes or their risk factors. As an example, in liver diseases, the genome wide association studies, transcriptome analyses, meta analyses and other clinical studies in different populations and ethnic backgrounds are until 2019 concordant in polymorphisms of a single gene *PNPLA3*.⁴¹ However, liver pathologies remain a major health burden of modern societies where sex dimorphism remains crucial, yet a neglected factor.^{42,43} Generally speaking, complex disease phenotypes can rarely be explained by a single gene, and genomic analysis integrated with protein-protein interaction networks have evidenced the role of groups of genes and variants, and new pathways in multiple diseases.⁴⁴ Therefore, the need for new disease risk models has emerged, including not only genetic factors, transcripts and proteins but also elements such as metabolites, the metabolome being closer to the phenotype. Metabolomics, described as a global analysis of small molecules present in a biofluid (blood, urine, saliva...), produced or modified as a result of stimuli (intervention, drug, genetic perturbations, etc.),^{45,46} is giving an integrated view of metabolism. Among different approaches, the untargeted strategy is a data-driven approach dedicated to biomarker discovery. Based on the use of multiple analytical platforms, such as mass spectrometry, it allows the detection of thousands of features and offers the possibility of characterizing global alterations associated with disease conditions.⁴⁷ It has been widely applied in epidemiology for metabolic disease diagnosis and candidate biomarker discovery, pathophysiological exploration of underlying mechanisms and for diagnosis and prognosis.^{48,49} It is now recognized as a powerful phenotyping tool to better understand not only the biological mechanisms involved in pathophysiological processes but also the complexity of

regulations in interaction with environmental factors. The concept of the exposome was defined to characterize the environmental exposure in a broad sense of 'non-genetic' factor, considering internal, specific external and general external exposure.⁵⁰ In particular, important advances have also been done for the identification of the contribution of the microbiome to the human metabolome and to study their interactions.^{51–53} Associations between nutrition, microbiota, and immune system are being actively studied as contributors to chronic metabolic diseases.⁵⁴

The application of the multi-omics approach has been shown of great interest to better characterize the complexity of phenotypes in human cohorts, but its translation to the clinical setting remains to be developed. Technical advances in biomarkers and personal monitoring devices open the door to translate the concept to utility and increase the completeness of the human system. Integrating communication tools and the exposome as a full part of systems in medicine is now under development, as are analytics that can make full use of the complexity of multidimensional omics data (see the section about data integration).

Big data availability and information systems

The emergence of powerful approaches allowed large datasets to be produced and analyzed, in the perspective of developing decision-making tools for health management. One of the challenges is the security of personal and private health data.⁵⁵ Moreover, the generation of high volumes of big omics data, combined with healthcare provider's high rate of data generation (also known as data velocity) constitute a critical challenge for supporting research and practical implementation of system medicine and tools. Securing personal and private health data is an additional crucial challenge for managing systems. Therefore, the future development of systems medicine requires advanced informatics tools for merging different nature of data to be shared among different communities.

FAIR, Privacy and federated machine learning

Big data also harbors risks to the safety of sensitive clinical data, in particular, when such data needs to be copied to clouds to provide software for learning statistical models with the required large-volume, high-quality data. The barrier for secure health data exchange over the internet is perceived to be insurmountable, thus posing a massive bottleneck hampering big data and prohibiting progress in computational systems medicine. Therefore, it makes the development of medical artificial intelligence (AI) tools for prognosis, response prediction or treatment optimization *de facto* impossible, as sharing and cloud-based storage of health data is ethically problematic and often legally prohibited. Modern omics technologies have paved the way for large-scale quantitative profiling of all kinds of biomolecules (genome, mRNAs, proteins, small molecules). With such data for many patients, we can build computational models that can predict medically relevant features (biomarkers). The PAM50 gene signature and the MammaPrint panel are such biomarker models,^{56,57} helping clinicians to determine whether a breast cancer patient will benefit from chemotherapy, and from what kind of chemotherapy.

However, recent results raise concerns regarding their predictive clinical value.^{58,59} The major problem is the selection of biomarkers due to the small number of samples compared to a very high number of features. The Cancer Genome Atlas (TCGA)⁶⁰ and the International Cancer Genome Consortium data portal (ICGC)⁶¹ are by far the most comprehensive repositories for clinical cancer omics data worldwide. For breast cancer, gene expression data for less than 2,000 patients are available. These few thousand samples, however, stand against more than 20,000 genes that artificial intelligence may combine to predict the outcome. Even when following best practice in machine learning (ML), the consequence is model overfitting and a significantly reduced impact of such kind of AI-based medical diagnostics tools. Big data is clearly in its infancy, even in oncology (the most advanced research area of precision and systems medicine). At the same time, one in eight women (ca. 12%) will develop breast cancer. In the EU, there are over 350,000 new cases per year.⁶² How come that we need to train ultra-high-dimensional AI models with >20,000 features (genes) on <2,000 breast cancer samples, while in the EU alone >350,000 new cases occur every year that are often investigated using gene expression (PAM50, MammaPrint)? Even if only 20% of them were analyzed computationally, and even if only 50% of the samples would be of sufficient quality to be used for AI learning, over the last five years we could have accumulated >300,000 samples - in contrast to the abovementioned 2,000 samples, exemplifying how far we are away from big data analytics in precision and systems medicine. Legal and ethical considerations dictate these circumstances. Patient data may not be shared, in particular neither the molecular data and nor the electronic health records, and most certainly not over the internet. In contrast, conventional AI tools require access to all data locally for training, resulting in the need to aggregate available data in a centralized cloud repository. But data protection legislation usually prohibits depositing sensitive medical patient data in central storage outside the hospital, with massive consequences. For example, the EU's laudable attention to privacy and respective national legislation is further feeding this problem and creating contradicting requirements: The General Data Protection Regulation (GDPR) and its national implementations, as well as the criminal laws on confidential medical communication and the restrictions in terms of data "ownership" prohibit the exchange of sensitive patient data,⁶³ while at the same time the FAIR principles are enforced e.g., in the H2020 programme, where projects are required to make research data publicly available.⁶⁴ One potential way out of this dilemma is federated machine learning. However, many challenges are to be overcome until client-sided ML becomes ubiquitous;⁶⁵ but recent experiments on deep learning demonstrated that it can be made practical and that there are many intriguing opportunities.⁶⁶ Europe has dedicated research projects to this task, e.g., FeatureCloud (<https://featurecloud.eu>), but no applications for federated systems medicine (e.g., for federated network enrichment, federated composite biomarker extraction, or federated mechanotyping) exist yet, but would be necessary to make systems medicine big-data-ready.

From a Healthcare Services perspective, data availability is a sensitive matter. For example, in the United Kingdom, the National Health Services (NHS) is financed by the government but healthcare customers' data are not centralized. In contrast, in some countries, such as Denmark^{67,68} and Israel although the healthcare system is also financed and regulated by the government, healthcare customers' data is

centralized. In Israel, these data are centralized by the Healthcare Management Organizations (HMO) while some of the data reside at hospitals and the Healthcare ministry. The data available to the HMO include, among others, socio-demographic data, and information on biological tests, clinical examinations, pharmacological treatments and communication channels. These data have been continuously collected and stored at the healthcare customer level for the last 25 years⁶⁹ In Israel, the healthcare data *de facto* fit the FAIR standards.^{70,71}

Standardization

Large data sets are often generated at great cost, consuming significant time and resources. Critical to realizing their full value is that we can quickly and easily deploy a diverse and well-developed set of software tools for analysis. This is best achieved when the data sets are made available following a common set of data standards used by a wide range of software tools. Conversely, the incentive for developing software tools is invariably stronger when a rich landscape of suitable data already exists that is easily accessible *via* data standards. Hence, data standards are vital to realizing the potential in large data sets. The adoption of BAM and VCF file formats, for example, has underpinned the explosive growth in the availability of genome data and the software tools for analysis. As we move into the era of systems medicine, data standards will be central to maximizing the value we derive from systems-level data sets.

Larger and more complex data sets require more sophisticated analysis and, as analysis grows in sophistication, it becomes increasingly challenging to reproduce. This is in part due to the network of software dependencies associated with the analytical tools and in part due to the array of design choices that form part of the analytical workflow. Hence, we not only need standards to optimize data availability, but we also need standards that make analyses reproducible and verifiable.

Amongst the key current standards that will support the growth of systems medicine are the Systems Biology Graphical Notation (SBGN), a set of symbols and usage rules that have been developed through open community action, as tools for mapping out the network of molecular interactions between genes, proteins and small molecules.⁷² Although many mapping systems exist, SBGNs strength lies in its lack of ambiguity and its machine-parsable structure, which means that maps can be translated straight to mathematical models. Three different flavors of SBGN have been created with the Process Diagram (SBGN-PD) providing the most detail and the highest level of parsability while the activity Flow (SBGN-AF) and Entity Relationship (SBGN-ER) provide greater levels of abstraction and lower levels of parsability (for details see <https://sbgn.github.io/>). Such models typically require great effort to develop^{73,74} and to facilitate their reuse, expansion and refinement, they can be made available using the SBGN-ML file format that captures the structure of the maps and encodes the biological meaning of the symbols.⁷⁵ It is based on coded use of plain text and therefore it can be edited not only in SBGN-ML compliant software but also manually in any text editor. The mathematical model that facilitates simulation of the pathways can be encoded and disseminated using the Systems Biology Markup Language (SBML). Similar to SBGN-ML, it makes use of a coded plain text file format that can be edited either by SBML compliant software

or manually in any text editor.⁷⁶ However, SBML captures the system of Ordinary Differential Equations that describe the kinetics of all the interactions between genes, proteins and small molecules in a machine parsable form along with their structure and meaning (for details see <http://sbml.org/>).

Ensuring the reproducibility of simulations and computational experiments requires the adoption of the above standards for the maps and models as well as a further set of standards to describe how they were used. In particular, the Minimum Information About a Simulation Experiment (MIASE) standard requires users to (i) specify and make available the exact model used, (ii) specify exactly how the model is simulated, and (iii) specify how the outputs are calculated from the model.⁷⁷

The MIASE standard is descriptive and therefore exposed to the subjectivity of the author. A more comprehensive approach is to capture (i) all the code that has been used to analyze the maps/models, (ii) all the outputs from the code, and (iii) a copy of the software used to run the code all in one place, so that this collection can be disseminated and other users can rerun the analysis and edit to experiment with the map or model. This is now possible with the advent of interactive scripting. Amongst the most prominent examples is the Jupyter Notebook in which authors can create word processor standard documents with code and code outputs embedded in the document along with a programming environment that enables the code to be executed.⁷⁸ Jupyter Notebooks originally supported the Python, R, Haskell and Ruby programming languages, but the list of supported languages has since grown considerably. MATLAB supports similar scripting with its Live scripts, though they only support MATLAB's scripting language (<https://www.mathworks.com>).

The ultimate approach to disseminating maps, models and how they have been analyzed is to take a snapshot of the computer on which they have been run and to transmit this snapshot. The snapshot contains a copy of the code and software needed to run the code, all the relevant files from the hard disk of the computer and all the relevant parts of the operating system needed to run the software. The Docker platform achieves this by creating minimal virtual machines called containers that hold everything needed to run the analysis.⁷⁹ These container files can be distributed to other users who can run them on their computer using the Docker engine. Because the containers hold everything needed to run the analysis, other users do not need any compatible software beyond the engine. This not only reduces the challenge of rerunning the analysis, but also eliminates the challenges of cross-compatibility between platforms and operating systems.

Therefore, one of the most important points in systems medicine is to develop conceptual models for their integration.⁸⁰ Once a computational model of the disease pathophysiology is available, a systems medicine model allows the setting of experiments that would not otherwise be possible for logistical or ethical reasons, especially around the iterative development and refinement of new mono- or multi-drug therapies.

Embracing the challenge of data integration and validation

Several discussion articles and reviews exist on omics data integration, from the perspective of model organisms,^{81,82} including microbes⁸² and bacteria,⁸³ or from the perspective of humans,^{84–86} and host-microbe interactions.⁸⁷ The available tools and methods of integrative omics analytics are not sufficient, and they even fail to successfully integrate, let alone analyze, different levels and sources of omics data. Important lessons can be learned from smaller-scaled analysis efforts. For instance, only adding one level of complexity to GWAS, namely multiple marker interaction analysis, has been a sobering lesson.⁸⁸ It has pointed towards problems that need to be tackled in omics integration efforts as well, as they are expected to be elevated when dealing with multiple non-independent, possibly interacting, dimensions. These problems include significance assessment, heterogeneity modelling in meta-analysis to increase power, replication, validation and replication^{88,89} and are widely applicable to systems medicine modelling in general.

Traditionally, integrative analysis techniques have focused on combining evidence derived from real data combined with public database knowledge.^{90,91} The field then moved on, from exploiting the combination of private and publicly available knowledge to accelerate drug discovery,⁹² to combining multi-dimensional views in, for instance, gene mapping.⁹³ Method developers are only gradually pacing up with the vast amount of heterogeneous data sources that become available and with introducing the necessary complexities into the models.^{94,95} With omics data increasingly being collected on the same set of individuals, it becomes theoretically possible to connect different layers of cellular or molecular information (for instance in causal models⁹⁶), while combining analytics to available expert knowledge. Integrative tools for Big Data ideally combine kernel theory (to bring in notions of non-linearity), components theory (to reduce dimensionality), and graph theory (to handle dependencies and interactions in systems). From a practical point of view, it remains essential to understand the minimum requirements each analytic tool for omics integration should have for it to be able to distinguish “noise” from “signal” and to compensate for the intrinsic power deficits resulting from having relatively small numbers of individuals with huge numbers of omics measurements. The metabolomics community has recently published viewpoints and recommendations for the development of multi-omics integration in the context of systems biology.⁹⁷ “Special Issues” on omics integration (e.g., Genin and Devoto),⁹⁴ highlight remaining challenges including the integration of dependent and independent omics data sets in meta-analyses,⁹⁸ and the integration of omics with non-omics data.⁹⁹

Once omics data have been curated, Hamid *et al* have identified three general roads to travel.⁸⁴ Either the data are fused before modelling (Fig. 2-A), or the representation of each omics data sources is altered to make it more digestible before deriving an integrative solution to the problem of interest (Fig. 2-B), or each omics data set is modelled separately and results are integrated (Fig. 2-C). Although these stages of integration are often discussed in the framework of association analyses, they also apply to prediction and profiling (pattern recognition) contexts. Notably, most of the novel analytic approaches to integrate

multiple omics dimensions do not concern analysis on fused data (Fig. 2-A). The main reason is that such analyses imply quite many information technology (IT)-infrastructure and computational challenges as well as analytic challenges in that any model is believed to be a too severe over-simplification of the rich information the inter-related data potentially entails. There are different omics levels of informativity and errors to account for, as well as for different measurement types and patterns of missingness. Regardless, fusing data before analysis seems to be the only natural way to fully account for non-independence between omics data records and to analyze data according to the spirits of systems biology. Does this mean that omics integrative analysis will be hopeless for many years to come? We do not think so... let us be creative!

Out-of-the-box thinking

One of the ways to overcome the computational burden and analytic complexities described above is to re-define the boundaries of the system we wish to elucidate. Taking the example of gene mapping, we can take a “gene” as a mini-system (Fig. 3) and combine principles of data fusion (Fig. 2-A) with ideas to change omics data representation (Fig. 2-B). In particular, we first capture the relationship between a meaningful set of omics features (Fig. 3-A) and then change the representation of that set (Fig. 3-B) while converting it into a single aggregated feature (i.e., a multidimensional module). Structure within each set can be modelled *via* prior knowledge or analytically on the observed data, such as *via* partial least square (PLS)-based path modelling that offers more possibilities than classic principal components analysis.¹⁰⁰ Such a strategy can be applied to any meaningful “unit of analysis” with characterizing features that can be represented as a network. From our perspective, the most promising strategies within an omics integration process are based on components-based association modeling,¹⁰¹ diffusion kernels on graphs for prediction,^{102,103} and similarity network fusion profiling.^{104,105}

Validation

One critical step of modelling approaches is validation, as fitting a model to data does not prove that it will accurately capture the clinical outcome. In particular, the high dimensionality of datasets is a major challenge in data analysis, especially for model reliability, as it is prone to overfitting. Therefore, there is a special need to develop dedicated protocols for validation of integrative (systemic) analyses. This effort requires tools that enable simulating realistic and sufficiently complex data. Consequently, simulation-oriented approaches have been increasingly applied over the past 5 years. Different alternatives exist in terms of validation strategy:

- *In silico* data generation: As an example, a multi-omics data simulator for complex disease studies was developed and applied to evaluate multi-omics data analysis methods for disease classification.¹⁰⁶ Another tool, iOmicsPASS, allowing network-based integration of multi-omics data for predictive subnetwork discovery was recently published.¹⁰⁷
- Validation protocols and the interpretation of validation studies: In contrast to replication, validation in other samples does not require sampling from the same populations as the discovery study. This

poses particular challenges towards interpreting the results from a validation study due to sample heterogeneity. Especially when thousands of features from heterogeneous data types are being collected, the problem of heterogeneity between individuals - assessed *via* the collected data - may become more pronounced.

- Preclinical validation suffers from the limited interpretability of *in vitro* cellular or *in vivo* animal models. Currently, we do not know for most diseases the underlying mechanism, which makes it close to impossible to decide whether the animal model that mimics a human disease symptom is due to the same mechanism. Once we know the human mechanism, there will be almost no need for an animal experiment, alas a drug repurposing study requires this, e.g. for regulatory reasons.

- Clinical trials have to be viewed as ultimate validation and with the shift from imprecise symptom-based disease definitions and symptom-based therapies, we will be able to design much smaller highly precise mechanism-based interventions with small numbers needed to treat, up to n-1-trials.

Defining the network and systems medicine framework is now allowing us to disclose some current and future clinical applications. In the following section, we elaborate on how systems medicine is being implemented on the field and in the real world.

Clinical applications (on the horizon)

Systems medicine is starting to be greatly used in the context of cancer but also in pharmacology. It has opened the door to advanced personalized medicine in these areas, improving the clinical approaches.

Cancer pathways and personalized therapy

Recent personalized therapeutic approaches in oncology target multiple pathways within a mechanistically defined cancer type by combining several drugs with the aim to cure or at least significantly improve survival and quality of life beyond current symptomatic or cytotoxic approaches. In this context, the increasing availability of pathway knowledge relevant for human systems modelling, for instance from databases including Signor¹⁰⁸ and Reactome¹⁰⁹, provides quite extensive information for building cellular signaling networks that allow the analysis of cancer cell function. The conversion of such a cell fate decision Prior Knowledge Network to a Boolean model is in practice a relatively trivial task, starting with the use of the causal interaction information to generate the logical rules that define mathematically the interactions of the network as a whole. An accurately designed logical model of a cell will follow these rules to arrive at a stable state in which the activities of the model components will quite accurately represent the activities of their biological counterparts in the cell that is represented. Logical

models built using Signor data, complemented with some additional *ad hoc* literature curation, has allowed for instance the assembly of several versions of a colorectal cancer model that have quite significant predictive power in assessing the effect of combinations of targeted drugs (^{109,110}, <https://github.com/druglogics/cascade>) on cellular states, and can be used to identify potential synergistic drugs that together are more effective for inhibiting cell proliferation than separately. The procedure to do this is: a general logical model is configured to represent a specific cancer cell line using baseline biomarker data that informs the logical model about the activity states of Boolean network nodes (Active = 1, Non-Active = 0), and the resulting cell line specific model can be used to filter out *in silico* the combinations that are least likely to display synergy. The remaining potential synergistic drug pairs can subsequently be tested in cancer cell line cultures to validate the synergy predictions.¹¹¹ From this proven system the next challenge is to implement it in a clinical setting, and develop patient-specific logical models using biomarker data from tumor biopsies obtained from a cancer patient, use these to select potential synergistic drug pairs and test these on *in vitro* cultured spheroids or organoids derived from the same tumor material (Fig. 4, see also).^{112,113} The timeline needed to perform such an analysis would be a matter of weeks, during which the patient would receive standard postoperative chemotherapy.^{108,109}

Personalized drug therapy

Drug Dosing

Precision medicine will also be oriented to personalized drug dosing to improve their efficacy and safety. In this roadmap, two important problems need to be addressed. On one hand, it is necessary to design personalized drugs for each disease and patient. New opportunities are arising with the arrival of technologies that allow the printing of 3D drugs. Thus, new improvements such as personalized dosages, chewable pills, multi-active pills and fast-dissolving tablets have been proposed. The second problem in the improvement of treatment efficiency is related to the administration of drugs to the patient. In general, very low differentiation in dosage is done for the same disease in similar patients. However, even individuals that share similar characteristics have different responses to drug administration (inter-patient variability). Also, for the same patient, the response to drug administration can vary with time (intra-patient variability). This can occur due to changes in the patient condition during the treatment.

Systems medicine arises as an approach that can help in the personalization of drug dosing. One of the key ideas in the development of personalized drug dosing mechanisms is the concept of closed-loop or feedback control. This concept plays an important role both in engineering applications and in natural systems. The main idea behind it is the observation of the output variable (variable of interest) to decide how to modify the input that is applied, to change the value of the output variable.

One of the most relevant medicine areas where feedback control systems are being applied is anesthesiology. Three main variables are involved in general anesthesia: hypnosis, analgesia and neuromuscular blockade. For each of these variables, a different drug is applied. The anesthesiologist

needs to estimate the correct dose for each of these variables. In traditional clinical practice drug dosing is according to patient characteristics (body mass index, age, gender and height). During the process, anesthesiologist corrects the drug dose according to the patient response to drugs.

If more accurate and safer drug dosing is desired, closed-loop control appears as the best option to be considered. The design of advanced control systems in anesthesiology involves methodologies included in systems medicine approach. In particular, three main challenges need to be addressed:

- **Effect assessment:** The first step towards personalizing drug infusion is the availability of an index that correlates well with the variable of interest. For general anesthesia, different measures have been proposed to measure the unconscious level of the patient:^{114,115} Bispectral index (Medtronic), Spectral Entropy (Datex-Ohmeda), Narcotrend index (Monitortechnik), Patient State Index (Masimo) or Auditory Evoked Potentials index (Danmeter). For neuromuscular blockade monitoring there are also reliable techniques, most of them based on Train-of-four stimulation.¹¹⁶ However, one of the current challenging issues is the proposal of a reliable index to assess analgesia level in patients. This is a much-complicated problem due to the complexity of the involved mechanisms and the disturbances affecting the process. Current monitoring devices for analgesia focus only on one or two variables (Electroencephalogram signal, Electrocardiogram signal, respiratory frequency, pupil diameter, mean arterial pressure, photoplethysmographic signal, etc.) to generate a pain measurement for the patient.^{116,117} It seems that a more general focus should be considered at this point. Thus, systems medicine proposes new approaches based on the development of new indexes for drug effect assessment based on the integration of multiple sources of information. This could lead to more reliable indexes that can be used to implement efficient and safe feedback control systems.
- **Modelling of patient response:** the improvement in the design and personalized titration for drugs greatly depends on the availability of reliable models. The aim is to be able to predict patient response and use this information to design personalized drugs and to administer them. Different methodologies can be used to model patient response. Thus, main methods for this are physiological models (built in the basis of physiology, anatomy and biochemistry of the body), compartmental models (based on the assumption that the body can be represented as a set of interconnected compartments¹¹⁸ and black box models (representations of the functional relationships between system inputs and system outputs). Compartmental models are much simpler than physiological ones and have been intensively used in practice. In last year's many studies have been done using black box models, based mainly on neural networks, fuzzy logic, evolutionary computation and ML^{119,120}
- **Drug infusion control systems:** the third great challenge for personalized drug dosing is the design of efficient controllers to decide the correct drug dose that the patient needs. In the field of general anesthesia, three main possibilities can be found. The first option is the signal-based controllers. These strategies are mostly based on proportional integral derivative controllers. The

algorithm decides the drug dose according to the measured errors observed. The performance of these methods is satisfactory although they have the inherent limitation of using only information of the history of the patient. Alternatively, model-based controllers predict the response of the patient (using any of the methods described above) and compute the solution that optimizes the response of the patient.¹²¹ These controllers, also known as predictive controllers, greatly depend on the reliability of the prediction model. A third option is intelligent controllers that include all those methods based on artificial intelligence techniques. It is common to find applications using fuzzy logic control that are based on heuristic rules.^{121,122} These methods allow translating directly the expertise of the anesthesiologist to the computer. Besides, applications based on neural networks can be found.¹²³ ML techniques are also being used for the design of computer aided decision system for the anesthesiologists.¹²⁴

Current research in anesthesia control systems is focused on three main issues. The first is the design of a control system with robust capabilities to reject disturbances occurring in the operating room. This is of great importance as the patient is affected by many stimuli that are changing during the surgery. The second important issue is the study of the interaction effect of different drugs. If an optimal drug dose is desired, it is necessary to study the whole patient system and his response to the different drugs that are being administered. A third important problem during surgery is related to the changes in the response to drug infusion between different patients (inter-patient variability). This means that the controller must be able to offer a satisfactory response regardless of the patient profile. The problem also occurs for a given patient during the surgery, as his/her response to drug infusion changes with time (intra-patient variability). The solution is the inclusion of adaptive systems in the closed-loop system. The controller should be able to adapt to the observed patient response. The complexity of this problem makes it necessary to use systems medicine approaches to help not only in the description and prediction of patient responses but also concerning the design of robust and efficient controllers.

This discussion about key issues in anesthesiology and the need for new approaches based on systems medicine opens up new perspectives for future research. It is important to note that most of the concepts explained for general anesthesia drug infusion can be extended to any other discipline in medicine.

2D- and 3D drug printing

Network and systems medicine will not only enable individual and genetic diagnoses but also precisely designed therapies - currently mainly pharmacotherapy. Individual pharmacotherapy is not entirely new, e.g., has always been individual in infusion therapy. Patients suffering from a tumor, for example, receive a therapy tailored to their needs, in which the strength or dose is precisely matched to the patient. Type 1 diabetics inject exactly as much insulin as they need in the respective situation. In drug therapy with tablets or capsules, individualization stops. Patients receive the active ingredient or combinations of active ingredients available on the market, each in a standardized strength or dose. These active strengths and

doses are determined in phase I-III studies that do not represent the collective in which they will later be used. The one-size-fits-all philosophy dominates therapy with oral drugs and thus prevents possible individual pharmacotherapy. This is particularly problematic in patients with impaired organ function (e.g., kidneys) or with active substances with a narrow therapeutic window. In addition, various drugs are metabolized hepatically. This varies greatly between individuals and to a clinically relevant extent (e.g., Clopidogrel, Tamoxifen). Even the introduction and elimination of therapies is currently only sub-optimally possible. With the help of digital printing technology, it is now possible to print the active substances required by a patient in exactly the strength or dose and in the needed combination. Patients, in particular those who have to take several tablets, will benefit because they will no longer have to take several tablets but only one in the future, which will contain all the active ingredients in the corresponding strength and release kinetics. 3D printing technology¹²⁵ uses digitally controlled devices for formulating active pharmaceutical ingredient (API) and excipients in a layer-by-layer pattern for developing a suitable personalized drug delivery system as per the need of the patient. It includes various techniques like inkjet printing (2D printing), fused deposition modelling (3D printing), which can further be classified into continuous inkjet system and drop on demand. To formulate such dosage forms, scientists have used various polymers to enhance their acceptance as well as therapeutic efficacy. Polymers like polyvinyl alcohol, poly (lactic acid), poly (caprolactone) etc. can be used during manufacturing. A varying number of dosage forms can be produced using 3D printing technology including immediate-release tablets, pulsatile release tablets, and transdermal dosage forms etc. Thus, it is foreseeable that in a few years, patients will not only receive highly individualized diagnoses but also personalized, precision therapies, which by combining APIs into one printed tablet will also increase compliance. With the help of digital printing technology, it will be possible to perform compounding with oral drug therapy. Ultimately, this technology will lead to higher efficiency and effectiveness and consequently significantly reduce overall therapy costs.

Drug repurposing

To have an impact and induce the necessary changes in our approach to medicine, network and systems medicine needs to provide clinical evidence. If this would involve new targets and depend on drug discovery and drug development, the proof-of-concept for network and systems medicine would take at least another 15-20 years. This gap can, however, be overcome by drug repurposing, i.e. the reuse of a registered drug for a new indication. By repurposing a registered (set of) drug(s) for a new indication, nearly the complete lead optimization and most or all of the clinical phase I is eliminated. Ideally, immediate phase II clinical trialing is ethically possible and medically justified, provided solid pre-clinical evidence on the target and drug can be provided. Compared to having to start from lead discovery, the net gain in time is at least 9 years on average. Depending on the indication (acute or chronic) and resulting trial length, the gain may be even more. This process is not new but has so far rather been serendipitous and projects like the EU funded Horizon 2020 project REPO-TRIAL (repo-trial.eu) takes this to another level and makes it more predictable and precise. REPO-TRIAL, a 5-year project, focuses on indications that allow

short-duration trials, either because the patient-relevant outcomes can be observed within days or weeks (stroke, myocardial infarction, resistant hypertension) or predictive biomarkers are available (diastolic heart failure, gout). Ultra-short or short trials are increasingly common and acceptable from a regulatory point-of-view, in particular for the most likely initial phase II, safety phase, with efficacy parameters rather being secondary outcomes. In REPO-TRIAL, a cluster of comorbid disease phenotypes has been associated with dysregulated reactive oxygen and cyclic guanosine monophosphate signaling. Patients are stratified based on biomarkers indicating this dysregulation and then treated with repurposed registered drugs that target these signaling pathways. The first trials on stroke (REPO-STROKE) and heart failure with preserved ejection fraction (REPO-HFPEF) are expected to be finalized in 2021 and 2022 respectively. With 2,538 approved small molecule drugs (Drugbank), the likelihood is high that for any given causal network at least one drug would be available. Indeed, this is the case for most targets. A fascinating recent observation, based on the PISCES dataset, is that registered drugs bind with high affinity to conserved binding pockets of, on average, 39 proteins.^{126,127} Thus, small molecule drugs are highly promiscuous and, in all likelihood, can be repurposed from one to many other target proteins with similar binding pockets. Repurposing registered drugs with known safety profiles may be so powerful that it may rapidly address therapeutic needs in many other causal disease pathways and thus outcompete classical drug discovery. Moreover, drug repurposing has occurred before, but mostly in a serendipitous manner; with network medicine this will become highly predictable, pathway by pathway.

Enlarging network and systems medicine applications

Improving patient engagement and treatment adherence

Network and systems medicine applications are further enlarged by improving patient engagement and treating adherence. Day-to-day healthcare services are not based on genotyping but rather on phenotyping. How to treat a patient is generally based on a physical examination and understanding a patient's behavior. Accordingly, improving patient engagement and treatment adherence strongly relates to the concept of the exposome, which, among others, deals with the complexity of patient-caregivers' interactions and other environmental such as the sociological and economic factors.

In recent years, the number of channels allowing healthcare customers and practitioners (a.k.a. providers) to communicate has grown dramatically. These channels are one dimension of the exposome, allowing measuring the strength of the interactions between the healthcare systems actors. Historically, patients and providers used face-to-face meetings as a standard means of communication. When phones were added to the healthcare organization arsenal, they were used for scheduling appointments or asking for services such as prescription renewals or medical recommendations.^{128–137} Since the beginning of the 2010's, the advent of the Internet, and the popularization of smartphones and Social Media, the rules of

communication between healthcare customers and healthcare providers have profoundly changed.^{138,139} This digital revolution is also allowing the healthcare system to integrate new tools supporting teleconsultation and tele-diagnostic systems, and to continuously develop and integrate innovating tools for both patients and healthcare providers.^{140,141} The main purpose of a large number of communication channels available today is to provide new ways to search, get, and share health-related information and knowledge. Nevertheless, the level by which healthcare customers and practitioners used these channels depends on numerous environmental factors such as economics, culture and regulations. The interactions between healthcare customers and practitioners must, therefore, be tracked and integrated as part of the system medicine data, as exposed-generated data, to provide an overall understanding of the patient so that the treatment and the educational and therapeutic messages are delivered to each patient in the most suitable way.¹⁴² Consider the following real-life example: HMOs record medical data and their interactions with their insured healthcare customers. One way by which patients' engagement and treatment adherence can be improved is based on the identification of subpopulations of patients by considering their communication usages and then characterizing each one with socio-demographic and bio-clinical data for improving treatment effectiveness and treatment adherence. This approach has been implemented, in 2015, on 309,460 patients with diabetes and 7 dominant profiles have been discovered and characterized to help healthcare decision-makers to improve follow-up policies and tools. Personalized services focusing on patients' needs and preferences were implemented based on this analysis.^{140,141}

Altogether, to increase the frequency of successful translational stories, the research enterprise needs to re-design research studies by considering the complexity and variability of human physiology, and by collecting high-dimensional datasets that will allow researchers to identify confounding variables and to stratify populations at early phases of biomarker discovery. As the "omics" term is expanding to wider systems, all of these have to be interrelated.

One of the remaining challenges of healthcare systems is patient accessibility. One way to improve this is to find the most suitable communication channel(s) to interact with a patient based on his/her profile, which combines socio-demographics, clinical, biological, and therapeutics data over time. This approach induces, at least for part of the population, proactive behavior and engagement in follow-up and treatment when relevant.^{140,141} HMOs around the world are developing digital services, such as online counselling services, which integrate video-call to physicians when the clinics are closed. This kind of consultation is based on the overall patient's data shared over the electronic medical record, thus allowing any healthcare practitioner to have a clear view of the patient anamnesis and therefore delivering low biased recommendations and treatment. Sharing data is an essential part of developing and delivering personalized medicine. As an example, when searching for patterns of interactions of patients with diabetes an Israeli HMO allowed pointing-out the need to tune its communication tools and messages to

patients, more particularly to those with special needs, such as elderly people, immigrants and minorities, who are not fluent in the local language, and those with low socioeconomic status. Matching a communication tool and message to the patient will improve patient's accessibility to HMO services, generate a better patient engagement and responsiveness to treatment, and improve the quality of treatment and treatment experience within existing budgetary constraints. Particularly, for patients with diabetes communication is a key dimension of systems medicine, which will provide an opportunity, for example, to collect more Patient-Reported Outcome Measures¹⁴³ for some basic follow-up measurements such as glycaemia values, weight (for computing body mass index), and smoking status.

Training

There is a consensus that systems medicine-specific training is a need, recognized by trainers, students, but also by the authorities. A major challenge of today's medicine is the ability to integrate the technological revolution, expansion of data collection which comes in multiple formats and is stored in different computers at different clinical sites, into the coordinated everyday clinical practice. Many of us believe that one, or even two generations of new medical doctors (MDs) and researchers might be needed for this to be achieved. We also believe that a society must educate their new generations on the data and technology revolution in medicine. The younger generations are already sensitized to comprehend and adapt to these changes due to their experience dealing with new technologies (smartphones and other gadgets, social media, etc.)¹⁴⁴ While the medical community is becoming increasingly aware of these educational needs, the how (and when) to introduce these new subjects is not so obvious. One view is to apply systems biology approaches and tools to biomedical problems, and to start educating biomedical students in an interdisciplinary manner as early as possible. In addition, these educational efforts have to take into account ethical concerns as well as economic circumstances and specific aspects of the different health care systems. Therefore, a joint effort of all (bio-)medical education and health care delivery stakeholders is key in this process. It seems evident that systems medicine training of future physicians cannot include a deep education in programming, mathematical modelling or computational sciences - given the wealth of medical information that has to be tackled in the course of medical studies. Instead, medical students need to learn the skills of using professional software solutions that have been developed by specialists in an interdisciplinary manner, together with practicing doctors. To state it simply: a car driver does not need to know and understand in detail how the engine of a modern car works but has to know how to drive the car. Similarly, future MDs have to know how to apply systems medicine solutions that have been developed by specialists in their daily medical routine.

Undergraduate education

Despite the recognized need to change and adapt the education programs of (bio)medicine studies, there is no agreement on the best practices and ways to achieve this goal.

The reason lies in the generally fragmented approach in the European higher education system, where

even within a single country, Universities teach similar subjects by different principles and keep the decision autonomy. Universities are independent in offering novel courses; accreditation for these is requested in countries that follow the Bologna process. Some steps towards the implementation of a systems medicine education have already been tested within the FP7 CASyM Coordinated Actions Systems Medicine and later within EASyM (European Association for Systems Medicine, <https://easym.eu/>). Similarly, the International Network and Systems Medicine Association (<https://insma.net>) provides such resources at an international level and is a direct spin-off of the COST action OpenMultiMed. Several medical schools in Europe teach subjects that are relevant to systems medicine. What is missing is the combination of relevant subjects into modules that would receive the formal name of “systems medicine”. If such modules are provided mostly into elective courses, we should ensure that students receive the proper information regarding the systems medicine subjects. For example, at the Faculty of Medicine, University of Ljubljana, systems medicine topics are currently covered within computational and practical/research elective courses in (bio)informatics, mathematics and computer-supported approaches, and e-learning, while in senior years interdisciplinary courses are given in functional genomics and pharmacogenetics. At Maastricht University, the Netherlands, medical students can choose for a Network and Systems Medicine elective.

Doctoral education

Similar to the undergraduate situation, doctoral education is also dispersed in Europe. Three possibilities appear feasible in the future:

- Introducing (accredited or non-accredited) systems medicine concepts or subjects into the existing biomedical doctoral programmes.
- Introducing a novel interdisciplinary systems medicine research training networks for doctoral students within the ITV Marie Curie or similar programmes.
- Establishing a formal systems medicine doctoral programme at individual Universities. This option has not yet been tested in Europe but is active in the USA. The Georgetown University MS degree in Systems Medicine is designed for students interested in bringing systems medicine into biomedical science and clinical practice and setting the stage for bridging research and clinical care (<https://systemsmedicine.georgetown.edu/>). The MD/MS dual program is designed for students already accepted to medical school, and who will take an additional year beyond the four required for the traditional MD to complete the MS. Students will be accepted into the program after completion of their second year in medical school. The program educates physicians to understand and apply new approaches to diagnose, prevent or delay disease manifestation and improve clinical outcomes for patients. The MS and MD/MS Dual Degree programs in Systems Medicine teach students to use

cutting-edge technology to train the next generation of physicians and biomedical scientists. Students learn a new language, which is the application of -omics technology and Big Data to patient care. In addition to credits gained by courses, students carry out a semester-long Capstone internship where they gain hands-on work experience in renowned institutions and are matched with a mentor based on their career goals and interests. The Capstone project culminates in presentations or even journal articles. This model could be used as an example for future implementation in institutions across Europe and North America.

Education of medical specialists

We also need to develop training opportunities for established MDs, medical specialists, to promote timely integration of systems medicine topics into the clinical practice. This is a more demanding task, since MDs have limited time available for education and training. However, experience shows that they want to gain this knowledge once they see the benefits for their patients, such as better diagnostics and treatments. To reach this target group, a variety of lifelong education possibilities has to be offered, such as systems medicine meetings, expert guided workshops and summer schools, targeted lecture series, etc. MDs could better be attracted if the courses are accredited with the Continuous Medical Education (CME) credits, which are required for maintaining the practitioner license in several European countries.

SWOT analysis

To outline the different important issues for network and systems medicine, a SWOT analysis (Fig. 5, for Strengths, Weaknesses, Opportunities and Threats) was performed. This thorough analysis is key for the design and development of a strategic plan that would contribute to the implementation of systems medicine in a wide spectrum of clinical applications within precision healthcare. In terms of strengths, i.e. the innate advantages of systems medicine, the availability of multidimensional data and computational tools are important and solid elements for the field, as well as the input of the big pharma industry for applications. However, the lack of standardization in methods and data storage, as well as inter-individual variability and populations limit hypothesis generations and clinical applications. Nonetheless, major opportunities were identified: the development of multidisciplinary communities and collaborations will result in a rapid advance in knowledge and translation that should be strengthened with education, which will finally lead to a truly personalized healthcare system. Lastly, prompt actions will be necessary to overcome the tradition and way of thinking in medicine in order to reinforce participation of stakeholders and funding agencies.

Concluding remarks and outlook

A considerable number of obstacles still need to be overcome for more profitable and successful implementation of practical systems medicine in the clinical setting.

Among these issues, with no claim of completeness, we include the fact that a high number of scientific hypotheses can be generated *via* systems medicine methods, and that corresponding reliable testing and solid validations - essential before safe clinical practice - are still lacking. This is mainly due to the limitation of resources to test many of such hypotheses. A comprehensive validation practice should ultimately steer towards the adoption of certified, harmonized and partly machine-operated workflows and protocols, finally capable and designed to function in dynamical clinical contexts.

Another problem concerns the vast imbalance of systems biology studies that still focus on smaller biological systems, over the systems medicine ones, targeting larger systems/whole organisms. Along with this, there is also the interrelated difficulty to scale up to the whole organism tier, clearly necessary in the clinical practice, due to the intrinsic limits of the conclusions related to the narrow experimental/biological context (e.g., gene regulatory networks acting in cellular processes, signaling pathways analysis, etc., whose analyses often provide views too limited to be relevant in the clinic).

Several issues occur when collecting supporting, comprehensible, and secured (another ‘buzzword’ itself, nowadays) datasets in clinical settings. Indeed, the type, arrangement and nature of medical and clinical data have their peculiar production methods, schemas, ontologies, standards, complexity and access limits, which often conflict with the requirement and the complications to tie unambiguously such clinical data with the clinical samples.

An often-overlooked consideration resides in the circumstance that *in silico* methods, models and research outcomes should not be excessively complicated or abstruse to medical doctors, personnel and policymakers. About this point, it may be relevant to refer here to the problem of AI, i.e. machine and deep learning approaches, used to perform predictive analyses in the clinical practice. It is well known that AI algorithms often work and produce results as “black box”, i.e. for its nature, it conceals the relationship and the importance of a set of data features from the output, which should finally represent the biological/medical interpretation. This is mainly because such correlations are multi-dimensional and not reducible enough to be easily grasped by the human mind (which is exactly why AI is used). This peculiarity of AI methods can often hamper or delay the deployment of predictive models because humans simply do not understand, and thus trust them.¹⁴⁵ Relevant efforts are being made to overcome this issue by providing interpretable ML methodologies capable of balancing accuracy, human interpretability and, last but not least, computational viability.^{146,147}

As a final consideration, it is now clear that the structure of the healthcare system will have to adjust radically to be able to run with highly interdisciplinary crews, operating daily with multi-omics, multi-source data, large-scale databases and storage facilities, complex analytical processes, and -clearly- effective managerial and organizational frameworks. Such practices call for tailored education programs and continuous, complimentary training for hospital personnel as well as for systemic scientists.^{36,148,149}

Acknowledgments

This article is based upon work from COST Action OpenMultiMed (CA15120), supported by COST (European Cooperation in Science and Technology). This review is also supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777111 (REPO-TRIAL). This reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

1. Loscalzo J. Personalized Cardiovascular Medicine and Drug Development. *Circulation* 2012;125:638–645.
2. Nosengo N. Can you teach old drugs new tricks? *Nature* 2016;534:314–316.
3. Wieseler B, McGauran N, Kaiser T. New drugs: where did we go wrong and what can we do better? *BMJ* 2019;l4340.
4. Schork NJ. Personalized medicine: Time for one-person trials. *Nature* 2015;520:609–611.
5. Scannell JW, Blanckley A, Boldon H, et al. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery* 2012;11:191–200.
6. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nature reviews. Drug discovery* 2011;10:712.
7. Kleinschnitz C, Mencl S, Kleikers PWM, et al. NOS knockout or inhibition but not disrupting PSD-95-NOS interaction protect against ischemic brain damage. *Journal of Cerebral Blood Flow & Metabolism* 2016;36:1508–1512.
8. Kleikers PWM, Hooijmans C, Göb E, et al. A combined pre-clinical meta-analysis and randomized confirmatory trial approach to improve data validity for therapeutic target validation. *Scientific Reports*;5 . Epub ahead of print 2015. DOI: 10.1038/srep13428.
9. Dornas WC, Silva ME. Animal models for the study of arterial hypertension. *Journal of Biosciences* 2011;36:731–737.
10. Segal-Lieberman G, Rosenthal T. Animal Models in Obesity and Hypertension. *Current Hypertension Reports* 2013;15:190–195.
11. Kleinschnitz C, Fluri F, Schuhmann M. Animal models of ischemic stroke and their application in clinical research. *Drug Design, Development and Therapy* 2015;3445.
12. Shanks N, Greek R, Greek J. Are animal models predictive for humans? *Philosophy, Ethics, and Humanities in Medicine* 2009;4:2.
13. O'Collins VE, Macleod MR, Donnan GA, et al. 1,026 Experimental treatments in acute stroke. *Annals of Neurology* 2006;59:467–477.
14. Seok J, Shaw Warren H, Cuenca AG, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences* 2013;110:3507–3512.

15. Patrick Vallance. *Nature Reviews Drug Discovery* 2010;9:834–834.
16. Root AA, Smeeth L. NNTs and NNHs: handle with care. *British Journal of General Practice* 2017;67:133–133.
17. Contopoulos-Ioannidis DG, Ntzani EE, Ioannidis JPA. Translation of highly promising basic science research into clinical applications. *The American Journal of Medicine* 2003;114:477–484.
18. Menche J, Sharma A, Kitsak M, et al. Uncovering disease-disease relationships through the incomplete interactome. *Science* 2015;347:1257601–1257601.
19. Kiel C, Lastrucci C, Luthert PJ, et al. Simple and complex retinal dystrophies are associated with profoundly different disease networks. *Scientific Reports*;7 . Epub ahead of print 2017. DOI: 10.1038/srep41835.
20. Sanchez-Vega F, Mina M, Armenia J, et al. Abstract 3302: The molecular landscape of oncogenic signaling pathways in The Cancer Genome Atlas. *Bioinformatics and Systems Biology* . Epub ahead of print 2018. DOI: 10.1158/1538-7445.am2018-3302.
21. Casas AI, Hassan AA, Larsen SJ, et al. From single drug targets to synergistic network pharmacology in ischemic stroke. *Proceedings of the National Academy of Sciences* 2019;116:7129–7136.
22. Laifenfeld D, Drubin DA, Catlett NL, et al. Early Patient Stratification and Predictive Biomarkers in Drug Discovery and Development. *Advances in Experimental Medicine and Biology* 2012;645–653.
23. Carrigan P, Krahn T. Impact of Biomarkers on Personalized Medicine. *New Approaches to Drug Discovery* 2015;285–311.
24. Liu R, Wang X, Aihara K, et al. Early Diagnosis of Complex Diseases by Molecular Biomarkers, Network Biomarkers, and Dynamical Network Biomarkers. *Medicinal Research Reviews* 2014;34:455–478.
25. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology* 2008;4:682–690.
26. Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *N Biotechnol* 2012;29:613–624.
27. Auffray C, Chen Z, Hood L. Systems medicine: the future of medical genomics and healthcare. *Genome Med* 2009;1:2.
28. Nookaew I. *Network Biology*. Springer. 2017.
29. Ma’ayan A. Introduction to network analysis in systems biology. *Sci Signal* 2011;4:tr5.

30. Kitano H. Systems biology: a brief overview. *Science* 2002;295:1662–1664.
31. Vidal M, Cusick ME, Barabási A-L. Interactome Networks and Human Disease. *Cell* 2011;144:986–998.
32. Chuang H-Y, Lee E, Liu Y-T, et al. Network-based classification of breast cancer metastasis. *Mol Syst Biol* 2007;3:140.
33. Berlin R, Gruen R, Best J. Systems Medicine Disease: Disease Classification and Scalability Beyond Networks and Boundary Conditions. *Frontiers in Bioengineering and Biotechnology*;6 . Epub ahead of print 2018. DOI: 10.3389/fbioe.2018.00112.
34. Hansen J, Iyengar R. Computation as the mechanistic bridge between precision medicine and systems therapeutics. *Clin Pharmacol Ther* 2013;93:117–128.
35. Delhalle S, Bode SFN, Balling R, et al. A roadmap towards personalized immunology. *NPJ Syst Biol Appl* 2018;4:9.
36. Flores M, Glusman G, Brogaard K, et al. P4 medicine: how systems medicine will transform the healthcare sector and society. *Per Med* 2013;10:565–576.
37. Furman D, Campisi J, Verdin E, et al. Chronic inflammation in the etiology of disease across the life span. *Nat Med* 2019;25:1822–1832.
38. Renzi C, Kaushal A, Emery J, et al. Comorbid chronic diseases and cancer diagnosis: disease-specific effects and underlying mechanisms. *Nat Rev Clin Oncol* 2019;16:746–761.
39. Targher G, Lonardo A, Byrne CD. Nonalcoholic fatty liver disease and chronic vascular complications of diabetes mellitus. *Nat Rev Endocrinol* 2018;14:99–114.
40. Wafi A, Mirnezami R. Translational –omics: Future potential and current challenges in precision medicine. *Methods* 2018;151:3–11.
41. Naik A, Košir R, Rozman D. Genomic aspects of NAFLD pathogenesis. *Genomics* 2013;102:84–95.
42. Lorbek G, Perše M, Jeruc J, et al. Lessons from hepatocyte-specific Cyp51 knockout mice: impaired cholesterol synthesis leads to oval cell-driven liver injury. *Sci Rep* 2015;5:8777.
43. Cvitanović Tomaš T, Urlep Ž, Moškon M, et al. LiverSex Computational Model: Sexual Aspects in Hepatic Metabolism and Abnormalities. *Front Physiol* 2018;9:360.
44. Micheel CM, Nass SJ, Omenn GS, et al. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials; Board on Health Care Services; Board on Health Sciences Policy;

Institute of Medicine. Evolution of translational omics: Lessons learned and the path forward. Evolution of translational omics: lessons learned and the path forward Washington, DC: National Academies Press.

45. Nicholson JK, Lindon JC, Holmes E. "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999;29:1181–1189.
46. Fiehn O, Kopka J, Dörmann P, et al. Metabolite profiling for plant functional genomics. *Nat Biotechnol* 2000;18:1157–1161.
47. Mamas M, Dunn WB, Neyses L, et al. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Arch Toxicol* 2011;85:5–17.
48. Lindon JC, Nicholson JK. The emergent role of metabolic phenotyping in dynamic patient stratification. *Expert Opin Drug Metab Toxicol* 2014;10:915–919.
49. Zhang A-H, Qiu S, Xu H-Y, et al. Metabolomics in diabetes. *Clin Chim Acta* 2014;429:106–110.
50. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24–32.
51. Shaffer M, Armstrong AJS, Phelan VV, et al. Microbiome and metabolome data integration provides insight into health and disease. *Transl Res* 2017;189:51–64.
52. Shoaie S, Nielsen J. Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Front Genet* 2014;5:86.
53. Greenblum S, Chiu H-C, Levy R, et al. Towards a predictive systems-level model of the human microbiome: progress, challenges, and opportunities. *Curr Opin Biotechnol* 2013;24:810–820.
54. Hood L, Heath JR, Phelps ME, et al. Systems biology and new technologies enable predictive and preventative medicine. *Science* 2004;306:640–643.
55. Galetsi P, Katsaliaki K, Kumar S. Values, challenges and future directions of big data analytics in healthcare: A systematic review. *Soc Sci Med* 2019;241:112533.
56. Lænkholm A-V, Jensen M-B, Eriksen JO, et al. PAM50 Risk of Recurrence Score Predicts 10-Year Distant Recurrence in a Comprehensive Danish Cohort of Postmenopausal Women Allocated to 5 Years of Endocrine Therapy for Hormone Receptor–Positive Early Breast Cancer. *Journal of Clinical Oncology* 2018;36:735–740.
57. Slodkowska EA, Ross JS. MammaPrint™ 70-gene signature: another milestone in personalized medical care for breast cancer patients. *Expert Review of Molecular Diagnostics* 2009;9:417–422.

58. Bösl A, Spitzmüller A, Jasarevic Z, et al. MammaPrint versus EndoPredict: Poor correlation in disease recurrence risk classification of hormone receptor positive breast cancer. *PLoS One* 2017;12:e0183458.
59. Walter VP, Taran F-A, Wallwiener M, et al. A high-risk 70-gene signature is not associated with the detection of tumor cell dissemination to the bone marrow. *Breast Cancer Res Treat* 2018;169:305–309.
60. Weinstein JN, The Cancer Genome Atlas Research Network, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics* 2013;45:1113–1120.
61. Zhang J, Baran J, Cros A, et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. *Database* 2011;2011:bar026–bar026.
62. International Agency for Research on Cancer. International Agency for Research on Cancer Biennial Report 2004-2005. Diamond Pocket Books (P) Ltd. 2006.
63. Kühling J. Datenschutz im Gesundheitswesen. *Medizinrecht* 2019;37:611–622.
64. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018.
65. Sattler F, Wiedemann S, Muller K-R, et al. Robust and Communication-Efficient Federated Learning From Non-i.i.d. Data. *IEEE Transactions on Neural Networks and Learning Systems* 2019;1–14.
66. Fang J, Fu H, Yang G, et al. RedSync: Reducing synchronization bandwidth for distributed deep learning training system. *Journal of Parallel and Distributed Computing* 2019;133:30–39.
67. Schmidt M, Schmidt SAJ, Sandegaard JL, et al. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol* 2015;7:449–490.
68. Schmidt M, Schmidt SAJ, Adelborg K, et al. The Danish health care system and epidemiological research: from health care contacts to database records. *Clinical Epidemiology* 2019;11:563–591.
69. Jaffe DH, Flaks-Manov N, Benis A, et al. Population-based cohort of 500 patients with Gaucher disease in Israel. *BMJ Open* 2019;9:e024251.
70. Boeckhout M, Zielhuis GA, Bredenoord AL. The FAIR guiding principles for data stewardship: fair enough? *Eur J Hum Genet* 2018;26:931–936.
71. Wittig U, Rey M, Weidemann A, et al. Data management and data enrichment for systems biology projects. *Journal of Biotechnology* 2017;261:229–237.

72. Le Novère N, Hucka M, Mi H, et al. The Systems Biology Graphical Notation. *Nat Biotechnol* 2009;27:735–741.
73. Benson HE, Watterson S, Sharman JL, et al. Is systems pharmacology ready to impact upon therapy development? A study on the cholesterol biosynthesis pathway. *British Journal of Pharmacology* 2017;174:4362–4382.
74. Parton A, McGilligan V, Chemaly M, et al. New models of atherosclerosis and multi-drug therapeutic interventions. *Bioinformatics* 2019;35:2449–2457.
75. van Iersel MP, Villéger AC, Czauderna T, et al. Software support for SBGN maps: SBGN-ML and LibSBGN. *Bioinformatics* 2012;28:2016–2021.
76. Hucka M, Finney A, Sauro HM, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 2003;19:524–531.
77. Waltemath D, Adams R, Beard DA, et al. Minimum Information About a Simulation Experiment (MIASE). *PLoS Comput Biol* 2011;7:e1001122.
78. Perkel JM. Why Jupyter is data scientists' computational notebook of choice. *Nature* 2018;563:145–146.
79. Boettiger C. An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review* 2015;49:71–79.
80. Benson M. Clinical implications of omics and systems medicine: focus on predictive and individualized treatment. *J Intern Med* 2016;279:229–240.
81. Joyce AR, Palsson BØ. The model organism as a system: integrating “omics” data sets. *Nat Rev Mol Cell Biol* 2006;7:198–210.
82. Zhang W, Li F, Nie L. Integrating multiple “omics” analysis for microbial biology: application and methodologies. *Microbiology* 2010;156:287–301.
83. De Keersmaecker SCJ, Thijs IMV, Vanderleyden J, et al. Integration of omics data: how well does it work for bacteria? *Mol Microbiol* 2006;62:1239–1250.
84. Hamid JS, Hu P, Roslin NM, et al. Data integration in genetics and genomics: methods and challenges. *Hum Genomics Proteomics*;2009 . Epub ahead of print January 12, 2009. DOI: 10.4061/2009/869093.
85. Gomez-Cabrero D, Abugessaisa I, Maier D, et al. Data integration in the era of omics: current and future challenges. *BMC Syst Biol* 2014;8 Suppl 2:I1.

86. Zierer J, Menni C, Kastenmüller G, et al. Integration of “omics” data in aging research: from biomarkers to systems biology. *Aging Cell* 2015;14:933–944.
87. Kint G, Fierro C, Marchal K, et al. Integration of “omics” data: does it lead to new insights into host-microbe interactions? *Future Microbiol* 2010;5:313–328.
88. Van Steen K, Moore JH. How to increase our belief in discovered statistical interactions via large-scale association studies? *Hum Genet* 2019;138:293–305.
89. Gusareva ES, Van Steen K. Practical aspects of genome-wide association interaction analysis. *Hum Genet* 2014;133:1343–1358.
90. Figeys D. Combining different “omics” technologies to map and validate protein-protein interactions in humans. *Brief Funct Genomic Proteomic* 2004;2:357–365.
91. Wachter A, Beißbarth T. pwOmics: an R package for pathway-based integration of time-series omics data using public database knowledge. *Bioinformatics* 2015;31:3072–3074.
92. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? An introduction and overview. *Yearb Med Inform* 2001;10:83–100.
93. Chari R, Coe BP, Wedseltoft C, et al. SIGMA2: a system for the integrative genomic multi-dimensional analysis of cancer genomes, epigenomes, and transcriptomes. *BMC Bioinformatics* 2008;9:422.
94. Génin E, Devoto M. Integration of Omics Data in Genetic Epidemiology. *Hum Hered* 2015;79:109–110.
95. López de Maturana E, Pineda S, Brand A, et al. Toward the integration of Omics data in epidemiological studies: still a “long and winding road.” *Genet Epidemiol* 2016;40:558–569.
96. Cai X, Bazerque JA, Giannakis GB. Inference of gene regulatory networks with sparse structural equation models exploiting genetic perturbations. *PLoS Comput Biol* 2013;9:e1003068.
97. Pinu FR, Beale DJ, Paten AM, et al. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites*;9 . Epub ahead of print April 18, 2019. DOI: 10.3390/metabo9040076.
98. Kaefer A, Landesfeind M, Feussner K, et al. Meta-analysis of pathway enrichment: combining independent and dependent omics data sets. *PLoS One* 2014;9:e89297.
99. López de Maturana E, Alonso L, Alarcón P, et al. Challenges in the Integration of Omics and Non-Omics Data. *Genes* ;10 . Epub ahead of print March 20, 2019. DOI: 10.3390/genes10030238.

100. Fouladi R. From Statistical to Biological Interactions towards an omics-integrated MB-MDR framework. PhD Doctor in Electrical Engineering and Computer Science. Université de Liège, Liège, Belgique Available from: <https://orbi.uliege.be/handle/2268/228579>. 6-Sep-2018.
101. Tenenhaus M, Vinzi VE, Chatelin Y-M, et al. PLS path modeling. *Comput Stat Data Anal* 2005;48:159–205.
102. Risi Imre Kondor JL. Diffusion kernels on graphs and other discrete structures. In: In Proceedings of the ICML Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.128.4966>. 2002. Accessed February 24, 2020.
103. Morota G, Koyama M, Rosa GJM, et al. Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. *Genet Sel Evol* 2013;45:17.
104. Zhang Y, Hu X, Jiang X. Multi-View Clustering of Microbiome Samples by Robust Similarity Network Fusion and Spectral Clustering. *IEEE/ACM Trans Comput Biol Bioinform* 2017;14:264–271.
105. Wang B, Mezlini AM, Demir F, et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014;11:333–337.
106. Chung R-H, Kang C-Y. A multi-omics data simulator for complex disease studies and its application to evaluate multi-omics data analysis methods for disease classification. *Gigascience*;8 . Epub ahead of print May 1, 2019. DOI: 10.1093/gigascience/giz045.
107. Koh HWL, Fermin D, Vogel C, et al. iOmicsPASS: network-based integration of multiomics data for predictive subnetwork discovery. *NPJ Syst Biol Appl* 2019;5:22.
108. Licata L, Lo Surdo P, Iannuccelli M, et al. SIGNOR 2.0, the SIGNaling Network Open Resource 2.0: 2019 update. *Nucleic Acids Res* 2020;48:D504–D510.
109. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. *Nucleic Acids Res* 2020;48:D498–D503.
110. Flobak Å, Baudot A, Remy E, et al. Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling. *PLoS Comput Biol* 2015;11:e1004426.
111. Flobak Å, Niederdorfer B, Nakstad VT, et al. A high-throughput drug combination screen of targeted small molecule inhibitors in cancer cell lines. *Sci Data* 2019;6:237.
112. Ooft SN, Weeber F, Dijkstra KK, et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*;11 . Epub ahead of print October 9, 2019. DOI: 10.1126/scitranslmed.aay2574.

113. Weeber F, Ooft SN, Dijkstra KK, et al. Tumor Organoids as a Pre-clinical Cancer Model for Drug Discovery. *Cell Chem Biol* 2017;24:1092–1100.
114. Musizza B, Ribaric S. Monitoring the depth of anaesthesia. *Sensors* 2010;10:10896–10935.
115. Reboso JA, Gonzalez-Cava JM, León A, et al. Closed loop administration of propofol based on a Smith predictor: a randomized controlled trial. *Minerva Anesthesiol* 2019;85:585–593.
116. Hund HC, Rice MJ, Ehrenfeld J. An Evaluation of the State of Neuromuscular Blockade Monitoring Devices. *J Med Syst* 2016;40:281.
117. Cowen R, Stasiowska MK, Laycock H, et al. Assessing pain objectively: the use of physiological markers. *Anaesthesia* 2015;70:828–847.
118. Martín-Mateos I, Méndez Pérez JA, Reboso Morales JA, et al. Adaptive pharmacokinetic and pharmacodynamic modelling to predict propofol effect using BIS-guided anesthesia. *Computers in Biology and Medicine* 2016;75:173–180.
119. Mathis MR, Kheterpal S, Najarian K. Artificial Intelligence for Anesthesia. *Anesthesiology* 2018;129:619–622.
120. Marrero A, Méndez JA, Reboso JA, et al. Adaptive fuzzy modeling of the hypnotic process in anesthesia. *J Clin Monit Comput* 2017;31:319–330.
121. Eskandari N, van Heusden K, Dumont GA. Extended habituating model predictive control of propofol and remifentanil anesthesia. *Biomedical Signal Processing and Control* 2020;55:101656.
122. Moore BL, Quasny TM, Doufas AG. Reinforcement Learning Versus Proportional–Integral–Derivative Control of Hypnosis in a Simulated Intraoperative Patient. *Anesthesia & Analgesia* 2011;112:350–359.
123. Eshghi N, Aliyari M, Teshnehlal M. Anesthesia Control Based on Intelligent Controllers. 2009 3rd International Conference on Bioinformatics and Biomedical Engineering . Epub ahead of print 2009. DOI: 10.1109/icbbe.2009.5162370.
124. Gonzalez-Cava JM, Reboso JA, Casteleiro-Roca JL, et al. A Novel Fuzzy Algorithm to Introduce New Variables in the Drug Supply Decision-Making Process in Medicine. *Complexity* 2018;2018:1–15.
125. Afsana, Jain V, Haider N, et al. 3D Printing in Personalized Drug Delivery. *Curr Pharm Des* 2018;24:5062–5071.
126. Wang G, Dunbrack RL. PISCES: a protein sequence culling server. *Bioinformatics* 2003;19:1589–1591.

127. Chartier M, Morency L-P, Zylber MI, et al. Large-scale detection of drug off-targets: hypotheses for drug repurposing and understanding side-effects. *BMC Pharmacology and Toxicology*;18 . Epub ahead of print 2017. DOI: 10.1186/s40360-017-0128-7.
128. Axén I, Bodin L, Bergström G, et al. Clustering patients on the basis of their individual course of low back pain over a six month period. *BMC Musculoskeletal Disorders*;12 . Epub ahead of print 2011. DOI: 10.1186/1471-2474-12-99.
129. Rai A, Chen L, Pye J, et al. Understanding determinants of consumer mobile health usage intentions, assimilation, and channel preferences. *J Med Internet Res* 2013;15:e149.
130. Hoffman AS, Volk RJ, Saarimaki A, et al. Delivering patient decision aids on the Internet: definitions, theories, current evidence, and emerging research areas. *BMC Medical Informatics and Decision Making*;13 . Epub ahead of print 2013. DOI: 10.1186/1472-6947-13-s2-s13.
131. Beck F, Richard J-B, Nguyen-Thanh V, et al. Use of the Internet as a Health Information Resource Among French Young Adults: Results From a Nationally Representative Survey. *Journal of Medical Internet Research* 2014;16:e128.
132. Moick M, Terlutter R. Physicians' Motives for Professional Internet Use and Differences in Attitudes Toward the Internet-Informed Patient, Physician–Patient Communication, and Prescribing Behavior. *Medicine 2.0* 2012;1:e2.
133. Kritz M, Gschwandtner M, Stefanov V, et al. Utilization and Perceived Problems of Online Medical Resources and Search Tools Among Different Groups of European Physicians. *Journal of Medical Internet Research* 2013;15:e122.
134. Dugdale DC, Epstein R, Pantilat SZ. Time and the patient-physician relationship. *Journal of General Internal Medicine* 1999;14:S34–S40.
135. Weiner JP. Doctor-patient communication in the e-health era. *Israel Journal of Health Policy Research*;1 . Epub ahead of print 2012. DOI: 10.1186/2045-4015-1-33.
136. Peleg R, Avdalimov A, Freud T. Providing cell phone numbers and email addresses to Patients: the physician's perspective. *BMC Research Notes*;4 . Epub ahead of print 2011. DOI: 10.1186/1756-0500-4-76.
137. Peleg R, Nazarenko E. Providing cell phone numbers and e-mail addresses to patients: The patient's perspective, a cross sectional study. *Israel Journal of Health Policy Research*;1 . Epub ahead of print 2012. DOI: 10.1186/2045-4015-1-32.
138. Henao R, Murray J, Ginsburg G, et al. Patient clustering with uncoded text in electronic medical

- records. AMIA Annu Symp Proc 2013;2013:592–599.
139. Sewitch MJ, Leffondré K, Dobkin PL. Clustering patients according to health perceptions. *Journal of Psychosomatic Research* 2004;56:323–332.
140. Benis A, Harel N, Barkan R, et al. Identification and Description of Healthcare Customer Communication Patterns Among Individuals with Diabetes in Clalit Health Services: A Retrospective Database Study. *Stud Health Technol Inform* 2017;244:18–22.
141. Benis A, Harel N, Barkan RB, et al. Patterns of Patients’ Interactions With a Health Care Organization and Their Impacts on Health Quality Measurements: Protocol for a Retrospective Cohort Study. *JMIR Research Protocols*;7 . Epub ahead of print November 2018. DOI: 10.2196/10734.
142. Nelson EC, Dixon-Woods M, Batalden PB, et al. Patient focused registries can improve health, care, and science. *BMJ* 2016;i3319.
143. Nelson EC, Eftimovska E, Lind C, et al. Patient reported outcome measures in practice. *BMJ* 2015;350:g7818–g7818.
144. Rozman D, Acimovic J, Schmeck B. Training in Systems Approaches for the Next Generation of Life Scientists and Medical Doctors. *Methods Mol Biol* 2016;1386:73–86.
145. Camacho DM, Collins KM, Powers RK, et al. Next-Generation Machine Learning for Biological Networks. *Cell* 2018;173:1581–1592.
146. Lakkaraju H, Bach SH, Jure L. Interpretable Decision Sets: A Joint Framework for Description and Prediction. *KDD 2016*;2016:1675–1684.
147. Yu MK, Ma J, Fisher J, et al. Visible Machine Learning for Biomedicine. *Cell* 2018;173:1562–1565.
148. Hood L. Systems biology and p4 medicine: past, present, and future. *Rambam Maimonides Med J* 2013;4:e0012.
149. Noell G, Faner R, Agustí A. From systems biology to P4 medicine: applications in respiratory medicine. *Eur Respir Rev*;27 . Epub ahead of print March 31, 2018. DOI: 10.1183/16000617.0110-2017.
150. Auffray C, Balling R, Benson M, et al. From Systems Biology to Systems Medicine, European Commission, DG Research, Directorate of Health. Brussels 14-15 June 2010. In: Workshop report. 2010.

Figure legends

Figure 1. Medical knowledge gaps and the ground-breaking nature of Network and Systems Medicine.

(a) Time-course of most chronic diseases without knowing the causal mechanism. Diagnosis relies on signs and symptoms pointing to (b) specific organs. Therapy focuses on achieving patient-relevant outcomes only in (c) a small fraction (green) of patients.⁵ (d) Network and Systems Medicine aim to define and diagnose a disease mechanistically, and treat it with higher precision, based on (e) mechanism-based diagnostics and therapeutics (i.e., theranostics).

Figure 2. Traditional choices to handle different omics data sources before deriving an integrated solution.

Different colors and symbols represent different data origins. A – Data fusion, which allows accounting for structure between omics data. Evidence for such structural relationships may be derived from biological knowledge or analytically (full lines), or maybe deduced from the latter (dashed lines). B – Changing the representation of each data source. This may be based on principals of dimensionality reduction or the identification of communities (cf. corresponding data corresponding symbols with gradient fill). C – Obtaining a data-specific solution, hereby ignoring detailed inter-relationships between data sources as part of obtaining an integrative solution. Once data are represented as in B, cross-data source relationships may be accounted for (A) or specific within-data source solutions may be targeted first (C), before to obtaining an integrative solution. This is indicated by the arrows connecting panel B with respectively A and C.

Figure 3. Fully acknowledging inter-relationships between omics data on reduced genomic sets when deriving integrative solutions.

A – Per meaningful genomic concept, such as a gene, create a network of inter-relationships between omics elements “mapped” to that concept. B – Represents the concept-based integrated data by using kernel-based principal components, where the kernel is chosen in such a way that the structure of the data is optimally captured. This leads to a new integrated concept-related signature for each individual in the sample. Each concept, therefore, gives rise to a new variable. The combined set of concepts (new variables) is submitted to subsequent analyses to obtain an integrated solution to the problem of interest.

Figure 4. The proposed implementation of logical model predictions and patient-derived spheroid testing of drug therapies.

An individual patient’s tumor material (top row) is used to produce spheroid cultures for small scale drug combination screening. In parallel, biomarkers are produced from these spheroids (bottom row) and used to configure a generic logical model so that it optimally represents the tumor of the patient. This model is used for a large scale *in silico* screening of the complete available drug

combination space, resulting in a limited set of potential synergistic drugs that are tested in the spheroids. This whole procedure can be completed in a couple of weeks. Validated drug combinations can be considered by a clinician for therapy decision.

Figure 5. SWOT analysis (Strengths, Weaknesses, Opportunities and Threats) for network and systems medicine.¹⁵⁰