EU-STANDS4PM
standards for in silico models for personalised medicine

A European standardization framework for data integration and data-driven *in silico* models for personalised medicine – EU-STANDS4PM

White Paper

Towards *in silico* approaches for personalised medicine – Recommendations for verifying and validating predictive computational models in EU collaborative research

Imprint

Authors

Alphabetical order:, Kirstine Belling[1], Marina Caldara[2], Catherine Bjerre Collin[1], Tom Gebhardt[3], Martin Golebiewski[4], Tugce Karaderi,[2] Faiz M. Khan[3], Marc Kirschner[5], Sylvia Krobitsch[5], Lars Küpfer[6], Heike Moser[7], Flora Musuamba Tschinanu[8], Mariam Nassar[3], Tito Poli[2], Philip Rosenstiel[9], Dagmar Waltemath[10], Olaf Wolkehnauer[3] and the EU-STANDS4PM consortium*

[1] Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Denmark
[2]University of Parma, Italy
[3]University of Rostock, Germany
[4]HITS gGmbH, Germany
[5]Forschungszentrum Jülich GmbH, Germany
[6]Bayer AG, Germany
[7]German Institute for standardization, Germany
[8]Federal Agency of Medicines and Health Products, Belgium
University of Kiel, Germany
[10]Medical Informatics Laboratory, Institute for Community Medicine, University Medicine Greifswald, Germany

[*] A list of authors and their affiliations appears below

EU-STANDS4PM consortium

Alphabetical order: Rolf Apweiler[1], Stephan Beck[2], Kirstine Belling[3a], Marina Caldara[4], Catherine Bjerre Collin[3a], Niklas Blomberg[1], Søren Brunak[3], Eugenijus Gefenas[5], Martin Golebiewski[6], Tom Gebhardt[7], Kalle Günther[8], Mette Hartlev[3b], Vincent Jaddoe[9], Marc Kirschner[10], Ingrid Kockum[11], Sylvia Krobitsch[10], Lars Küpfer[12], Stamatina Liosi[2], Jurate Lekstutiene[5], Vilma Lukaseviciene[5], Ali Manouchehrinia[11], Arshiya Merchant[1], Neha Mishra[13], Heike Moser[14], Miranda Mourby[15], Wolfgang Müller[6], Flora Musuamba Tshinanu[16], Katharina Eva Ó Cathaoir[3b], Uwe Oelmüller[8], Tito Poli[4], Philip Rosenstiel[13], Dagmar Waltemath[17], Olaf Wolkenhauer[7], Amonida Zadissa[1]

[1] European Bioinformatics Institute (EBI-ELIXIR), United Kingdom
[2] University College London, United Kingdom
[3a] Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Denmark
[3b] Centre for Legal Studies in Welfare and Market, University of Copenhagen, Denmark
[4] University of Parma, Italy
[5] Vilnius University, Lithuania
[6] HITS gGmbH, Germany
[7] University of Rostock, Germany
[8] Qiagen GmbH, Germany
[9] Erasmus University Rotterdam, The Netherlands
[10] Forschungszentrum Jülich GmbH, Project Management Jülich, Germany
[11] Karolinska Institutet, Sweden
[12] Bayer AG, Germany
[13] University of Kiel, Germany
[14] German Institute for Standardization, Germany
[15] Univeristy of Oxford, United Kingdom
[16] Federal Agency for Medicines and Health Products, Belgium
[17] Medical Informatics Laboratory, Institute for Community Medicine, University Medicine Greifswald, Germany

Review

Legal and ethical review by: Katharina Ó Cathaoir[1], Eugenijus Gefenas[2], Mette Hartlev[1], Miranda Mourby[3]

[1] University of Copenhagen, Denmark
[2] Vilnius University, Lithuania
[3] University of Oxford, United Kingdom

Publisher and contact

EU-STANDS4PM administrative office on behalf of the EU-STANDS4PM consortium

Forschungszentrum Jülich GmbH, Project Management Jülich, Germany

Contact: Marc Kirschner (m.kirschner@fz-juelich.de)

Using this content

Please note that the content of this document is property of the EU-STANDS4PM consortium. If you wish to use some of its written content, make reference to:
EU-STANDS4PM consortium (2020); White Paper: Towards in silico approaches for personalised medicine – Recommendations for verifying and validating predictive computational models in EU collaborative research

Links to external websites

This document may contain links to external third-party websites. These links to third-party sites do not imply approval of their contents. Forschungszentrum Jülich GmbH or Project Management Jülich has no influence on the current or future contents of these sites. We therefore accept no liability for the accessibility or contents of such websites and no liability for damages that may arise as a result of the use of such content.

Table of content

Abbreviations

ADME	Absorption, distribution, metabolism, and excretion (phenotype)
AI	Artificial intelligence
ASME	American Society of Mechanical Engineers
CEN	European Committee for Standardization
CENELEC	Comité Européen de Normalisation Électrotechnique
COMBINE	Computational Modelling in Biology Network
DICOM	Digital Imaging and Communication in Medicine
EMA	European Medicines Agency
EHR	Electronic health record
EPR	Enterprise Resource Planning
FAIR	Findable, Accessible, Interoperable and Reusable
FDA	US-Food and Drug Administration
FHIR	Fast Healthcare Interoperability Resources
FMA	Foundational Model of Anatomy
GEUVADIS	Genetic European Variation in Health and Disease (consortium)
GWAS	Genome-wide association studies
HTA	Health Technology Assessment
IT	Information Technology
ICD	International Classification of Diseases
ICRP	International Commission on Radiological Protection
ICT	Information and communication technology
IBD	Inflammatory Bowel Diseases
ISCT	*In silico* clinical trials
ISO	International Standardization Organization
LOINC	Logical Observation Identifiers Names and Codes
MedDRA	Medical Dictionary for Regulatory Activities
MIRIAM	Minimum information requested in the annotation of biochemical models
ML	Machine learning
NPU	Nomenclature for Properties and Units
ODE	Ordinary differential equation
OMOP	Observational Medical Outcomes Partnership
PBPK	Physiologically based pharmacokinetic
PD	Pharmacodynamic
PK	Pharmacokinetic
PRS	Polygenic risk scores
popPK	Population pharmacokinetic
QALYs	Quality life years
QSP	Quantitative systems pharmacology
SBML	Systems Biology Markup Language
SNOMED CT	Systematized Nomenclature of Medicine, Clinical Terms
SNPs	Single nucleotide polymorphisms
STARD	Standards for Reporting of Diagnostic accuracy Studies
STROBE	Strengthening the Reporting of Observational studies in Epidemiology
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis

Executive Summary

Harmonization of data integration is the key to standardization efforts in personalised medicine, which would also facilitate cross-European studies. Standardization of the models themselves is less essential within a research context, where new models are created and tested in line with research progress, harmonization and/or standardization of input data is both feasible and necessary.

We argue that model validation should receive more attention, and other measures should be implemented such that validation of models within personalised medicine becomes easier, also across borders. While this is an evident necessity within the context of models implemented as medical devices or decision tools, which are regulated by the European Medicines Agency and national competent authorities, we argue that model validation should be a higher priority at research level also, facilitating assessment by peers and by medical doctors – who themselves should receive better training in assessment of research using *in silico* models. This will also ease the implementation of translational research results in the clinic.

Acceptance by doctors and the relevant medical specialties is a key hurdle for *in silico* models in personalised medicine. Any medical product - device, algorithm or drug - has to prove itself safe and effective to be licensed for use by regulators; however, it has also to be accepted by medical experts as being a good choice, and be recommended within clinical specialties.

EU-STANDS4PM joined forces to examine to what extent existing standards or standards under development for both format and semantics can be used to link clinical and health as well as research data to computational models relevant for personalised medicine. As all requirements should be equally understood and fulfilled by users it is important to define them uniformly in an international context. To achieve this the conclusion of our work shall be also discussed in international standardization and technical committees, especially in the case of standards that are still being drawn up, and new standardization projects shall be initiated where necessary.

We present a White Paper featuring recommendations for standardization of data integration as well as recommendations for standardization of model validation within a collaborative research context, such that health-related data can be optimally used for translational research and personalised medicine across Europe.

As such the White paper showcases the approach that takes big data in health through harmonized data integration to the most relevant predictive computational models for personalised medicine. As they are refined and validated these models can provide guidance not just how to use data, but also how to best cope with disease and preserve wellbeing in the daily lives of patients.

Introduction

Aims of EU-STANDS4PM

EU-STANDS4PM is a Horizon 2020 funded Coordinating and Support Action tasked to develop harmonized transnational standards, recommendations and guidelines that allow a broad application of predictive *in silico* methodologies in personalised medicine across Europe. EU-STANDS4PM will assess and evaluate national standardization strategies for health data integration, as well as data-driven *in silico* modelling approaches for personalised medicine, with the aim to bundle European standardization efforts. A major goal is to develop harmonized standards as well as recommendations and guidelines for predictive data-driven *in silico* (here: mathematical and computational modelling, (Wolkenhauer et al. 2014))[1] methodologies applied in personalised medicine.

The challenge of making sense of big data

Although Big Data already drives fundamental medical/scientific applications and the associated socioeconomic potential forward, a large-scale future exploitation of Big Data in research and health care represents a major challenge (Apweiler et al. 2018)[2]. This concerns both technical and safety issues, as well as legal, ethical, social and cultural aspects (Ó Cathaoir K. 2020) in dealing with personal health-relevant and large-volume data sets that differ to a great extent in Europe. In addition to country-specific heterogeneities in regard to e.g., technical or cultural, legal and ethical issues, Europe currently lacks well-functioning, standardized and interoperable Information and Communication Technology (ICT) infrastructure, capable of linking the databases of basic and clinical research with different registries (including those from environmental, food and social sciences and humanities), whilst addressing both legal and ethical frameworks to maintain public trust (Horizon-2020-Advisory-Group 2018-2020) in a proper legal, ethical and privacy-protective data environment. Publicly accepted strategies and governance frameworks for integrated Big Data to further develop healthcare –quality and –system performance are key factors that have to be implemented before Big Data can unfold its full medical and research potential through *in silico* analysis and interpretation. Regarding Big Data in health the specific challenges are thus to:

- Develop data-driven computational approaches that are tailored (personalised) to the individual or stratified patient groups addressing clinically relevant questions.
- Harness, utilize and understand (exploit) high volume, high diversity biological, clinical, environmental, and lifestyle information.
- Develop European harmonized standardization guidelines for data integration strategies.
- Ensure that integration of personal and patient-derived data is performed lawfully, ethically and fairly, including in full respect of patients' rights.
- Standards on a European level for secure interoperable data integration and predictive computational models are therefore essential to utilize the wealth of information that Big Data contain — specifically and efficiently to push a pro-active personalised medicine forward.

To achieve these goals both (i) the "technical" element of standardizing data input and modelling, and (ii) the legal and ethical framework supporting data integration and interoperability must be addressed. The legal and ethical framework is analysed in the EU-STANDS4PM paper "Legal and ethical review of in

[1] In the context of EU-STANDS4PM the term *in silico* refers to mathematical and computational models of biological systems, such as molecular modelling, modelling of subcellular processes, individual-cell or cell-based models, tissue/organ level models, body systems level models (Wolkenhauer et al. 2014).

[2] We refer to the following definition "*Big data in health* encompasses high volume, high diversity biological, clinical, environmental, and lifestyle information collected from single individuals to large cohorts, in relation to their health and wellness status, at one or several time points" (Auffray et al. 2016)

silico modelling" (Ó Cathaoir K. 2020). The broader context of this White Paper is thus to promote and initiate the development of technical standards as well as to provide recommendations for data-driven modelling approaches for personalised medicine.

In a preceding EU-STANDS4PM paper "Towards standardization guidelines for in silico approaches in personalized medicine" (Brunak et al. 2020) we initiated this development and addressed crucial requirements associated with the implementation of computational models in personalized medicine, including a first set of general recommendations to key actors (please see info box 5, annex). The current White Paper follows this direction and provides a comprehensive analysis of computational models that are of relevance for personalized medicine. Each modelling approach contains a collection of specific challenges and corresponding recommendations for data input and model validation.

Motivation for standardization in personalised medicine

Common standards support communities (see also info box 2) with a basis for mutual understanding and information exchange and common standards are indispensable for collaborative work. Data from different sources and recorded at different times must be integrated in order to setup computer models in personalised medicine. Consistent documentation of data, models and simulation results based on standards ensure that the data and corresponding metadata (data describing the data and its context), as well as models, methods and visualizations are structured and described in a "FAIR" manner: Findable, Accessible, Interoperable and Reusable (Info box 1) (Wilkinson et al. 2016).

Info box 1: FAIR guiding principles: Findable, Accessible, Interoperable and Reusable

According to Joint Declaration of Data Citation Principles (JDDCP; https://www.force11.org/fairprinciples):

To be **Findable** any Data Object should be uniquely and persistently identifiable

- The same Data Object should be re-findable at any point in time, thus Data Objects should be persistent, with emphasis on their metadata
- A Data Object should minimally contain basic machine actionable metadata that allows it to be distinguished from other Data Objects
- Identifiers for any concept used in Data Objects should therefore be Unique and Persistent

Data is **Accessible** in that it can be always obtained by machines and humans

- Upon appropriate authorization
- Through a well-defined protocol
- Thus, machines and humans alike will be able to judge the actual accessibility of each Data Object.

Data Objects can be **Interoperable** only if:

- (Meta) data is machine-actionable
- (Meta) data formats utilize shared vocabularies and/or ontologies
- (Meta) data within the Data Object should thus be both syntactically parseable and semantically machine-accessible

For Data Objects to **be Reusable** additional criteria are:

- Data Objects should be compliant with principles 1-3
- (Meta) data should be sufficiently well-described and rich that it can be automatically (or with minimal human effort) linked or integrated, like-with-like, with other data sources
- Published Data Objects should refer to their sources with rich enough metadata and provenance to enable proper citation

A central aim of EU-STANDS4PM is to develop ISO Technical Report(s) together with the existing ISO bodies as guidelines and recommendations for the data and models to be used - as best practice. These activities will take the FAIR principles into account.

Hence, data and model standards support the reliable exchange of health-related data, making the data FAIR for their integration into computer models used in personalised medicine. Such data and model standards, together with harmonized ways to describe their metadata, also ensure the interoperability of tools used for data integration and modelling, as well as the reproducibility of the simulation results of the models. In that sense modelling standards are agreed ways of consistently structuring, describing

and associating models and data, their respective parts, their graphical visualization, as well as information about applied methods and the outcome of model simulations. Such standards also assist with describing how constituent parts interact together, or are linked, and how they are embedded in their physiological context.

Major challenges in the field of personalised medicine are to harmonize the standardization efforts that refer to different data types, approaches and technologies, as well as to make the standards interoperable so that the data can be compared and integrated into models. Reproducible modelling in personalised medicine requires a basic understanding of the modelled system, as well as of its biological and physiological background. There is a relevant checklist that provides guidelines on the minimum amount of metadata information required in order to understand a model (minimum information requested in the annotation of biochemical models (MIRIAM), (Le Novere et al. 2005)). This information about data and models can be transferred by using metadata in the form of semantic annotations. These annotations can improve the shareability, and interoperability of the data or model (Neal et al. 2019). To render data and models FAIR, it is important that all their elements (entities) in their context are understood in exactly the same way, independently from the individual or tool that process or analyses them. For this purpose, it is necessary to consistently use the defined terminologies, such as controlled vocabularies and domain ontologies that can be defined and applied independently of the data/model format.

For many different data types used in personalised medicine domain-specific annotation standards and terminologies are available. For example, UniProt[3] or the Protein Ontology[4], can be used to uniquely identify proteins in a particular biological context which can then be linked to specific entities in the computational model. Similarly, the Gene Ontology[5] could be used to identify specific genes or cellular components whereas the Foundational Model of Anatomy (FMA) (Rosse and Mejino 2003) can be used to localize an entity in the computational model to specific spatial location or anatomical structure. If not found completely or partially unstructured, which is often the case, health-related data is most commonly structured and codified by specific formatting standards for medical data. These can be the interoperability standard HL7 Fast Healthcare Interoperability Resources (FHIR) (Bender and Sartipi 2013), or the standard for electronic health records (openEHR (Kalra, Beale, and Heard 2005)). Semantical content is usually annotated with domain-specific clinical terminologies, e.g., International Classification of Diseases (ICD)[6], Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT, see table 1, annex), or Logical Observation Identifiers Names and Codes (LOINC, see table 1, annex). Thus, the wheel does not have to be reinvented for the semantic data annotation in personalised medicine, but existing annotation standards have to be consistently applied.

Info box 2: Community vs. formal standards

Community standards usually reflect the results of a grass-roots standardization effort from a specific user group and are created by individual organizations or communities. They can cover a broad variety of different topics and are typically domain-specific with a focus on the community of their origin, whereas there is no formally prescribed process for creating, agreeing and consensus-building.

Formal standards are created by official national or international standardization bodies (e.g., CEN/CENELEC or ISO) based on the consensus principle, in a defined procedure with the participation of all interested stakeholders. Once completed and released, formal standards are internationally respected and recognized as state of the art – also from a quasi-legal perspective, although they do not have regulatory functions. Nevertheless, sometimes standards are referred to in directives or laws and thus acquire a "law-like" character.

[3] UniProt: https://www.uniprot.org/;
[4] Protein Ontology: https://www.ebi.ac.uk/ols/ontologies/pr
[5] Gene Ontology: http://geneontology.org/
[6] ICD: https://www.who.int/classifications/icd/en/

Data-driven computational models for personalised medicine

The future development of personalised medicine is dependent on a vast exchange of data from different sources, as well as harmonized integrative analysis of large-scale personalised medicine data (Big Data in health). Computational modelling[7] approaches play a key role in order to understand and predict the underlying molecular processes and pathways that characterize human biology, but they also lead to a more profound understanding of the mechanisms and factors that drive disease, hence they allow personalised treatment strategies that are guided by central clinical questions.

Computational models have the potential to translate in vitro, nonclinical and clinical results (and their related uncertainty) into descriptive or predictive expressions. Over the last decades, the added value of such models, also called digital evidence, in medicine and pharmacology has increasingly been recognized by the scientific community (Apweiler et al. 2018; Wolkenhauer et al. 2014; The CASyM consortium 2014; Thiel et al. 2015), as well as regulatory bodies such as the European Medicines Agency (EMA guideline on physiologically based pharmacokinetic (PBPK) reporting[8]) or the US Food and Drug Administration (FDA guidance document: Reporting of Computational Modeling Studies[9]) – irrespective of their ultimate use or application. Computational models are now integrated in different fields in medicine and drug development expanding from disease modelling, biomarker research to assessment of drug efficacy and safety. In this context models are regulated as medical devices, an area explored in detail by the Virtual Physiological Human Institute commentary paper "Verifying and Validating Quantitative Systems Pharmacology and *In silico* Models in Drug Development: Current Needs, Gaps, and Challenges" (Musuamba et al. 2020) and regulated by relevant authorities (e.g., EMA and/or FDA guidelines). *In silico* approaches are also expanding in neighbourhood fields such as pharmacoeconomics (Dasbach and Elbasha 2017; Ademi et al. 2013), analytical chemistry (Oliveri 2017; Zaborenko et al. 2019) and biology (Hood and Tian 2012; Weston and Hood 2004).

Patients will greatly benefit from this development that equips personalised medicine with predictive capabilities to simulate *in silico* clinically relevant questions, such as the effect of therapies, the response to drug treatments or the progression of disease. Currently there are a number of computational modelling approaches in pre-clinical and clinical research that are able to address these questions in greater detail and, therefore, play a leading role for the future development of personalised medicine.

However, despite the growing popularity of computational modelling approaches (Morrison et al. 2018; Saez-Rodriguez and Blüthgen 2020; Wolkenhauer et al. 2014; Wolkenhauer et al. 2013; Apweiler et al. 2018), there are still many hurdles to overcome. Especially the integration of clinical and life science data from multiple sources and types is a highly complex task. Figure 1 illustrates a typical workflow followed in personalised medicine starting from the clinical question, and followed by identification, access and harmonization of relevant data, the development of data model(s), the model validation, and finally the application in a clinical setting.

[7] Computational modelling in this context refers to mathematical and computational models of biological systems, such as molecular modelling, modelling of subcellular processes, individual-cell or cell-based models, tissue/organ level models, body systems level models (Wolkenhauer et al. 2014)

[8] https://www.ema.europa.eu/en/reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation

[9] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/reporting-computational-modeling-studies-medical-device-submissions

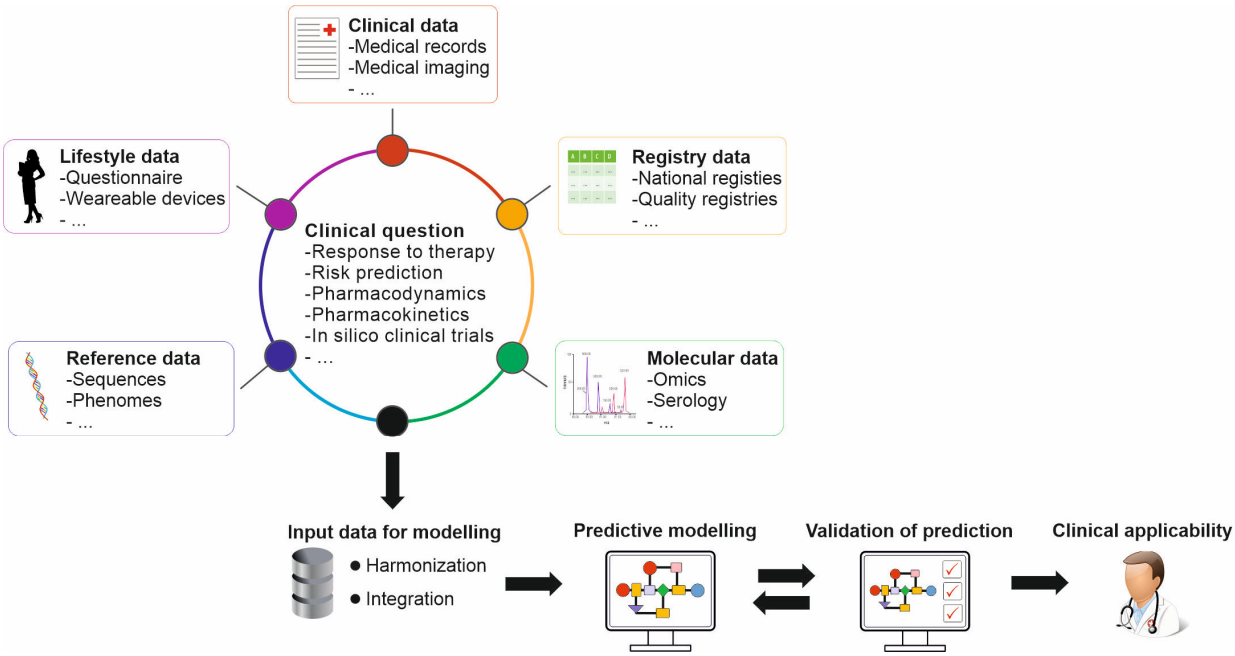

Figure 1: Modelling workflow for personalised medicine studies.
Model creation starts with a clinical question followed by data collection, which can be derived from a knowledge-driven (e.g., literature research) or data-driven (experimental data) approach. The data employed need harmonized approaches for data integration to start the model construction. The initial model usually undergoes several refinement and improvement iterations to enhance predictive capabilities. The use of common standards (summarized in tab. 1, annex) is usually necessary for model building and curation process. Accuracy measurements and validation processes are key, and should be transparent, while model output and function should ideally be interpretable and explainable.

Model building key factors

Clinical and health related data are usually generated in different medical environments and by various sources and systems; with the advent of health data created outside the health care setting, sources and formats become even more diverse. Such heterogeneous data require harmonized strategies for data integration utilizing broadly applicable standards that allow for a reproducible data exploitation to generate new knowledge for medical benefits. For the development of validated predictive computational models in personalised medicine two key factors of the model building process, for which standardization is essential, are:

Data integration – Harmonized strategies and methods for integration of data
Model validation – Validation of models and simulations through the underlying clinical question

Standardization of data input is based on many common requirements. All model building needs harmonized strategies to ensure comprehensive, high quality and unbiased data. This also includes common languages and procedures to store, share and communicate results. In addition overarching collaborations are key to increase diversity and sample size and all require clear concepts as well as harmonized procedures to ensure reproducibility and interoperability for model building.

Common to all *in silico* models is a need for validation (Debray et al. 2015) and accuracy[10], however, in contrast to data input model validation methods are considered to be individual and type-specific. It is important that any algorithm performs well on novel data that have not been used in training the algorithm; i.e. the model should be able to generalize to new data from the same domain (Baldi et al.

[10] http://www.clinicalpredictionmodels.org/doku.php?id=validation:start

2000). There are guidelines and methods for validating models which are accurate and confident in predictions, both in terms of accuracy and confidence in predictions. Performance evaluation of *in silico* models should be transparent and consistent to existing guidelines, explaining the reasons for not doing so when alternative methods are chosen.

It is important that the standards to be developed for model building (data input) and validation (model validation) are commonly acknowledged by all the involved parties (clinicians, Health Technology Assessment agencies, academia, industry, regulators and patient organizations but also funding organization such as the EU-Commission investing heavily in sustainable research and technological developments) and are relevant for all the types of models used in the clinic.

Data and model integration

Data integration: Currently there are no widely accepted, overarching strategies or concepts to harmonize the integration of heterogeneous health and disease data for *in silico* models to interpret Big Data for personalised medicine. What is lacking are standardization documents on a European or international level (e.g., European Committee for Standardization, CEN; International Standardization Organization, ISO), and recommendations that allow exploitation of Big Data to develop true medical benefits for an individual patient or stratified patient groups. In order to be effective, such recommendations need to be disseminated broadly and adopted by the relevant scientific and clinical communities. *In silico* modelling of all kinds requires high quality data integration to create reliable output, but the vast quantities of research, medical and health data in existence are too disparate to be harnessed optimally. Data can be integrated at different levels, also with regard to personal identifiability. These represent different opportunities:

- *Individual level integration:* Data linked at the personal ID level can be used for personalised prediction of disease progression for the individual, i.e. personalised therapy based on past disease and health data. Individual level data integration is important for research in rare diseases, where patient numbers are low and counting a patient more than once would affect data quality.
- *Integration of variables*: Data can be combined at the variable level, e.g., diagnoses, lab values, therapeutic intervention, symptoms, scores, outcomes or omics.
- *Integration of unstructured data*: unstructured data can be integrated and processed together.
- *Federation of data or validation of findings*: Data sets can contribute to joint results by training an algorithm sequentially on the data sets without combining them, or by using new data sets for validation. While there are models that can process non-harmonized data and learn from them, this is often inefficient, and for most systems of federation of clinical data, e.g., Observational Medical Outcomes (OMOP) Partnership[11], data sets still have to be harmonized and interoperable to return useful results.

While some challenges (e.g., legal, ethical) are connected to national boundaries and limitations in European and national data protection regulation as well as patients' rights laws, e.g., requiring the consent of individuals (e.g., patients), many others don´t.

However, EU countries have achieved different levels in terms of data standardization. Here, we examine some examples taken from the use case studies table presented in the annex, in terms of level of data standardization and the impact on the research performed. Notably, lack of standardization in data prevents both the re-use of data in later projects, and the validation of models with other data sets.

Harmonized strategies for data integration is key to standardization efforts in personalised medicine and standardization of input data is both feasible and necessary. Data harmonization can take place during data creation or post hoc, and is usually to some extent necessary for data integration. Data harmonization is necessary because of:

[11] OMOP: https://www.ohdsi.org/data-standardization/the-common-data-model/

- Differences in Information Technology (IT) systems used in data generation, e.g., Enterprise Resource Planning (EPR) systems and lab result software or hardware, at national, regional or local clinical centre level.
- Adoption of different standards such as Nomenclature for Properties and Units (NPU) or LOINC.
- Differences in implementation of international terminologies such as ICD.
- Language differences in, e.g., unstructured text.

Data harmonization requires canonical interoperability (placement of the data in an ontology or structure), syntactic interoperability (data packaging) physical interoperability, which can be a challenge in cases of legal restrictions or very large data sets and semantic interoperability, which usually requires mapping of data sets onto each other for comparison and/or analysis. For example, multiple data types can be mapped into the OMOP common data model federation. Each transformation and mapping step carries the risk of changing the meaning of the term slightly (e.g., use of synonyms in ontologies or lack of terms), undocumented knowledge can be lost, and errors being created.

Alternatively, raw data can be pre-processed in similar ways, and harmonized mathematically to produce larger data sets that can be analyzed together. Standard terminologies exist for the core elements of clinical practice (diagnoses, symptoms and observations; interventions, procedures, treatments and medication; health outcomes e.g., disability, quality life years (QALYs), symptom status), such as Medical Dictionary for Regulatory Activities[12] (MedDRA), SNOMED CT[13] – or Digital Imaging and Communication in Medicine[14] (DICOM, ISO 12052) (DICOM_Secretariat 2020). For a more comprehensive overview of already existing standards relevant for personalised medicine (please see table 1, Annex).

These and other core standards are widely used, and most systems either incorporate these standards in data generation, or produce data which can be mapped to them. However, these standards also contain regional or national implementations that require mapping. One core area which is missing comprehensive standards is patient identification. Data from disparate clinical sources can be linked at the national level, but not across borders; conversely, data from research projects can be linked across research centres, but not necessarily to clinical data. Harmonization of patient identifiers could have the potential to be ethically problematic as they could facilitate combination of data sets and re-identification of data subjects. However, with an ethically sound regulation of patient data sharing, a common identifier in itself is not regarded as problematic.

Model integration: Given the increasing flood and complexity of data in personalised medicine, standardization of these data and their documentation are crucial. This comprises the consistent description of the applied diagnostic and therapeutic methods and also the workflows for:

- Data processing, analysis, exchange and integration (e.g., into model codes and calculations).
- Biological sources (the individual or patient, organs, tissues, etc.).
- Corresponding medical information (e.g., electronic health records, diagnostic results and values, as well as drug concentrations, responses to treatments, biomarkers).
- The setup, handling and simulation of the models.

Hence, standards for formatting and describing data, workflows and computer models have become important, especially for data integration across the biological scales for multiscale approaches (Schreiber et al. 2019).

To this end, the corresponding scientific communities have defined many grassroots standards to consistently structure and format data, models and their metadata for modelling in the life sciences (Golebiewski 2019). These standardization efforts are driven by standardization initiatives, such as the

[12]MedDRA: https://www.ich.org/page/meddra
[13] SNOMET CT: http://www.snomed.org/
[14] DICOM: https://www.dicomstandard.org/

Computational Modelling in Biology Network (COMBINE)[15] (Myers et al. 2017; Hucka et al. 2015). For providing the potential users with an overview and comparable information about such standards, web-based information resources have been developed and are publicly available, such as the NormSys[16] registry for modelling standards.

For facilitating the integration of data and models, formatting and description standards have to be harmonized to become interoperable and allow interfacing between the often heterogeneous data sets and/or model parts. To support this, novel standards are defined by the International Organization for Standardization (ISO) in its technical committee 276 – Biotechnology[17] (ISO/TC 276). One example is the emerging standard ISO 20691 "Requirements for data formatting and description in the life sciences for downstream data processing and integration workflows", which defines a guideline and framework for interoperable community data standards in the life sciences with emphasis on their application. Such standards aim at enhancing the harmonization and interoperability of standards for life science data and models and therefore facilitate complex and multiscale data integration, as well as model building with heterogeneous data gathered across the domains.

The following table contains a summary of the main challenges for data integration (data input) that occur during model building:

General challenges for data integration and model building

- High degree of variability regarding data types (structured, unstructured, molecular, clinical, patient-reported, etc.)
- Differences in coding and calculation within data types (between machine variability, different measurements, etc.)
- Heterogeneous utilization of existing data
- Costs of data harmonization efforts are currently high, in terms of time, resources and data quality loss
- Models relevant for clinical use need to fit for purpose
- Differences in IT systems used in data generation, e.g., EPR systems and lab result software or hardware, at national, regional or clinical centre level
- Adoption of different standards such as NPU or LOINC
- Differences in implementation of international terminologies such as ICD
- Language differences in unstructured text, and other factors

Common recommendations for data integration

Taken the above discussed challenges into account a common set of recommendations applicable to most model building approaches can be defined:

R1: Develop tools to standardize and harmonise the data from different centres and laboratories that work with or develop similar modelling approaches
R2: Develop data quality assessment frameworks to evaluate data used for modelling
R3: Develop clear and harmonized reporting (Artificial Intelligence (AI) Section)
R4: Exclude non-harmonized or inappropriately pre-processed data (compare AI Section)
R5: Make sure that the standards to be developed for model building (data input) and validation (model validation) are commonly acknowledged by all the involved parties e.g., clinicians, Health Technology Assessment (HTA) agencies, academia, industry, regulators and patient organizations but also funding organization such as the EU-Commission investing heavily in sustainable research and technological developments

[15] COMBINE: http://combine.org
[16] NormSys: http://normsys.h-its.org
[17] ISO/TC 276: https://www.iso.org/committee/4514241.html

Info box 3: Are formal standards used in modeling and where are the deficits?

"The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement includes a 22-item checklist, which aims to improve the reporting of studies developing, validating, or updating a prediction model, whether for diagnostic or prognostic purposes. The TRIPOD Statement aims to improve the transparency of the reporting of a prediction model study regardless of the study methods used. This explanation and elaboration document describes the rationale; clarifies the meaning of each item; and discusses why transparent reporting is important, with a view to assessing risk of bias and clinical usefulness of the prediction model. Each checklist item of the TRIPOD Statement is explained in detail and accompanied by published examples of good reporting. The document also provides a valuable reference of issues to consider when designing, conducting, and analyzing prediction model studies" (Collins, Reitsma, et al. 2015a; Moons et al. 2015).

The guidelines have been repeatedly found to be inadequately followed, such that few machine learning (ML) studies meet basic reporting standards for clinical prediction tools, and even fewer make their models available such that they can be reproduced or adequately evaluated (Heus et al. 2018b; Liu et al. 2019; Wang et al. 2020).

However, the situation is evolving in the right direction. The ML community working with health data is increasingly aware of reporting standards. The digital health journals do not uniformly require them – but the medical journals do, usually either

- Standards for Reporting of Diagnostic Accuracy Studies (STARD, Cohen et al. 2016),
- Strengthening the Reporting of Observational studies in Epidemiology (STROBE, Cuschieri 2019) or
- Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD, Collins, Reitsma et al. 2015b).

With ML research increasingly being published in medical journals, such as the Lancet Digital Health, we foresee that reporting standards will be followed more frequently. Journals' requirements and funders' enabling should thus progress hand-in-hand.

Computational models addressing clinically relevant questions

The current paper reflects on modelling approaches that are able to address central clinically relevant questions. In the following section the most relevant modelling approaches able to address these questions will be discussed. We will provide an overview focusing on state of the art, challenges and recommendations to these modelling approaches.

Computational model approaches relevant for clinical applications

Model: Cellular systems biology
Purpose: Dynamic mechanistic models of complex biological processes and networks

Model: Risk prediction for common diseases
Purpose: Prediction of disease prognosis

Model: Disease course and therapy response prediction
Purpose: Prediction of disease prognosis

Model: Pharmacokinetic/-dynamic modelling and in silico trial simulations
Purpose: Prediction/simulation of drug exposure and effects in (different components of) living systems; simulation of 'virtual populations' for testing of pharmacological therapies and devices

Model: Artificial intelligence
Purpose: Data-driven approaches utilizing AI and ML

Cellular systems biology

There are numerous experimental methods to observe and quantify processes within an organism at an unprecedented level of detail. Some of these molecular approaches have already entered clinical diagnostics ((Hastings et al. 2020; Briganti and Le Moine 2020), adding information about the patient, generating vast amounts of individualized data. *In silico* processing and interpretation of clinical measurements can be either data-driven or mechanism-based (also called top-down and bottom-up approaches of systems biology (Bruggeman and Westerhoff 2007; Edwards and Thiele 2013)).

Mechanism-based concepts aim for a structural representation of the governing physiological processes based on model equations with limited amount of data, which are required for base model establishment (Saez-Rodriguez and Blüthgen 2020) or, alternatively, on static interacting networks (Vidal, Cusick, and Barabasi 2011; Fiers et al. 2018). Therefore, on one side *mechanism-based concepts* could rely on gathering information from diverse resources, ranging from literature to experimental data, and combining them to build an equation based model; a process that can be personalised by changing the parameters according to the patient data, and simulate potential outcomes and treatment approaches. The construction of *mechanism-based models* requires time-consuming manual efforts. Often, these efforts are disease-specific, although models for different diseases may overlap. On the other side strategy, *mechanism-based concepts* combine a variety of data like genome-wide association studies, identified risk factors single nucleotide polymorphisms (SNPs), pathway(s), symptoms and phenotypes to build a simple but informative network that provides information on the consequences of a genetic modification, guides precision medicine, explains and avoids drug resistance in cancer chemotherapy, and suggests the most efficient drug to hit a target (Halu et al. 2019; Zhou et al. 2014; van der Wijst et al. 2018).

Data-driven approaches (Wolkenhauer et al. 2014; Apweiler et al. 2018) require, as the name implies, sufficiently rich and quantitative time-course data to train and to validate the model. Due to its black-box nature, the model validation process relies on performance tests against known results. Thus, personalised data-driven models cannot rely exclusively on the input of one patient. To properly build the model, one needs a large set of personal data from a variety of patients with defined inclusion criteria (e.g., similar symptoms) and an additional data set for comparison, e.g., from healthy persons. Thus, the main challenges are the unavailability of sufficient data in high quality, incomplete data sets and sparse training data (Alber et al. 2019). Availability of data furthermore depends on legislation, data security matters and relies on safe anonymization strategies (see also AI section below).

Both, data-driven approaches and mechanism-based concepts share the general requirements for data harmonization, data integration, data and method standardization and documentation (Apweiler et al. 2018; Schreiber et al. 2019; Neal et al. 2019). Ultimately, from a technical point of view, the basic workflows are largely complementary.

Challenges

Abstraction vs size

One challenge is to create models that balance the level of abstraction (size) with comprehensiveness. Thus, one challenge is to make the efforts reproducible and reusable. Modelling the processes in smaller modules and connecting them automatically would improve the situation and accelerate future model creation and adaption. It would also support collaborations for model construction.

Data for validation

In addition to rules for data security, anonymization and consent the challenge here is the availability of data which highly depends on:

- Development of prediction models that can easily be adopted to individual patient profile
- Efficient parameter estimation tools to cope with population and disease heterogeneity

- Not overfitting the model to the experimental/patient data
- Optimization method for model predictions in a realistic parametric uncertainty
- More flexibility in models to cope with missing data (e.g., diverse patient profiles)
- Scaling from cellular to organ and to organism levels (high clinical relevance, high hurdles for regulatory acceptancy)

Recommendations

Data integration: See common recommendations for data integration in section "Data and model integration"

Model validation: Mechanistic models can serve multiple purposes, which makes it difficult to have a strict set of guidelines. Models in systems biology are encoded and curated by community defined standards including machine-readable format Systems Biology Markup Language (SBML) (Hucka et al. 2018) and MIRIAM (Novère et al. 2005) respectively; but this does not imply a direct application for clinical practice.

Performing numerical predictions, sensitivity and identifiability analysis on mechanistic models (Ingalls and Sauro 2003; Raue et al. 2009) are important systems theoretic tools to validate a model. For instance, identifiability analysis provides conclusions about the uniqueness of the model parameters regarding model structure, i.e. whether there could be others sets of parameter values that lead to the same behavior/predictions. This could be described as a validation of the parameter values.

The most basic form of validation for a mechanistic model is how simulations numerically match experimental/patient data.

R1: Develop a standardized protocol for patient history and information and integrate with clinical standard protocols, e.g., electronic health records (EHRs) and FHIR.
R2: Agree on the use of patient expression profile e.g., mRNA or protein concentration.
R3: Develop prediction models with a fitting hypothesis.
R4: Use model replication and reproduction before considering clinical trials.
R5: Compare model predictions (e.g., biomarkers selection) with established clinical ones.
R6: Develop of user friendly graphical interfaces to ease the use of models in clinic.

Info box 4: Hybrid or grey box models

A promising option to combine data-driven statistical learning and mechanistic, knowledge-based models are hybrid models – also called grey box models or semiparametric models (Frohlich et al. 2018). In this approach, structurally well-understood sub-models are linked through input- and output relations (Muller and Schuppert 2011; Balabanov et al. 2013). The functioning of these connecting nodes, in turn, needs to be trained with experimental data according to concepts from statistical learning. It is thus possible to combine sub-models with very different levels of prior mechanistic understanding and to compensate in particular for a lack of knowledge in some of these sub-systems. Hybrid modelling thereby supports a modular modelling approach where models with different levels of granularity. If new data and information requires an extension of a specific building block, the corresponding sub-model can be revised and afterwards be re-integrated into the overall model network. Thereby it is possible to also combine very heterogeneous model parameters representing for example transfer of mass and energy on the one hand and flow of information on the other. Potential applications are models of the cardio-renal systems (Hallow and Gebremichael 2017), which combine mechanistic descriptions of the cardiovascular system and nephron filtration and reabsorption with neurohormonal regulation.

Risk prediction for common diseases

Common complex diseases are multifactorial and polygenic; there are multiple genetic and environmental (e.g., diet, smoking) factors that affect an individual's risk of having a disease. The polygenic model assumes that the genetic variance of a disease is a combination of small effects of multiple variants across the allele frequency spectrum. Genome-wide association studies (GWAS) that scan the genomes of thousands of individuals offer a very powerful method to identify these multiple genetic risk factors for having the disease.

The emergence of GWAS with the 'Common Disease, Common Variant' hypothesis (Pritchard and Cox 2002) soon proved not to comprehensively capture the polygenic and multifactorial nature of common diseases. In addition to the continuous need for larger sample sizes to increase statistical power to detect more genetic associations with diseases of interest, the *missing heritability* [i.e. the missing portion of phenotypic variance in a population attributable to additive genetic factors; (Manolio et al. 2009)] issue has highlighted the additional factors to be considered in genetic studies such as:

- Accuracy of heritability estimates and contribution of the environmental factors
- Allelic architecture of common diseases
- Contribution by the rare (frequency<1%) and low frequency variants (1%<frequency<5%)
- Contribution by structural variants such as copy number variants
- Epigenetic effects
- Parent-of-origin effects
- Accurate phenotype definitions
- Lack of population diversity in genetic studies

Taking these factors into account to determine the genetic risk factors for common diseases is critical for the estimation of polygenic risk scores (PRS) as well. However, the application of polygenic scores are still proved to be of importance for: (i) Association testing with the composite PRS that can provide evidence for a polygenic model for a disease even when the statistical power of a GWAS is limited as in schizophrenia (International Schizophrenia et al. 2009); and (ii) Disease risk prediction for individuals, which has been less successful due to the limited power of PRS for distinguishing affected individuals from unaffected individuals, hence, could not be translated into clinical use. Use of data for genome-wide variants also did not improve the clinical discriminatory power of PRS (Dudbridge 2013).

PRS is estimated to predict the risk of an individual having a disease based on the individual's genetic make-up. The potential of PRS is that it can aid in early diagnosis as well as precision medicine, leading to better health outcomes for individuals. There are different approaches developed for the estimation of PRS that may lead to some improvement over time (Marquez-Luna et al. 2017; Duncan et al. 2019; Choi, Heng Mak, and O'Reilly 2018). Overall, the main approaches rely on:

- The genetic structure of the relevant population, i.e. the linkage disequilibrium structure across the genome to minimize the correlation among the variants included in the PRS estimation
- Known genetic associations with the disease of interest and the effect sizes from GWAS to be used as weights to calculate the PRS
- Adjusting for ancestry by using principal components
- A training GWAS using a discovery sample and an independent test sample. PRS is calculated by summing risk alleles per individual and weighing each allele with its effect size (Grinde et al. 2019; Pare, Mao, and Deng 2017; Vilhjalmsson et al. 2015; Okser et al. 2014; Wray et al. 2013; Dudbridge 2013; International Schizophrenia et al. 2009; Wray, Goddard, and Visscher 2007)

Challenges

Genetic analysis

Although genetic analyses themselves are, in most part, based on well-defined and well-accepted standards, there are still multiple challenges regarding how these analyses are performed across studies and populations. In many cases, the challenges are due to problems *simultaneously* relating to input data, models and model validation; such a key challenge is the lack of ethnic diversity in *input data*, making it difficult to have adequate *model validation* across different populations, and hence, to extrapolate risk scores to these populations. Many studies have highlighted the limitations of applying PRS obtained from European cohorts to other populations due to bias and reduced predictive accuracy (International Consortium for Blood Pressure Genome-Wide Association et al. 2011; International Schizophrenia et al. 2009; Brown et al. 2016; Marquez-Luna et al. 2017).

Choice of appropriate methods

Methodological choices when preparing the input data preparation such as use of different filtering thresholds is not stringent and harmonized across studies, and therefore, model assumptions may vary. A review of the first decade of PRS studies (2008-2017), it has been demonstrated that methodological choices relating to the appropriate use of linkage disequilibrium structure and varying variant frequencies across populations are highly relevant for more accurate application of PRS and its predictive performance (Duncan et al. 2019).

Access to genetic data

For both European and non-European populations, a comprehensive consideration of the complete allele frequency and effect size spectra is still required – genetic associations with common and low frequency variants are detected with the availability of powerful studies with increasing sample sizes. But rare variants still remain understudied and their contribution to polygenic risk is not well accounted for. In addition, structural variants (e.g., copy number variants, currently not even listed in GWAS catalogue) are also not studied as widely as SNPs. Exome and whole genome sequencing is becoming more affordable for larger sample sizes yet, it is still not widely available in many cases. Access to cellular/tissue-relevant data for functional analyses is also key for further progress in application of PRS. Moreover, environmental factors such as life style are also necessary to understand the contribution of genetic variants to disease risk. However, measurement/record of such factors are often missing, or not as accurate as the genetic factors. Clinical characteristics of individuals including known/potential co-morbidities and family history have traditionally been considered in risk assessment of patients. Effectively combining these with PRS have the potential to significantly improve the predictive power further (Khera et al. 2018; Yang et al. 2010). National biobanks and international consortia providing large sets of different data types (e.g., omics, clinical, environmental) from diverse sample populations will aid in the emergence of useful genetic predictors. Data sharing in a harmonised manner as well as transparency of methods used for analysing such large data sets will be of great importance for reproducible studies and translatable/interpretable results across populations.

In summary

- Access to individual-level genetic data from diverse populations and harmonised summary statistics across published studies (i.e. publicly available data)
- Limited replication of genetic associations and poor application of PRS *diverse populations* (e.g., too poorly represented to be of interest for specific analyses), specifically of mixed or non-European ancestry
- Varying transparency of methodological choices and reproducibility
- Limited cellular/tissue context and harmonized functional data availability across populations/studies
- Missing environmental information coupled to genetic data

Recommendations

Data integration: See also common recommendations for data integration in section “Data and model integration”.

R1: Ensure highest possible diversity and sample sizes further for all the genetic studies on complex diseases as well as for performing functional studies.

R2: Enable more transparent, standardized and detailed methods clearly stating methodological choices made with necessary justifications to enhance reproducible research.

Disease course and therapy response prediction

Despite all advances in medicine predicting the disease course and treatment response of an individual patient remains a major challenge. For complex diseases, such as chronic inflammatory diseases, degenerative disorders or cancer entities, it is still virtually impossible to predict the disease behavior (mild vs. severe, stable vs. progressive) early in the disease course. Disease classifications often use morphometric scores, which reflect rather the secondary tissue damage (e.g., metastasis and tumor size in cancer or stenosing inflammation in inflammatory bowel disease (IBD)) or late stage processes (e.g., memory loss in Alzheimer´s disease) than molecular measurements which would allow a more precise modelling of disease course (Apweiler et al. 2018). It is thus important to develop biomarkers and dynamic models which allow an improved timing of therapy introduction and choice of therapy scheme ("wait and see" vs "hit hard and early" (Schultze, consortium, and Rosenstiel 2018)). For many complex diseases, parallel first-line targeted treatment options have emerged. Although such targeted therapies (e.g., anti-cytokine antibodies (biologics) or different kinase inhibitors) have led to significant improvements in disease control and quality of life of patients, both diseases suffer from high rates of non-response to the approved therapies. It is self-evident that companion diagnostic tests are mandatory to avoid unnecessary exposure to non-efficacious treatments. New diagnostic standards for targeted therapies could include also more complex and expensive sets of markers and standardized models of individual pharmacokinetics as -besides the potential effect on individual affliction and quality of life- an individually optimized treatment would lead to significant reduction of health care costs. Currently, simple stratifying molecular tests can only be performed for certain cancer types, as recurrent mutations in driver genes are unequivocally linked to success of specific treatments (e.g., activating kinase mutations in EGFR and EGFR tyrosine kinase inhibitors (Maemondo et al. 2010; Mok et al. 2009)). Other molecular data layers (transcriptome, microbiome, epigenetic modifications, and single cell heterogeneity) are clearly on the horizon. A systems-oriented, model-based medicine promises to make these layers available for clinical medicine and to provide a prediction of 'multi-factorial' diseases at unprecedented resolution, in a way that clinicians can use the information in their daily decision making.

Challenges

Standardized clinical information for measuring the disease of interest

All algorithms, which make use of complex molecular data sets, usually are trained on simple clinical categories (e.g., response or non-response, clinical scores), which depend on clinical assessment and may differ between national guidelines or even from hospital to hospital. Disease scores used in phase III pivotal trials are often chosen for historical reasons, ease of assessment in a multicentre setting or are tailored to the specific compound. It is important to understand that the same scrutiny, which is needed for developing a molecular method, must also be put into the development of clinical parameters, which measure disease activity. A physician´s assessment of disease (e.g., endoscopic picture in IBD) may be very different from the patient´s perspective (e.g., fatigue as a major symptom of the underlying chronic inflammatory process). Successful models in systems-oriented medicine have to start with a harmonization process for a precise, multi-dimensional definition of disease severity for any given disease under study.

Clinical challenges of standardized production of data and transparent models

It is self-evident that, in order to be transferrable to the clinical situation, all molecular data generation and interpretation must follow a transparent and quality-controlled workflow. Foreseeable specific challenges for clinical translation are:

- The procurement of samples and preparation of biological material (e.g., single cell RNA extraction) are key to standardized results. Consortia efforts must be developed (such as in: (Lappalainen et al. 2013; t Hoen et al. 2013; Alioto et al. 2015)) to harmonize sample extraction standards and to develop feasible ring trial formats for given molecular analysis types. As technologies are developing fast modular certification solutions (example gene panels) needs to be considered.

- Transparent reduction of contents and definition of appropriate marker sets and dynamic models are key for clinical translation. Currently, usually small clinical cohorts are typed for a given molecular data layer (e.g., single cell RNA profiles). The size of the endeavour is defined by academic economic rules (i.e. available budget, track record in the field, chances to obtain future grants). The leading principle is competition, as new results and algorithms are expected to be part of a "successful" scientific story. Publishing negative results (e.g., failure of replication of a previous result) are suppressed by the academic community and journal editors (lack of novelty), despite all public declarations. As an example, more than 50 papers on molecular indicators of anti-TNF response/non-response have been published, none of which have been formally validated in independent multi-centre replication studies or led to a prospective transnational clinical trial. Likewise, head-to-head trials comparing different therapeutic principles stratified by molecular models are necessary, which do not fit into a usual pharma-driven multicentre scheme, as it would require that different companies interact and potentially share a therapeutic space by molecular definition and not by marketing competition. Investigator Initiated Trials which are financed by usual funding schemes of academia (national research agencies, EC) are currently underpowered to achieve this goal.
- Translation of complex information into simple clinical language. Current education of physicians does not convey enough knowledge on the emerging field of systems medicine. With current efforts to transform medical curricula into a more practice-oriented approach, this problem will be further aggravated. It is obvious that neither the increasingly complex molecular background can be taught in depth nor can the comprehensive mathematical or informatics knowledge be fully covered. Yet, it is important to include principles of this potentially disruptive approach into a "normal" medical education (Schultze, consortium, and Rosenstiel 2018). Systems medicine has to be perceived as mainstream for medical students. Most importantly, all disciplines have to develop a common language, which ultimately will also be used to communicate results to patients. Systems medicine results will only be used in clinical practice when results can be interpreted by clinicians and conveyed as a rationale decision making element to the individual patient. This does not only involve the development of intuitive visualization of rigorously tested results, but also insights into molecular analyses and critical appraisal of limitations of models by the physicians. Curricular postgraduate programs, specialty training and faculty programs are needed.

Recommendations

Data integration: See common recommendations for data integration in section "Data and model integration"

Model validation:

R1: Harmonize disease specific scores including objective parameters of disease activity and progression.

R2: Support the development and validation of innovative patient-reported outcome tools for clinical trials including the safe and easy use of app- and wearable based technologies.

R3: Define minimum clinical criteria for a systems medicine trial combined using harmonized scores and quantitative, validated patient reported outcomes into account.

R4: Develop concepts for integrating of model-based prediction and AI content into curricular education in medicine and medical life sciences.

R5: Commence stakeholder discussions including political decision makers to overcome financial and intellectual hurdles in large European consortia trials, ideally involving both academic and European industry partners (current IMI schemes could serve as a starting point).

R6: Develop new models of public-private partnerships for a strong European health care economy on IP participation and protection of mutual interests.

Pharmacokinetic/-dynamic modelling and *in silico* trial simulations

Pharmacokinetic/pharmacodynamic (PK/PD) modelling can usefully translate in vitro, nonclinical and clinical PK/PD data into meaningful information to support decision making during drug development. At individual level, drug PKs can either be described by non-compartmental analysis, compartmental PK modelling or by physiologically-based PK (PBPK) modelling. At population level, population PK (popPK) models have become the most commonly used top-down models that derive a pharmaco-statistical model from observed systemic concentrations. PopPK models are now widely accepted in the context of drug development and regulatory assessment: they have a relatively simple mathematical structure and a high impact in the understanding of the benefits and harmful effects one can expect from a given drug. PopPK models describe the time course of exposure (drug concentrations) to the drug of interest at a population level, accounting for their variability. When mixed-effects approaches are used, model parameters are characterized by a fixed and a random component. Random effects enable us to describe inter-individual variability usually explained by patient's characteristics such as the weight, the age, the sex and the renal function.

In contrast to population PK approaches, PBPK modelling aims to reproduce the physiology of an organism at a larger level of detail. Different organs are explicitly represented in a PBPK model and they are assigned specific physiological properties such as volumes, composition and blood flow rates. The values assigned to model parameters need to be scientifically sound. Of note, PBPK models allow to simulate drug concentration profiles in plasma and different tissues which can be seen as the upstream input of any drug-induced effect.

With regard to personalised medicine PBPK modelling offers a large variety of possibilities due to large level of detail of the underlying model structure. PBPK model building usually starts with the development of a base reference model for an average individual based on mean values of physiological parameters. However, the integration of parameter values from single patients is directly possible, once this information is available. This is in particular supported by the large granularity of PBPK models which may represent physiological information from different levels of biological organization. Therefore, very diverse patient-specific information can be considered ranging from a the absorption, distribution, metabolism, and excretion (ADME) phenotype of a patient (Lippert et al. 2012) to specific pathophysiological alterations at organ level (Edginton and Willmann 2008). There are guidelines by EMA[18] and FDA[19] for model development and model evaluation as well as risk assessment of simulation outputs. Consideration of these guidelines is mandatory for regulators submissions including PBPK simulations.

PK/PD modelling involves on the one hand a quantification of drug absorption and disposition (PK) and on the other hand a description of the drug-induced effect (PD). Of note, PK/PD models (Gerlowski and Jain 1983; Pérez-Urizar et al. 2000) need to cover both drug disposition that drive drug concentrations in different components of the body as well as specific aspects of the disease itself to capture the relevant therapeutic effect. This can be done either using a top-down approach which is very often data-driven. Alternatively, bottom-up approaches can be applied for mechanistic description of drug disposition as well as drug action. PK/PD models and quantitative systems pharmacology (QSP) both aim for mechanistic and quantitative analyses of the interactions between a drug and a specific biological system (van der Graaf and Benson 2011).

Classical PK/PD models largely focus on the pharmacology of single, isolated pathways through the consideration of dose-effect correlations (Derendorf and Meibohm 1999). Vice versa QSP aims at the integration of network models from computational systems biology to capture the effect of a drug at network level (Danhof 2016). The connection between the response variable (biomarker) in PK/PD or

[18] https://www.ema.europa.eu/en/reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation
[19] https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/modeldata-format

QSP models and the clinically relevant outcome is not always straightforward. These models would in these cases need to be complemented by biomarker analyses or dose-exposure-response analyses.

Another important concept to mention in drug effects modelling are semi-mechanistic or physiologically-motivated models. Such reduced models may be preferable if only key aspects of physiological functions need to be described. This can for example become necessary when important parameters are not accessible experimentally, the considered physiological function is too complex, for example in the case of multi-tissue cross-talk at whole-body level. In recent years several semi-mechanistic or physiologically-motivated models have been published which describe amongst others tumour growth (Simeoni et al. 2004), chemotherapy-induced myelo-suppression (Friberg et al. 2002) or cardio-renal crosstalk (Hallow et al. 2018) to name a few.

Building upon individualized PBPK models, QSP models provide a unique possibility to integrate personalised omics data in a whole-body context. It is hence a plausible expectation, that PBPK-based QSP models or similar modelling approaches will in the future be used to represent personalised patient data in digital twins and to optimize therapeutic outcomes of drug treatments. Such approaches may thereby help to overcome the currently prevailing "one-size-fits-all" paradigm in drug treatment through model-informed precision dosing. This includes in particular tailored patient-specific therapies with maximum efficacy yet minimum adverse side effects.

However, successful examples for a standardized import of individual patient data into PK models in clinical practice or even basic research are unknown to date. It should be noted that in the case of PBPK modelling, the integration of patient data from multiple levels of biological organization ranging from the molecular scale to whole-body level into the computational model structure is –from a technical point of view– straight forward. A long-term goal would be a virtual twin, were personalised patient information is integrated into a computational model for an individualized design of treatment schedules.

Already today, PK and PBPK modelling are used for simulations for virtual patient populations in *in silico* clinical trials (ISCT). This term refers to: "The use of individualized computer simulation in the development or regulatory evaluation of a medicinal product, medical device, or medical intervention" (see also Avicenna roadmap[20]). The concept is that computer simulations are proposed as an alternative source of evidence to support drug development to reduce, refine, complement or replace the established data sources including in vitro experiments, in vivo animal studies, and clinical trials in healthy volunteers and patients.

The technical definition of clinical trial simulations (CTs) includes the generation of a response for a virtual subject by reproducing the trial design, the disease progression, the drug and the patient's characteristics and behaviour using mathematical models and numerical methods (Teutonico et al. 2015). The input in an ISCT simulation tools (also called clinical trial simulators) include:

- A disease and drug effects model component (disease model (e.g., system biology model) + drug exposure and effect model (e.g., PBPK, QSP model, agent-based models, etc.)
- A patient intrinsic characteristics component: including the relevant characteristics (biometry, disease stage, co-morbidities, genotypes for relevant enzymes, phenotypes for relevant pathways, etc.) of the patients whose outcome/response need to be simulated
- Trial design component: including other design factors such as dosing regimen, treatment duration, compliance, co-medications, dropouts, etc.

Thanks to appropriate computational methods, this input will permit to simulate/predict predefined response variables in (subgroups of) the target population or in particular types of patients, in order to address different questions (e.g., dose selection, prediction of efficacy or safety, prediction of drug-drug

[20] Avicenna roadmap *in silico* clinical trials (download): https://www.vph-institute.org/avicenna.html

interactions, identification of sensitive subset of patients, etc.). The most relevant ISCT components are discussed below:

Disease and drug effect model component: This component of the ISCT can potentially include various types of models to describe the time course of the disease and/or drug exposure and response. Technically wise, these models include not only the mechanism-based and fully mechanistic models described above: i.e. systems biology/medicine models, PK/PD, PBPK, and QSP models but also semi-mechanistic models (e.g., agent-based models, multi-physic models) and rather empirical and fully data-driven models (e.g., approaches based on machine learning and artificial intelligence). All these models, irrespective of their complexities are developed to address similar applications in the context of drug development or treatment individualization. These include:

- Determination of drug efficacy in special populations (renally or hepatically impaired, children, elderly, rare diseases, etc.) with very limited or efficacy data generation in these target populations.
- Determination of the impact of drug-drug interaction without clinical data.
- Study design optimization: dose selection or power computation for Phase I, II and III studies.

Over the last decades, the added value of *in silico* models (also called digital evidence) in medicine/pharmacology has increasingly been recognized by the scientific community irrespective of their ultimate use/application. *In silico* models are now integrated in different fields in medicine and drug development expanding from disease modelling, biomarker research to assessment of drug efficacy and safety. *In silico* approaches are also expanding in neighbourhood fields such as pharmacoeconomics, analytical chemistry and biology.

Virtual patient generation component: In order for the simulation output to be relevant, the response variables need to be generated for virtual patients who carry the same characteristics as the target population. For this purpose, covariate distribution models play an essential role. They describe patient-specific aspects defining e.g., patient demographic information, baseline disease characteristics, co-morbidity and concomitant medication. In the context of personalised medicine, these details may be considered as a vector containing the patient information. Disease epidemiology registries can be useful to inform this component of the simulators.

Trial design component: This is the so called 'trial execution model'. It represents the design variables of interest in a simulation exercise (e.g., dosing regimens, selection criteria, stratification rules, study duration, etc.). This not only reflects the patient's behaviour and comprises those factors that determine the trial execution features, such as adherence and missing records, but also reflects sponsors', physicians' or investigators' decisions such treatment duration.

While ISCT applications involving pharmaco-statistical PK and PKPD models are now well established and largely accepted, it is expected that the ISCT of tomorrow should include more mechanistic models incorporating available physiological and biological knowledge and capturing the feature of individual patients, and introducing the concept of the patient-specific model. The latter is however considered to still be at its infancy essentially due to the issues described above, that hamper the acceptancy of mechanistic models.

Challenges

Availability of public repositories

PBPK models can in particular be specified to present the PK of various patient cohorts (e.g., paediatrics, elderly patients, different genotypes, hepatically impaired patients, etc.). For simulation of such virtual patient cohorts it is mandatory, that biometric information (gender, age, body mass index) as well physiological information (organ volumes, blood perfusion rates, etc.) is available. At population level such physiological information can be for example taken, for example, from public databases such as the International Commission on Radiological Protection (ICRP) reports. However, such public repositories have not been developed for PBPK models, rather their applicability for such purposes is a side product.

Availability of data

As of now, the availability of adequate data sets (including patient biometry, physiology, phenotypes, genotypes as well as the corresponding electronic health record) is the limiting factor. Generally, patient database provide a useful repository for personalised patient data. However, most of such database are not comprehensive, i.e. they only focus on specific aspects of physiology and missing important aspects of meta information. With regard to PD modelling clinical endpoints, as well as corresponding omics data need to be documented. For the PD effect, challenges from cellular systems biology apply here analogously. For PBPK and QSP models the main challenges are:

- Reliable data sources for 'systems' related parameters are currently limited
- Methods for data generation, collection and integration are not standardized
- Reporting of results are very heterogeneous and inconsistent
- Tools to be used and criteria for model evaluation are very variable across projects
- Very limited platforms ('systems' model) are currently considered reliable and qualified for regulatory submission

In addition, the following challenges are listed in the Avicenna roadmap:

- ISCT is being developed mostly through findings during research projects not specifically targeting the subject.
- The lack of coordinated research and a technological development roadmap prevents the consolidation of the sector and encourages fragmentation.
- The adoption of ISCT requires the active participation of a number of different stakeholders from industry, regulatory agencies, patients' organizations, etc. This requires a balanced, pre-competitive setting where these discussions can be conducted without the risk of any unwanted bias.
- To be effective in a number of diseases ISCT must better predict the systemic responses; but more research is necessary to unravel systemic processes using VPH strategies, systems dynamics models, and the lessons learnt from process design.
- The use of *in silico* methods to translate from animal models to humans is promising in principle, but requires a lot more research and technological development before it can be used effectively.
- The adoption of ISCT requires a significant investment in validation studies to identify those approaches that work reliably, but when conducted publicly and openly, will help to establish some trust among stakeholders.

The development of ISCT is a grand science. Because of its extreme interdisciplinary that can be tackled only in very large research institutes, we need to support their formation, but also explore virtual organization approaches where small groups can join forces and work together to tackle complex problems.

Recommendations

Data integration: See also common recommendations for data integration in section "Data and model integration"

- R1: Enable standardized patient databases to develop personalised PBPK, personalised PK/PD and QSP models that can inform optimal and personalised dosing in patient´s/subgroups (such repositories, e.g., from national cohorts, should ideally encompass all kinds of patient data covering both patient anthropometry and biometry as well as the corresponding meta-information).
- R2: Implement a common standard for such collections to ensure interoperability of data.
- R3: Report clinical endpoints for specific therapies (if available and relevant) in order to develop complementary personalised PD and QSP models (omics data need to be likewise provided to inform the corresponding effect models).

Model validation (see also section: Cellular systems biology):

- R1: Develop appropriate in vitro-in vivo extrapolations from targeted assays to aid in validating PK/PD models (a full validations of PK/PD models requires comprehensive measurements of drug PK as well as of the resulting therapeutic effect).
- R2: Implement an agreement among relevant stakeholders on required criteria for model adequacy for a well-defined context of use.
- R3: Ensure that the standards (to be) used for model development, evaluation, reporting and related decision making are commonly acknowledged by all the involved parties (regulators, HTA agencies, academia, research centres, industry, regulators and patients) and are relevant for all the types of models that can be used[21].
- *R4:* Moreover, the requirements for adequate implementation of the two other components (patient´s and other design characteristics) should also be standardized and discussed by the involved stakeholders.

[21] Of interest is the standard recently published by the American Society of Mechanical Engineers (ASME) on assessment of credibility of computational modelling through Verification and Validation, applied to medical devices (V&V40) which was proposed to be extended to other biomedical products such as drug and combined medical products (Viceconti et al. 2020). In the current situation, a similar initiative oriented to disease description and drug development would be of great value.

Artificial Intelligence approaches

Data-driven approaches, with AI and ML as the most prominent concepts, treat the mechanism as unknown and aim to model a function that operates on data input to predict the outcome, regardless of the unknown physiological processes. The mechanisms operating in the complex systems being modelled, i.e. which factors together drive outcomes, are considered too complex to be determined (*black-box* models). The quality of *black-box* models is assessed through the accuracy of their predictions, tested in a variety of ways. These data-driven models can be applied in a hypothesis-naïve way, made as to which factors drive the causal mechanism. Such models perform well with large inputs of all potentially significant data. AI/ML approaches provide an opportunity to move away from current interpretive attempts to apply group associations and to predict responses in individual patients. Group associations may be confused with prediction, and any group result must be interpreted in a largely unstructured way to the treatment of an individual. Thus, classical group-based clinical studies and AI/ML approaches represent two distinct paradigms: the first based on patient groups, while the second, AI/ML approach, builds a prediction based on the individual patient.

ML attempts ever more accurate clustering or classification, to a high level of accuracy, and a high level of confidence. Different types of ML are employed as appropriate according to input data and aim of analysis, i.e. according to the clinical question. Examples of state-of-the-art *in silico* risk models of relevance for personalised medicine, using varying input data with different results are given below. As input data and research questions are ever-expanding, these can only be examples and not a comprehensive list.

- Population-wide analyses of longitudinal disease trajectories identifying disease associations, comorbidities and dispositions on the basis of clinical or registry data to identify associations driving mechanistic hypotheses in future studies.
- PRS to assess both relative and absolute risks of diseases, and exploration of polygenic risk scores' association with phenotype or exposome.
- Estimating heritability and genetic correlations from large health data sets in the absence of genetic data.
- Pathway enrichment analysis using data fusion techniques for data interpretation and hypothesis generation, as well as discovery of enriched pathways across multiple data sets (clinical and genetic) to highlight associated genes. Integrative methods can be used to identify genes, mutations and prognostic molecular pathways in disease.
- Co-occurrence analysis of symptoms and diseases using terms of text-mining from clinical notes.
- Sub-grouping of conflicting associations for better understanding of phenotype and exposure data, such as dietary patterns and changes of anthropomorphic traits, using combined stratification and association discovery approaches. Conflicting rules of association adds complexity when finding genetic associations
- Identification of immune drivers through distribution profiling of the adaptive immune system.
- Assessment of metabolite association with comorbidities through linear and logistic regression, self-organizing maps and trajectory clustering.
- Neural network-based prediction of signal peptides from amino acid sequencing, optimized to handle incorrectly annotated input data; these results can contribute to disease understanding, to precision medicine and to drug development or repurposing.
- Prediction of receptor specificity for all members of a given protein family for better understanding of biological systems, including the immune system, harnessing deep artificial neural networks to integrate information from individual receptors into a single pan-receptor model.

ML approaches learn the theory automatically from the data through a process of inference, model fitting, or learning from examples (Baldi and Brunak 2001). ML can be supervised, unsupervised, or partially supervised. Unsupervised learning has the potential to take all features into account and identify clusters for precision medicine. Unsupervised learning also comprises dimensionality reduction, permitting feature elicitation, compression and Big Data visualization, all of which can allow for better understanding of big medical data and the factors driving disease initiation and progression. For general

understanding of the relationships between disease and cofactor, unsupervised learning offers the ability to discover new knowledge. More importantly, for the individual patient, unsupervised learning can uncover previously unrecognized phenotypes which can then be used to refine outcome prediction. In supervised learning, the model is supplied with (labelled) input features that are considered when predicting a predetermined outcome from new data, either through regression (for continuous results such as number of days or months before disease debut) or classification (for discrete results such survival/death, or for image classification). Semi-supervised learning employs both labelled and unlabelled data for training. Accuracy of ML results is verified using independent test sets. ML of the reinforcement type allows the algorithm to react to an environment, and is related to the concept of AI, permitting real-time decisions and individualized predictions. Supervised AI clinical decision support tools allow AI the possibility to tailor prediction to the individual patient to the extent that bias imparted by data summary. In this regard, AI predictions incorporating full data sets of disease comorbidities provide better prediction of patient outcome than those relying on bundled or synthesized comorbidity indices (Westergaard et al. 2019; Nielsen et al. 2019). A fundamental part of incorporating supervised AI/ML into patient care will be the creation of a consensus among clinicians and data scientists on how data inputs should be structured. The danger is that clinically accepted associations, especially those obtained through studies from patient groups will be incorporated into the AI models. This predetermined data organization can bias and limit the AI model.

Challenges

Input Data (see also recommendations for data integration):

Imprecise reporting: A significant challenge is confused reporting, which makes it difficult to harvest the full benefit of results, navigate biomedical literature and generate clinically actionable findings (Varga et al. 2020, in review) A study evaluating adherence to the TRIPOD Statement, a collection of guidelines for communicating results from multivariate prediction models (Collins et al. 2015), reported that only 2% of all examined articles had satisfactory reporting in their abstracts, and only 5% had satisfactory reporting in their titles (Heus et al. 2018).

Non-harmonized data: Data standardization is a major challenge, as most *in silico* methods require comparable input data. Moreover, validation of the models requires analysis of an independent test set, e.g., comparable data from a hospital in a separate geographical area, which also requires that secondary data (i.e. data produced for another purpose than the research at hand) be harmonized and used as comparable input data.
Some European countries provide insights into how the challenge of standardization of health data can be met, and their lead can be followed and improved upon. In the Nordic countries, population-based biorepositories and data from 8-10 million individuals are currently available for research. These countries have a long history of integrated healthcare and patient registries and biobanks (Njolstad et al. 2019) and have implemented systems facilitating data access and promoting re-use of data (e.g., Findata.fi[22], operating under the Ministry of Health, which aims to be a "one-stop shop for the secondary use of health and social data."

Inappropriately pre-processed data: Data based on group associations, or pre-determined understanding of clinical relationships, may bias and limit AI/ML predictions. In the ideal system, data inputs are provided to the ML model free from linkages. Certainly, associations resulting from grouped data should not be prioritized for input, since they often prove to be poorly predictive. In addition, group trial data assumes that the individual patient for whom the AI model is intended, shares the important attributes of the group previously studied.

Data lockdown: Proprietary systems are still a challenge for personalised medicine. *In silico* models for risk prediction have the ability to harness diverse data sources, which may or may not have significance for disease risk. At the same time, health care IT cannot integrate all modalities into one system; hence, it is necessary to have multiple systems which can fulfil the needs of subspecialties and be designed or acquired flexibly and rapidly. Sub-specialized imaging systems are an example, but something as common as electrocardiogram technology is another. These subsystems, often integrated software and hardware solutions, produce frequent challenges of data extraction, labelling, interpretation and standardization. The commitment to reasonable data extraction possibilities and service should be a component of any sale of health care technology within EU countries.

[22] https://www.findata.fi/

Legal issues: As discussed elsewhere it is imperative to find lawful routes for sharing data. As mentioned above, a full discussion of these issues is found in the EUSTANDS4PM paper "Legal and Ethical Review of *in silico* modelling" (Ó Cathaoir K. 2020).

Model Validation:

Testing: Model validation of AI/ML models is part of the development process. Often it is referred to as *testing*, while the algorithmic creation is called *training.* With model testing, one determines how well the used method performs for the given data set. Normally one would consider different ML methods and compare them by their test results. For internal validations, the model is tested against its underlying data. The most common techniques are distinguished by the division of the data set into train and test sets. In the most basic form, the complete data set is used to train and to test the model (*Apparent Validation*). This naturally leads to an optimistic impression of the model performance i.e. over-fitting; if not, this could indicate a faulty model training, e.g., a non-fitting algorithm. Other approaches aim to reduce the bias by randomly splitting the data set into training and test sets. The algorithm should be robust enough to produce a predictive model even if it is trained with a subset of the data. A split into 50% - 67% training data and 50% – 33% test set, respectively, is referred to as *Split-Validation*. A more reliable approach is *Cross-Validation*. Where the data set is split into N random groups for an N-fold validation. N-1 of the N subsets are used to train the model while the leftover set functions as the test data. This is repeated N times, so that each group is the test set once (*N-fold cross-validation*). To improve the stability of the result, this process can be repeated k-times, whereby different groups are formed in each process (*k x N-fold cross-validation*). *Bootstrap validation* provides an alternative approach; instead of taking a distinct subset from the sample S, a same size group S' is formed by selecting random data points with repetition. Theoretically, but unlikely, the bootstrap group could be exactly the same as the sample or could contain only one data point several times. S' functions as the training set while S - S' is the test set. The process can be repeated several times to improve the stability of results. When handling incomplete data sets, one can consider to estimate the missing values either in the original data set or in the random subsets. While the latter approach is expected to give a better impression of the method, it is also computationally harder. With those validation measures, one decides for the best method to build a model from the original data set. The best method is not necessarily the one that fits the training data perfectly i.e. over-fitting, but rather the one that can predict the outcomes of an unknown data set (i.e. test set) the most accurately. The developed models can also be validated by dividing the data manually instead of selecting a subset randomly from the original data set. This could be done by dividing the data set into newer and older cases before the model development. This is considered a *temporal validation*. Another option could be a *geographical validation,* in which the model is tested for other cohorts or hospitals. Even though this could be interpreted as non-random cross- or split-validation, it is a crucial step in ML model validation. The differences in how or when hospitals retrieve the patient data can have a high impact on the model prediction.

Recommendations

Data integration: See also common recommendations for data integration in section "Data and model integration"

For clinical decision support using AI, input data should be as comprehensive and unbiased as feasible. While there is a temptation to prioritize data that have been shown to be associated with an outcome from group analysis, such a prioritization biases the model and may amplify the error that associations are predictive of outcome. Similarly, ranking significant findings from group trials as more important to the AI model than other data amplifies the incorrect assumption that the individual patient can be fairly represented by an average patient of a group trial. Finally, the application of usual clinical wisdom in an attempt to pre-process information for the AI model may also result ignoring important data sets. In this regard, the application of older comorbidity indices (e.g., Elixhauser and Charlson) rather than offering the AI model access to the comorbidities provided in data sets such as the Past Medical History of the clinical record, or population disease registries, significantly limits the predictive power of the model (Nielsen et al. 2019; Moseley and Brunak 2019). A more rational process to align AI input may be to develop an EU-wide list of what data sets are accessible and fulfil basic criteria, across the Union. In this way, comparisons can be made in the least processed fashion and avoiding as much interference with the AI inputs as possible.

R1: Data should be as comprehensive and unbiased as feasible, rather than using selected data based on potentially flawed existing knowledge
R2: Data should be available in both, processed and unprocessed form.

Model validation:

R1: Model validation should involve a 3-phase process.
R2: In the first phase, outcome prediction for the AI model should be compared to standard measures that employ best clinical practice and using established evidence-based hierarchy which takes into account historical practice, clinical trials and systematic review/meta-analyses (Murad et al. 2016).
R3: If the AI model shows a predictive advantage, a second phase of validation could involve use of a High Confidence Off Policy Evaluation as previously described (Komorowski et al. 2018) where the AI model is compared to clinician decision making.
R4: If beneficial, phase 3 provides clinicians access to AI prediction points.

Annex

Common standards relevant for personalised medicine

Table 1: Common standards relevant for personalised medicine and *in silico* approaches.
Examples of common standards that have been developed by specific user communities and different stakeholders. Their use has been enhanced as they have been coupled to tools which have spread in the respective field of research. In addition, a current overview about data formats and standards for in-silico systems biology and quantitative modelling can be found in (Golebiewski 2019) and as a comprehensive reference in the annex of ISO 20691 (in preparation).

DNA, RNA, protein sequence formats	
FASTA	Widely used for representing nucleotide sequences or amino acid, developed for use in the FASTA program (Lipman and Pearson 1985; Pearson and Lipman 1988). The FASTA format is simple and lacks facility for extensive annotation.
Sequence Alignment/Map (SAM) and Binary Analysis Map (BAM) format	Capture of sequences that have been aligned to a reference genome. SAM is a tab delimited text format consisting of a header section, which is optional, and an alignment section. BAM is in a binary more condensed version while SAM has the same information in a series of tab delimited ASCII columns (Li et al. 2009). BAM files are compressed files.
CRAM	A compressed columnar file format also used for storing biological sequences mapped to a reference sequence, it has been developed to improve compression and hence save on storage costs (Hsi-Yang Fritz et al. 2011).
ISO/IEC 23092 (MPEG-G): Information technology — Genomic information representation	The ISO/IEC 23092 (MPEG-G) series of standards is a coordinated international effort to specify a compressed data format that enables large scale genomic data processing, transport and sharing. Interoperability and integration with existing genomic information processing pipelines is enabled by supporting conversion from/to the FASTQ/SAM/BAM file formats. It consists of currently (as of October 2020) six parts: Part 1: Transport and storage of genomic information Part 2: Coding of genomic information Part 3: Metadata and application programming interfaces (APIs) Part 4: Reference software Part 5: Conformance Part 6: Coding of genomic annotations
General feature format (GFF)	Stores DNA, RNA or protein genetic sequence data (Akanksha Limaye 2019). It stores the whole sequence for the relevant feature.
Variant call format (VCF)	A text format file storing the same data but only contains the sites which differ from a given reference and hence is more space efficient than GFF (GitHub_Community 2020). Originally designed to be used for SNPs and INDELs but can also be used for structural variation. A Variant represents a change in DNA sequence relative to some reference. For example, a variant could represent a Single Nucleotide Polymorphism (SNP) or an insertion. Variants belong to a VariantSet. This is equivalent to a row in VCF.

Binary variant call format (BCF)	A binary version of VCF and therefore is more space efficient, the relationship between BCF and VCF being similar to that between BAM and SAM.
Synthetic Biology Open Language (SBOL)	An RDF/XML format for representing, among other things, sequences for genetic circuit designs. It has a rich ability to express both sequence feature annotations and part/sub-part relationships. It is also designed to represent incomplete/partial sequences and relative ordering of parts in a genetic design.
Mass Spectrometry	
mzML	Stores the spectra and chromatograms from mass spectrometry in and eXtensible Markup Language (XML) format. Now a well-tested open-source format for mass spectrometer output files that is widely used (Martens et al. 2011).
mzTab	A more easily accessible format which could be used with R or Microsoft Excel tools in the field of proteomics and metabolomics. mzTab files can contain protein, peptide and small molecule identifications. In addition experimental meta-data and basic quantitative information (Griss et al. 2014).
Medical imaging, Digital Imaging and Communications in Medicine	
Digital Imaging and Communications in Medicine (DICOM)	Dominating standard used in medical radiology for handling, storage, printing and exchanges of images and related information. Specifies the file format and communication protocol for handling these files. Captures pixel data making up the image and how the image was generated (e.g., used machine and protocol, information regarding what patient the image is capturing. Living standard regularly maintained and modified (DICOM_Secretariat 2020), also adopted as ISO 12052 "Health informatics - Digital imaging and communication in medicine (DICOM) including workflow and data management".
The European Data Format (EDF)	A standard to archive, share and analyse data from medical time series (Kemp et al. 1992).
Semantic integrations	
BRIDG (Biomedical Research Integrated Domain Group Model)	An information model being used to support development of data interchange standards and technology solutions to enable semantic (meaning-based) interoperability within the biomedical/clinical research arena and between research and the healthcare arena. BRIDG is a collaborative effort engaging stakeholders from the Clinical Data Interchange Standards Consortium (CDISC), the HL7 BRIDG Work Group, the International Organization for Standardization (ISO), the US National Cancer Institute (NCI), and the US Food and Drug Administration (FDA). The goal of the BRIDG Model is to produce a shared view of the dynamic and static semantics for the domain of basic, pre-clinical, clinical, and translational research and its associated regulatory artifacts. The BRIDG Model is a hybrid of conceptual and logical models represented as UML Class Diagrams. It was built by harmonizing other project and domain models and each concept in the

	BRIDG model carries its provenance in the form of mapping tags indicating what data elements from other models map to that concept.
HL7 FHIR (Fast Healthcare Interoperability Resources)	A standard for exchanging healthcare information electronically.
Human Phenome Ontology (HPO)	Developed by the Monarch Initiative a consortium, carrying out semantic integration of genes, variants, genotypes, phenotypes and diseases in a variety of species allowing powerful searches based on ontology. HPO is a standardized vocabulary of phenotypic abnormalities associated with disease. Standard terminology for clinical "deep phenotyping" in humans, providing detailed descriptions of clinical abnormalities and computable disease definitions (Shefchek et al. 2020). The primary labels use medical terminology used by clinicians and researchers. These are complemented with laypersons synonyms. HPO is one of the projects in the Global Alliance for Genomics and Health (GA4GH) seeking to enable responsible genomic data sharing within a human rights framework (GA4GH_Community 2020).
SNOMED CT	The Systematized Nomenclature of Medicine (SNOMED) is a family of medical terminology systems. Originally conceived as a nomenclature, the latest version SNOMED CT can best be characterised as an ontology-based terminology standard. The goal of all SNOMED versions is to provide a language that represents clinical content as clearly and precisely as possible, regardless of its original language. This should enable search queries to be answered with high recall and high precision (see also www.snomed.org).
LOINC	The Logical Observation Identifiers Names and Codes (LOINC) is a database of common names and identifiers used to identify laboratory and clinical examination and test results. The aim is to facilitate the electronic exchange of data when transmitting medical examination results and findings data. LOINC is recommended (also by HL7 and DICOM) for the exchange of structured documents (CDA) and messages (see also https://loinc.org/).
Serial ISO/IEEE 11073	Personal Health Data (PHD) Standards, a group of standards addressing the interoperability of personal health devices (PHDs) such as weighing scales, blood pressure monitors, blood glucose monitors, etc. (see also: http://11073.org/).

Models and modelling tools	
CellML	A standard based on XML markup language (Lloyd, Halstead, and Nielsen 2004) used for storing and exchanging computer-based mathematical models allowing sharing of models even when different modelling tools are used (Schreiber et al. 2016). CellML is a description language to define models of cellular and subcellular processes and supports component-based modelling, allowing models to import other models, or subparts of models, therefore strongly encouraging their reuse and facilitating a modularized modelling approach. A CellML model typically consists of components, which may contain variables and mathematics that describe the behaviour of that component. The mathematical model is considered to be the primary data and biological context is provided by annotating the variables and equations with metadata using the Resource Description Format (RDF).
The Systems Biology Markup Language (SBML)	A standard model interchange languages that permits exchange of models between different software tools (Hucka et al. 2018). SBML is a machine-

	readable, XML (Extensible Markup Language) based model description and exchange format for computational models of biological processes. Its strength is in representing phenomena at the scale of biochemical processes, but it is not limited to that. The evolution of SBML proceeds in stages (levels). Since SBML Level 3 the format is modular, with the core usable in its own right and packages being additional "layers" that add features to the core. SBML core is suited to representing such things as classical metabolic models and cell signaling models. SBML packages that extend the core and are optional in their use, add additional model features, such as visualizations, distributions, constraint-based models (flux balance constraints), hierarchical model composition, special processes or grouping of elements.
The Synthetic Biology Open Language (SBOL)	A standard to support specifications and exchange of biological design information (Madsen et al. 2019). SBOL Data provides both an electronic format for representing this information, while SBOL Visual provides schematic glyphs to graphically depict genetic designs.
Simulation Experiment Description Markup Language (SED-ML)	A machine-readable, XML (Extensible Markup Language) based format for encoding the description of a computational simulation. Developed to capture the Minimum Information about a simulation experiment (MIASE), the minimal set of information needed to allow reproduction of simulation experiments (Waltemath et al. 2011; Schreiber et al. 2019). Typically used with an XML-based model description format (e.g. CellML or SBML), SED-ML allows for the description of applying a numerical algorithm to a mathematical model in order to perform a given task. Tasks may be nested to allow the composition of relatively simple tasks into increasingly complex simulations. Mechanisms exist in SED-ML to apply pre-processing steps to a model prior to executing a simulation task and also to apply post-processing to the raw simulation results (Nickerson et al. 2016).
Open Modelling EXchange format (OMEX)	OMEX supports the exchange of all the information necessary for a modelling and simulation experiment in the life sciences. An OMEX file is a ZIP container that includes a manifest file, an optional metadata file, and the files describing the model. The manifest is an XML (Extensible Markup Language) file listing all files included in the archive and their type. The metadata file provides additional information about the archive and its content. Although any format can be used, an XML serialization of the Resource Description Framework (RDF) is recommended (Bergmann et al. 2014).
NeuroML	XML-based standardized model description language to describe mathematical models of neurons and complex neuronal networks (Goddard et al. 2001). The focus of NeuroML is on models which are based on the biophysical and anatomical properties of real neurons.
PBPK/PD	Physiologically based Pharmacokinetic/Pharmacodynamic models allow a mechanistic representation of drugs in biological systems (Kuepfer et al. 2016).
Pharmacometrics Markup Language (PharmML)	A machine-readable, XML (Extensible Markup Language) based model description and exchange format used for encoding computational models, associated tasks and their annotation as used in pharmacometrics. It provides the means to encode pharmacokinetic and pharmacodynamic (PK/PD) models, as well as clinical trial designs and modelling steps (Nickerson et al. 2016).
Human Physiome Field Markup Language (FieldML)	A machine-readable, XML (Extensible Markup Language) based model description and exchange format for representing hierarchical models using generalized mathematical fields. FieldML can be used to represent the dynamic 3D geometry and solution fields from computational models of cells, tissues and organs (Nickerson et al. 2016).
Biological Pathways Exchange (BioPAX)	A machine-readable standard format that aims to enable integration, exchange, visualization and analysis of biological pathway data.

Numerical Markup Language (NuML)	A machine-readable, XML (Extensible Markup Language) based format for describing and exchanging multidimensional arrays of numbers to be used with model and simulation descriptions.
Analysis pipelines	
ISO 25720	Genomic Sequence Variation Markup Language (GSVML). The standard is applicable to the data exchange format that is designed to facilitate the exchange of the genomic sequence variation data around the world, without forcing change of any database schema, based on XML.
ISO/TS 20428	Data elements and their metadata for describing structured clinical genomic sequence information in electronic health records. The specification defines the data elements and their necessary metadata to implement a structured clinical genomic sequencing report and their metadata in electronic health records particularly focusing on the genomic data generated by next generation sequencing technology.
ISO/DIS 21393 (in preparation)	Omics Markup Language (OML). OML is a data exchange format designed to facilitate exchanging omics data around the world without forcing changes to existing databases.
ISO/DTR 21394 (in preparation)	Health informatics — Whole Genome Sequence Markup Language (WGML)

Recommendations to key actors

Info box 5: General recommendations to different key actors from the EU-STANDS4PM paper "Towards standardization guidelines for in silico approaches in personalized medicine" (Brunak et al. 2020)

Funders, including the EU-Commission: Key requirements of any grant funding for personalized medicine projects should be that: (i) Grant recipients make algorithms and pre-processed project data available to the community and (ii) algorithms are accompanied by documentation and follow approved standards. Standardization efforts shall also be fully fundable to ensure that appropriate and sufficient resources are made available to the scientific communities for developing standards that the researchers then could apply consistently to their workflows. This ensures establishment of standards that reflect best practice in their domain. Data processing, documentation, and subsequent sharing thereby become integral, obligatory deliverables of funded projects, included in the budget and planning. Data sharing and documentation thereby become less onerous than currently, where they are un-funded and altruistic.

Health care providers purchasing and developing electronic health care systems: State organizations purchasing health care systems should make data harvesting a criterion for system developers and providers. Many providers regard both the data produced and the algorithms involved as proprietary and create closed systems where analysis of data proceeds internally with key limitations in how data analysis can be performed. This is for example the case with some providers of Electronic Patient Record systems, where the business model seems to work against open systems. Instead we suggest that tools should be shared even across countries, health care providers, and with academic or industrial stakeholders involved in health data science. The negotiation power necessary to enforce harvestability of these data might arise only as a consequence of legislation making it compulsory.

Journals: A requirement of publication should be processed data deposition in recommended, preferably open data repositories. Where the nature of the data is such that deposition is not legal/ethical, a description of the data should be catalogued in such a repository. Restricted access models will also in many cases be needed and desirable.

Research groups: Documentation and data sharing tasks should be included in the preparation of grant applications for projects in the form of a data management plan. Once the project is initiated, documentation should be prioritized when pre-processing data to make it possible for others to re-use processed data. Algorithms should follow available standards unless there are clear reasons why not to use such existing standards. The advantage of being cited for re-use of pre-processed data and algorithms should be a focus point. Transparency and compliance with standards for algorithms and data should be a key quality parameter when assessing both one's own work, and work received for peer review.

National and regional health data providers: Options for sharing of pre-processed data originally provided by these actors should be facilitated. For example, the Health Data Authorities could provide a repository for pre-processed data and scripts, stipulating that the researchers having done the pre-processing must be credited in work building upon it. Too often, users are handed poorly annotated data requiring cleaning in the same way leading to substantial duplication of effort. Guidelines for returning clean and value-added data to data providers should be encouraged.

Policy makers: To ensure best adaptation and acceptance of mandatory standards, regulatory and governance bodies, as well as other policy makers releasing and monitoring such standards should work closely with the scientific communities when establishing official standards. The need for greater clarity regarding the scope of legal standards related to personalized medicine is clear. Treaties and recommendations should be reconsidered in light of big data driven healthcare. Yet, even newer legislation, namely the GDPR, is open to interpretation and national deviation, which can leave researchers and individuals unclear regarding processing of personal data. This should be addressed through legal guidance from, for example, the European Court of Justice. Furthermore, there is a need for greater transparency within the healthcare system regarding use of data for research, including informed opportunities to opt out of secondary use and information on data ownership. Governments should ensure that individuals are adequately protected from misuse of their data, including through proportionate fines. Although scientific research is vital, the individual's rights continue to weigh higher in international bio law.

Case Study Collection

Project	**BD2Decide: Big Data and Models for Personalized Head and Neck Cancer Decision Support (http://www.bd2decide.eu/)**
Methods	a) RNA: STAR/DeSEQ2 b) DNA methylome: Bismark/RnBeads c) Microbiome: Mothur/Qiime/Humman2/Metaphlan d) Marker identification and reduction to diagnostic sets are done by different ML approaches (random forest , Bayesian inference)" e) Machine learning (unsupervised) f) Omics data analysis g) Omics data analysis
Model function	a) Combined methods making survival prediction based on clinical factors in HNC process
Input data	a) Clinical and pathological b) Clinical and pathological c) Clinical, pathological, genomics (transcriptomics) and radiomics d) Clinical, pathological, genomics (transcriptomics) and radiomics e) Tissue samples f) Imaging data
Output	a) Survival prediction Purpose: To assess the impact, in terms of survival, of each clinical factor involved in the HNC process. b) Patient/cohort classification. Purpose: To identify relations among different variables that apparently are not related. c) Patient/cohort classification Purpose: To identify variables that are correlated to a certain group of population (new patterns). d) Patient/cohort classification Purpose: To uncover significant indicators associated to patient cohorts. e) Transcriptomic profiling Purpose: To deal with high amount of genes and discover relations between genes and patient cohorts. f) Radiomics profiling Purpose: To deal with high amount of radiomic features and discover relations between radiomics and patient cohorts.
Processing steps	a) Univariate (Log-rank test), Multivariate Cox Model, Survival trees b) Logistic regression, Support Vector Machine, Random Forest c) K-means, Birch, Ward, Spectral Cluster d) Principal Component Analysis, Independent Component Analysis, Non-negative Matrix Factorization e) RNAseq (next-generation sequencing) f) Radiomic feature extraction

Use cases	a-f) Patients affected by TNM stage III-IV head and neck cancers
Scale (tissue, organ, cell etc.)	a) All b) All c) All d) All e) Sub-cellular & Tissue scale f) Tissue scale & Organ scale
Challenges/Benefits/Limitations, Input Data standardisation	a) In order to avoid issues on data standardization, due to data comes from different hospitals, well-defined protocols and data cleaning processes have been adopted, inspired also in previous works. b) Laboratory problem were solved centralizing genomic tumour tissue samples analyses in only one clinical centre. c) Technical approaches has been applied to solve the problem of sharing data in various centres and to compare the data with other projects (like RARECAREnet). Dedicated and private services, and integrated approaches has been applied. d) IT restrictions in each hospital to share and collect the data were addressed to allow the usage of secure services in each centre. e) Ethical and data protection regulation has been followed-up to allow the correct use of the data management during the project execution.
Existing standards (formats, guidelines, ontologies)	a) BD2Decide ontology (mapped with external ontologies such as SNOMED-CT or ICD10). Also mapped with Gene ontology. b) Ontology was based on previous project (NEOMARK). c) Within the Decision Support System, as part as the Knowledge Management System, a set of rules has been defined to be used in future projects as guidelines. d) Ethical issues force to create informed consent to obtain the authorization by Ethical committee in each centre, complying each legal standard.
Model validation	no information available
Lessons and Comments	no information available

Project	**Computational Horizons In Cancer (CHIC): Developing Meta- and Hyper-Multiscale Models and Repositories for In Silico Oncology (http://chic-vph.eu/project/)**
Methods	no information available
Model function	-Adapting a Four Dimensional Nephroblastoma Treatment Model to a Clinical Trial Case Based on Multi-Method Sensitivity Analysis (Georgiadi et al. 2012) -The Technologically Integrated Oncosimulator: Combining Multiscale Cancer Modelling with Information Technology in the In Silico Oncology Context (Stamatakos et al. 2014).
Input data	Tomographic imaging data, clinical data, molecular data, pathology data
Output	Response to neoadjuvant treatment
Processing steps	Imaging data postprocessing, molecular data postprocessing, pathology data postprocessing
Use cases	Nephroblastoma
Scale (tissue, organ, cell etc.)	Multiscale heterogeneous data (clinical, imaging, molecular, pathology)
Challenges/Benefits/Limitations, Input Data standardisation	Poor standardization of DICOM data from different MRIs using different protocols, postprocessing of imaging data is needed to render the tumour, no automatic tools are available, combining different hypomodels to one hypermodel.
Existing standards (formats, guidelines, ontologies)	no information available
Model validation	Will be done by comparing the predictions with reality (imaging data after neoadjuvant chemotherapy and pathology data after surgery)
Lessons and Comments	This is done in close collaboration with Prof. Dr. Stamatakos and his group from ICCS, National Technical University of Athens, Greece. This shows that multidisciplinary approaches and team work are needed, including a legal and ethical framework. Sustainability is of importance after funding periods. The Medical Device Regulation needs to be taken into consideration if the model is to be used in clinical care.

Project	**iPlacenta: Integrative placenta: A systems biology approach towards phenotype-specific interactomes for placental function (https://www.iplacenta.eu/)**
Methods	a) Molecular interaction map construction and analysis; supervised ML b) In vitro cell work: expression of specific proteins/modified proteins in cells, differentiation of pluripotent stem cells into trophoblasts c) Use of small animal model for assessment of vasculature and endothelial function with the newly developed devices d) Doppler ultrasound in pregnancy women and collection of blood sample after the delivery. e) Molecular laboratory techniques, data analysis, bioinformatics f) Expression-Quantitative Trait Loci analysis by combining genotype and gene expression datasets from two cohorts of human placental samples. g) Recruitment patients and performing non-invasive CV assessment; supervised BT and AK
Model function	a) Identification of regulatory motifs; risk prediction b) Uncover mechanistic roles of redox modifications in angiogenic signalling and assess their functional effects in early pregnancy events c) Literature, in vivo data animals, knowledge industrial partner d) Identification of a epigenetic marker in maternal blood e) identification of senescence markers and disease progress f) Identification of eQTLs, of regulatory hot-spots g) Identification of abnormal CV findings; risk prediction
Input data	a) Primary Literature, public databases; clinical data (hospitalomics) b) Gene and protein expression, functional assays, proteomics, splicing micro-array c) Prototypes, in vivo validated d) Primary Literature, clinical data (from patients) e) Clinical and experimental data f) Genotype data (Acquired with InfiniumOmniexpress Illumina Array, from DNA samples) and gene expression data (Acquired with ClariomD microarray from Affymetrix, from RNA samples) g) Clinical data, biophysical and biochemical data
Output	a) Therapeutic drug target, biomarker, improved micro-arrays, knowledgebase; patient stratification & classification b) New insights on signaling pathways and cell functions c) Prototype development process d) Possibility of predicting postnatal neurological damage when the fetus is in utero. e) Understanding pathophysiology of adverse pregnacies affected by placental ageing. Biomarker identification and Senolytic therapeutics f) list of gene-snps that show statistically significant correlation and could indicate potential regulatory mechanisms in the placenta g) Patient stratification & classification; prediction; improving CV health in women
Processing steps	a) Various approaches b) --- c) Endothelial function of rodents, placenta vasculature in murine pregnancy

	d) Doppler ultrasound/Collection of blood samples - Analysis of samples/Analysis of results e) Various approaches f) Filtering and quality control of genotype and gene expression data g) Various approaches
Use cases	a) Pregnancy complications, here: preeclampsia intrauterine growth restriction b) Pathologies involving oxidative stress and angiogenesis, here preeclampsia c) Organ d) Pregnancy complications: late onset intrauterine growth restriction e) Pregnancy complications, here: preeclampsia intrauterine growth restriction f) Placental function pregnancy complications, here: preeclampsia intrauterine growth restriction g) Hypertensive disorders of pregnancy
Scale (tissue, organ, cell etc.)	a) Multiscale from placenta at the highest level to molecule on the lowest level; b) Molecules (single protein functions), cell (gene expression and functional effects, potentially interaction between several cell types in co-cultures) c) Animal study under home office license Comparison with other developments (similar technique) d) Multiscale (blood sample, extraction of micro-RNA) e) Multiscale from placenta at the highest level to molecule on the lowest level; f) Using DNA and RNA from the whole placenta (organ) g) Assessment of CV system and maternal heart
Challenges/Benefits/Limitations, Input Data standardisation	a) Dependency on results of partner projects results; Ultrasound image anonymization; hospital database access (law restrictions) b) Challenges/limitations: translation from molecular to tissue/organ scale, lack of physiological relevance for immortalised cell lines, questions towards precise identity of iPSC-derived trophoblasts; Benefits: precise focus on specific molecular pathways, easy access to and maintenance of cells c) --- d) Limitations: -Recruitment of patients and blood samples, follow-up of patients. -Benefits: possibility of predicting postnatal neurological damage when the fetus is in utero. -anonymization of the clinical data e) Sample collection and sample size; limited access to hospital data f) Main limitations are due to the small number of samples available. g) Recruitment and follow patients up/offering extra-care for patients/observational study
Existing standards (formats, guidelines, ontologies)	a) SBGN PD & AF (CellDesigner mix), GO, ChEBI, UniProt b) Published literature c) --- d) --- e) SBGN PD & AF (CellDesigner mix), GO, ChEBI, UniProt

	f) Study design based on previous studies found in the literature, integrating different approaches to better tackle our own dataset g) Not applicable
Model validation	a) Experimental validation, expert curation/validation b) Experimental in vitro validation with complementary studies, animal studies c) --- d) Experimental validation e) Experimental validation, expert curation/validation f) Overlap of findings with previous eQTL analyses in placenta, experimental validation g) External validation
Lessons and Comments	a) iPlacenta is an interdisciplinary project/training network that investigates pregnancy complications from with various approaches. Thus, the methods, data, standards and validation depend heavily on the sub project. Here we only list the information for b) --- c) --- d) iPlacenta is an interdisciplinary project/training network that investigates pregnancy complications from with various approaches. Thus, the methods, data, standards and validation depend heavily on the sub project. e) iPlacenta is an interdisciplinary project/training network that investigates pregnancy complications from with various approaches. Thus, the methods, data, standards and validation depend heavily on the sub project. f) --- g) iPlacenta is a great network among ESRs and among research teams all around Europe. I am very happy to be part of this.

Project	**LifeCycle: EU child cohort network (https://lifecycle-project.eu/)**
Methods	Regression analyses; Omic studies
Model function	no information available
Input data	Data collected in ongoing population-based cohorts studies in pregnancy and childhood (questionnaire, physical examinations, biomarkers, imaging, omics)
Output	Harmonized data for core exposure, covariate and outcome variables
Processing steps	Different approaches
Use cases	no information available
Scale (tissue, organ, cell etc.)	Mostly blood biomarkers and omics
Challenges/Benefits/Limitations, Input Data standardisation	Harmonisation steps to align 20 cohorts with over 300,000 participants is a challenge; lack of governance structure for data sharing through an federated data analysis approach
Existing standards (formats, guidelines, ontologies)	Not really available, we had to develop
Model validation	no information available
Lessons and Comments	harmonized data crucial for cross cohort collaboration; great opportunity to capitalize on existing data; GDPR complicates international collaboration

Project	**Multiple MS: Multiple manifestations of genetic and non-genetic factors in Multiple Sclerosis disentangled with a multi-omics approach to accelerate personalised medicine (https://www.multiplems.eu/)**
Methods	Unsupervised approach: e.g. Topological mapping, multi-partite "knowledge graph". Supervised approaches: e.g. cell-specific pathway analysis coupled to burden score.
Model function	Stratification of patients with MS
Input data	Genetic, lifestyle (questionnaires), established biomarkers. We will have two complementary approaches one using risk factors for MS and one using risk factors for disease severity measured in several different ways. For smaller part of cohort expression and methylome data will also be used.
Output	Clusters of patients that we then will be characterized clinically (i.e. Do the different clusters differ with regard to response to treatment or severity of disease?). Clinical data is already collected.
Processing steps	Data has been collected from several previous studies. Harmonization of data (both genetic, questionnaire and clinical).
Use cases	no information available
Scale (tissue, organ, cell etc.)	Genotyping done on blood. Biomarker analysis on blood or CSF. MRI of brain and spinal cord.
Challenges/Benefits/Limitations, Input Data standardisation	Harmonization of data from >30 previous studies have been challenging. Currently genetic, clinical, biomarker and MRI data have been harmonized. Lifestyle exposures have not been harmonized yet.
Existing standards (formats, guidelines, ontologies)	ICD10 used for comorbidities. Standard formats used for genotype data. Biomarker data as much as possible standardized to units used in clinical medicine. MRI, standardized pipeline used for processing DICOM images, this standard was developed in this project, but applied in several other project too.
Model validation	Models that are developed using data in the retrospective arm of the project will be validated in the prospective observational trial of newly diagnosed MS patients which is also part of the study.
Lessons and Comments	We have learnt that harmonization of data takes longer than we expected. We have also learnt that getting data processing agreements in place allowing sharing of data was complicated and took a lot of effort but is not impossible. We are only now starting the modelling part of the project.

Project	**Personalised treatment of anaemia in lung cancer patients (https://www.dkfz.de/en/systembiologie/AreasofInterest.html)**
Methods	no information available
Model function	Mathematical model (ODE) predicting outcome of treatment options based on hemoglobin and CRP values (longitudinal measurements and cohort data).
Input data	Lab values (Hgb, CRP)
Output	Treatment outcome predictions
Processing steps	no information available
Use cases	no information available
Scale (tissue, organ, cell etc.)	Modelling on the Epo pathway and coupling to whole body effects in anemia
Challenges/Benefits/Limitations, Input Data standardisation	Non-standardised input data, e.g. homemade medication name input data. Challenges: input data variation where input data was not recognized by the model as the same treatment under two different names. Challenges: importance of knowing precise death time and date including which day of the week; perceived conflict with principles of data minimization. Challenge: does the model take into account the subjective well-being of the patient, as an outcome?
Existing standards (formats, guidelines, ontologies)	no information available
Model validation	Model validation: Comparison of patient outcome with or without use of the model, e.g. survival. Design? RCT parallel populations? Difficult to test the model in different hospital because of homemade input data standards, but not impossible.
Lessons and Comments	no information available

Project	**SYSCID: Systems medicine for chronic inflammatory diseases (https://syscid.eu/)**
Methods	Machine learning. Analytical pipelines including but not limited to standard methods DNA: BWA/Samtools/GATK RNA: STAR/DeSEQ2 DNA methylome: Bismark/RnBeads Microbiome: Mothur/Qiime/Humman2/Metaphlan Marker identification and reduction to diagnostic sets are done by different ML approaches (random forest , Bayesian inference)
Model function	Biomarker identification; predict disease outcome and treatment response to guide therapy decisions on individual basis.
Input data	Tissue and blood coupled to clinical data from EHR and PRO. Exomes/genomes, transcriptomes, DNA methylomes and 16rRNA /metagenomics (microbiome) data.
Output	Several data and metadata formats one all analysed level according to standards (DNA/DNAm/RNA) according to IHEC guidelines.
Processing steps	Oriented towards research question, the above mentioned pipelines use standard data processing (e.g. Deseq2).
Use cases	Inflammatory diseases (IBD, RA and SLE)
Scale (tissue, organ, cell etc.)	Tissue, blood and single cell from peripheral leukocytes
Challenges/Benefits/Limitations, Input Data standardisation	Versioning of community standards (e.g. reference genomes and updates of mappers and count software)
Existing standards (formats, guidelines, ontologies)	Single cell field developing quickly, less standardized compared to genetic analyses. As IHEC and HMP (Raes) Partners SYSCID is well aware of data standards. For response analyses, longitudinal analyses and models building on regulatory /functional networks are necessary. Here, we feel that there is much less standardization. This is an unmet need in the systems immunology field.
Model validation	no information available
Lessons and Comments	Importance of standardizing outcome data that is meaningful for patients. Increase communication between modellers and medical specialization communities, patient organisations.

References

Akanksha Limaye, Dr. Anuraj Nayarisseri. 2019. "Machine learning models to predict the precise progression of Tay-Sachs and Related Disease." In *MOL2NET 2019, International Conference on Multidisciplinary Sciences, 5th edition session USEDAT-07: USA-Europe Data Analysis Training School, UPV/EHU*. Bilbao-JSU, Jackson, USA, 2019

Alber, Mark, Adrian Buganza Tepole, William R. Cannon, Suvranu De, Salvador Dura-Bernal, Krishna Garikipati, George Karniadakis, William W. Lytton, Paris Perdikaris, Linda Petzold, and Ellen Kuhl. 2019. 'Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences', *npj Digital Medicine*, 2: 115.

Alioto, T. S., I. Buchhalter, S. Derdak, B. Hutter, M. D. Eldridge, E. Hovig, L. E. Heisler, T. A. Beck, J. T. Simpson, L. Tonon, A. S. Sertier, A. M. Patch, N. Jager, P. Ginsbach, R. Drews, N. Paramasivam, R. Kabbe, S. Chotewutmontri, N. Diessl, C. Previti, S. Schmidt, B. Brors, L. Feuerbach, M. Heinold, S. Grobner, A. Korshunov, P. S. Tarpey, A. P. Butler, J. Hinton, D. Jones, A. Menzies, K. Raine, R. Shepherd, L. Stebbings, J. W. Teague, P. Ribeca, F. C. Giner, S. Beltran, E. Raineri, M. Dabad, S. C. Heath, M. Gut, R. E. Denroche, N. J. Harding, T. N. Yamaguchi, A. Fujimoto, H. Nakagawa, V. Quesada, R. Valdes-Mas, S. Nakken, D. Vodak, L. Bower, A. G. Lynch, C. L. Anderson, N. Waddell, J. V. Pearson, S. M. Grimmond, M. Peto, P. Spellman, M. He, C. Kandoth, S. Lee, J. Zhang, L. Letourneau, S. Ma, S. Seth, D. Torrents, L. Xi, D. A. Wheeler, C. Lopez-Otin, E. Campo, P. J. Campbell, P. C. Boutros, X. S. Puente, D. S. Gerhard, S. M. Pfister, J. D. McPherson, T. J. Hudson, M. Schlesner, P. Lichter, R. Eils, D. T. Jones, and I. G. Gut. 2015. 'A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing', *Nat Commun*, 6: 10001.

Apweiler, R., T. Beissbarth, M. R. Berthold, N. Bluthgen, Y. Burmeister, O. Dammann, A. Deutsch, F. Feuerhake, A. Franke, J. Hasenauer, S. Hoffmann, T. Hofer, P. L. Jansen, L. Kaderali, U. Klingmuller, I. Koch, O. Kohlbacher, L. Kuepfer, F. Lammert, D. Maier, N. Pfeifer, N. Radde, M. Rehm, I. Roeder, J. Saez-Rodriguez, U. Sax, B. Schmeck, A. Schuppert, B. Seilheimer, F. J. Theis, J. Vera, and O. Wolkenhauer. 2018. 'Whither systems medicine?', *Exp Mol Med*, 50: e453.

Auffray, C., R. Balling, I. Barroso, L. Bencze, M. Benson, J. Bergeron, E. Bernal-Delgado, N. Blomberg, C. Bock, A. Conesa, S. Del Signore, C. Delogne, P. Devilee, A. Di Meglio, M. Eijkemans, P. Flicek, N. Graf, V. Grimm, H. J. Guchelaar, Y. K. Guo, I. G. Gut, A. Hanbury, S. Hanif, R. D. Hilgers, A. Honrado, D. R. Hose, J. Houwing-Duistermaat, T. Hubbard, S. H. Janacek, H. Karanikas, T. Kievits, M. Kohler, A. Kremer, J. Lanfear, T. Lengauer, E. Maes, T. Meert, W. Muller, D. Nickel, P. Oledzki, B. Pedersen, M. Petkovic, K. Pliakos, M. Rattray, I. Mas JR, R. Schneider, T. Sengstag, X. Serra-Picamal, W. Spek, L. A. Vaas, O. van Batenburg, M. Vandelaer, P. Varnai, P. Villoslada, J. A. Vizcaino, J. P. Wubbe, and G. Zanetti. 2016. 'Making sense of big data in health research: Towards an EU action plan', *Genome Med*, 8: 71.

Balabanov, Stefan, Thomas Wilhelm, Simone Venz, Gunhild Keller, Christian Scharf, Heike Pospisil, Melanie Braig, Christine Barett, Carsten Bokemeyer, Reinhard Walther, Tim H. Brümmendorf, and Andreas Schuppert. 2013. 'Combination of a Proteomics Approach and Reengineering of Meso Scale Network Models for Prediction of Mode-of-Action for Tyrosine Kinase Inhibitors', *PLoS One*, 8: e53668.

Baldi, Pierre, and Søren Brunak. 2001. *Bioinformatics: the machine learning approach* (MIT Press).

Baldi, Pierre, Søren Brunak, Yves Chauvin, Claus Andersen, and Henrik Nielsen. 2000. 'Assessing the accuracy of prediction algorithms for classification: An overview', *Bioinformatics (Oxford, England)*, 16: 412-24.

Bender, D., and K. Sartipi. 2013. 'HL7 FHIR: An Agile and RESTful Approach to Healthcare Information Exchange', *2013 Ieee 26th International Symposium on Computer-Based Medical Systems (Cbms)*: 326-31.

Bergmann, F. T., R. Adams, S. Moodie, J. Cooper, M. Glont, M. Golebiewski, M. Hucka, C. Laibe, A. K. Miller, D. P. Nickerson, B. G. Olivier, N. Rodriguez, H. M. Sauro, M. Scharm, S. Soiland-Reyes, D.

Waltemath, F. Yvon, and N. Le Novere. 2014. 'COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project', *BMC Bioinformatics*, 15: 369.

Briganti, Giovanni, and Olivier Le Moine. 2020. 'Artificial Intelligence in Medicine: Today and Tomorrow', *Frontiers in medicine*, 7.

Brown, B. C., Consortium Asian Genetic Epidemiology Network Type 2 Diabetes, C. J. Ye, A. L. Price, and N. Zaitlen. 2016. 'Transethnic Genetic-Correlation Estimates from Summary Statistics', *Am J Hum Genet*, 99: 76-88.

Bruggeman, Frank J., and Hans V. Westerhoff. 2007. 'The nature of systems biology', *Trends in Microbiology*, 15: 45-50.

Brunak, S., C. Bjerre Collin, O. Cathaoir K. Eva, M. Golebiewski, M. Kirschner, I. Kockum, H. Moser, and D. Waltemath. 2020. 'Towards standardization guidelines for in silico approaches in personalized medicine', *J Integr Bioinform*, 17.

Choi, Shing Wan, Timothy Shin Heng Mak, and Paul F. O'Reilly. 2018. 'A guide to performing Polygenic Risk Score analyses', *bioRxiv*: 416545.

Collins, G. S., J. B. Reitsma, D. G. Altman, K. G. M. Moons, and TRIPOD Grp. 2015. 'Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement', *European Urology*, 67: 1142-51.

Danhof, M. 2016. 'Systems pharmacology - Towards the modeling of network interactions', *Eur J Pharm Sci*, 94: 4-14.

Debray, T. P., Y. Vergouwe, H. Koffijberg, D. Nieboer, E. W. Steyerberg, and K. G. Moons. 2015. 'A new framework to enhance the interpretation of external validation studies of clinical prediction models', *Journal of Clinical Epidemiology*, 68: 279-89.

Derendorf, H., and B. Meibohm. 1999. 'Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives', *Pharm Res*, 16: 176-85.

DICOM_Secretariat. 2020. 'Digital Imaging and Communications in Medicine '. https://www.dicomstandard.org/.

Dudbridge, F. 2013. 'Power and predictive accuracy of polygenic risk scores', *PLoS Genet*, 9: e1003348.

Duncan, L., H. Shen, B. Gelaye, J. Meijsen, K. Ressler, M. Feldman, R. Peterson, and B. Domingue. 2019. 'Analysis of polygenic risk score usage and performance in diverse human populations', *Nat Commun*, 10: 3328.

Edginton, A. N., and S. Willmann. 2008. 'Physiology-based simulations of a pathological condition: prediction of pharmacokinetics in patients with liver cirrhosis', *Clin Pharmacokinet*, 47: 743-52.

Edwards, L. M., and I. Thiele. 2013. 'Applying systems biology methods to the study of human physiology in extreme environments', *Extrem Physiol Med*, 2: 8.

Fiers, Mwej, L. Minnoye, S. Aibar, C. Bravo Gonzalez-Blas, Z. Kalender Atak, and S. Aerts. 2018. 'Mapping gene regulatory networks from single-cell omics data', *Brief Funct Genomics*, 17: 246-54.

Friberg, L. E., A. Henningsson, H. Maas, L. Nguyen, and M. O. Karlsson. 2002. 'Model of chemotherapy-induced myelosuppression with parameter consistency across drugs', *J Clin Oncol*, 20: 4713-21.

Frohlich, F., T. Kessler, D. Weindl, A. Shadrin, L. Schmiester, H. Hache, A. Muradyan, M. Schutte, J. H. Lim, M. Heinig, F. J. Theis, H. Lehrach, C. Wierling, B. Lange, and J. Hasenauer. 2018. 'Efficient Parameter Estimation Enables the Prediction of Drug Response Using a Mechanistic Pan-Cancer Pathway Model', *Cell Syst*, 7: 567-79 e6.

GA4GH_Community. 2020. 'The Global Alliance for Genomics and Health'. https://www.ga4gh.org/.

Georgiadi, E. C., D. D. Dionysiou, N. Graf, and G. S. Stamatakos. 2012. 'Towards in silico oncology: Adapting a four dimensional nephroblastoma treatment model to a clinical trial case based on multi-method sensitivity analysis', *Computers in Biology and Medicine*, 42: 1064-78.

Gerlowski, Leonard E., and Rakesh K. Jain. 1983. 'Physiologically Based Pharmacokinetic Modeling: Principles and Applications', *Journal of Pharmaceutical Sciences*, 72: 1103-27.

GitHub_Community. 2020. 'GitHub'. https://github.com/features.

Goddard, N. H., M. Hucka, F. Howell, H. Cornelis, K. Shankar, and D. Beeman. 2001. 'Towards NeuroML: Model description methods for collaborative modelling in neuroscience', *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 356: 1209-28.

Golebiewski, Martin. 2019. 'Data Formats for Systems Biology and Quantitative Modeling.' in, *Encyclopedia of Bioinformatics and Computational Biology*.

Grinde, K. E., Q. Qi, T. A. Thornton, S. Liu, A. H. Shadyab, K. H. K. Chan, A. P. Reiner, and T. Sofer. 2019. 'Generalizing polygenic risk scores from Europeans to Hispanics/Latinos', *Genet Epidemiol*, 43: 50-62.

Griss, J., A. R. Jones, T. Sachsenberg, M. Walzer, L. Gatto, J. Hartler, G. G. Thallinger, R. M. Salek, C. Steinbeck, N. Neuhauser, J. Cox, S. Neumann, J. Fan, F. Reisinger, Q. W. Xu, N. Del Toro, Y. Perez-Riverol, F. Ghali, N. Bandeira, I. Xenarios, O. Kohlbacher, J. A. Vizcaino, and H. Hermjakob. 2014. 'The mzTab data exchange format: communicating mass-spectrometry-based proteomics and metabolomics experimental results to a wider audience', *Mol Cell Proteomics*, 13: 2765-75.

Hallow, K. M., and Y. Gebremichael. 2017. 'A quantitative systems physiology model of renal function and blood pressure regulation: Model description', *CPT Pharmacometrics Syst Pharmacol*, 6: 383-92.

Hallow, K. M., P. J. Greasley, G. Helmlinger, L. Chu, H. J. Heerspink, and D. W. Boulton. 2018. 'Evaluation of renal and cardiovascular protection mechanisms of SGLT2 inhibitors: model-based analysis of clinical data', *Am J Physiol Renal Physiol*, 315: F1295-F306.

Halu, A., M. De Domenico, A. Arenas, and A. Sharma. 2019. 'The multiplex network of human diseases', *NPJ Syst Biol Appl*, 5: 15.

Hastings, J. F., A. Gonzalez Rajal, S. L. Latham, J. Z. Han, R. A. McCloy, Y. E. O'Donnell, M. Phimmachanh, A. D. Murphy, A. Nagrial, D. Daneshvar, V. Chin, D. N. Watkins, A. Burgess, and D. R. Croucher. 2020. 'Analysis of pulsed cisplatin signalling dynamics identifies effectors of resistance in lung adenocarcinoma', *Elife*, 9.

Heus, P., J. A. A. G. Damen, R. Pajouheshnia, R. J. P. M. Scholten, J. B. Reitsma, G. S. Collins, D. G. Altman, K. G. M. Moons, and L. Hooft. 2018. 'Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement', *Bmc Medicine*, 16.

Horizon-2020-Advisory-Group. 2018-2020. "Advice for 2018–2020 of the Horizon 2020 Advisory Group for Societal Challenge 1, "Health, Demographic Change and Well-being"." In, edited by "Horizon 2020 Advisory Group".

Hsi-Yang Fritz, Markus, Rasko Leinonen, Guy Cochrane, and Ewan Birney. 2011. 'Efficient storage of high throughput DNA sequencing data using reference-based compression', *Genome Research*, 21: 734-40.

Hucka, M., F. T. Bergmann, A. Drager, S. Hoops, S. M. Keating, N. Le Novere, C. J. Myers, B. G. Olivier, S. Sahle, J. C. Schaff, L. P. Smith, D. Waltemath, and D. J. Wilkinson. 2018. 'The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 2 Core', *J Integr Bioinform*, 15: 1-173.

Hucka, M., D. P. Nickerson, G. D. Bader, F. T. Bergmann, J. Cooper, E. Demir, A. Garny, M. Golebiewski, C. J. Myers, F. Schreiber, D. Waltemath, and N. Le Novere. 2015. 'Promoting Coordinated Development of Community-Based Information Standards for Modeling in Biology: The COMBINE Initiative', *Front Bioeng Biotechnol*, 3: 19.

Ingalls, B. P., and H. M. Sauro. 2003. 'Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories', *J Theor Biol*, 222: 23-36.

International Consortium for Blood Pressure Genome-Wide Association, Studies, G. B. Ehret, P. B. Munroe, K. M. Rice, M. Bochud, A. D. Johnson, D. I. Chasman, A. V. Smith, M. D. Tobin, G. C. Verwoert, S. J. Hwang, V. Pihur, P. Vollenweider, P. F. O'Reilly, N. Amin, J. L. Bragg-Gresham, A. Teumer, N. L. Glazer, L. Launer, J. H. Zhao, Y. Aulchenko, S. Heath, S. Sober, A. Parsa, J. Luan, P. Arora, A. Dehghan, F. Zhang, G. Lucas, A. A. Hicks, A. U. Jackson, J. F. Peden, T. Tanaka, S. H. Wild, I. Rudan, W. Igl, Y. Milaneschi, A. N. Parker, C. Fava, J. C. Chambers, E. R. Fox, M. Kumari, M. J. Go, P. van der Harst, W. H. Kao, M. Sjogren, D. G. Vinay, M. Alexander, Y. Tabara, S. Shaw-Hawkins, P. H. Whincup, Y. Liu, G. Shi, J. Kuusisto, B. Tayo, M. Seielstad, X. Sim, K. D. Nguyen, T. Lehtimaki, G. Matullo, Y. Wu, T. R. Gaunt, N. C. Onland-Moret, M. N. Cooper, C. G. Platou, E. Org, R. Hardy, S. Dahgam, J. Palmen, V. Vitart, P. S. Braund, T. Kuznetsova, C. S. Uiterwaal, A. Adeyemo, W. Palmas, H. Campbell, B. Ludwig, M. Tomaszewski, I. Tzoulaki, N. D. Palmer, C.

ARDIoGRAM consortium, C. KDGen Consortium, Consortium KidneyGen, consortium EchoGen, Charge-Hf consortium, T. Aspelund, M. Garcia, Y. P. Chang, J. R. O'Connell, N. I. Steinle, D. E. Grobbee, D. E. Arking, S. L. Kardia, A. C. Morrison, D. Hernandez, S. Najjar, W. L. McArdle, D. Hadley, M. J. Brown, J. M. Connell, A. D. Hingorani, I. N. Day, D. A. Lawlor, J. P. Beilby, R. W. Lawrence, R. Clarke, J. C. Hopewell, H. Ongen, A. W. Dreisbach, Y. Li, J. H. Young, J. C. Bis, M. Kahonen, J. Viikari, L. S. Adair, N. R. Lee, M. H. Chen, M. Olden, C. Pattaro, J. A. Bolton, A. Kottgen, S. Bergmann, V. Mooser, N. Chaturvedi, T. M. Frayling, M. Islam, T. H. Jafar, J. Erdmann, S. R. Kulkarni, S. R. Bornstein, J. Grassler, L. Groop, B. F. Voight, J. Kettunen, P. Howard, A. Taylor, S. Guarrera, F. Ricceri, V. Emilsson, A. Plump, I. Barroso, K. T. Khaw, A. B. Weder, S. C. Hunt, Y. V. Sun, R. N. Bergman, F. S. Collins, L. L. Bonnycastle, L. J. Scott, H. M. Stringham, L. Peltonen, M. Perola, E. Vartiainen, S. M. Brand, J. A. Staessen, T. J. Wang, P. R. Burton, M. Soler Artigas, Y. Dong, H. Snieder, X. Wang, H. Zhu, K. K. Lohman, M. E. Rudock, S. R. Heckbert, N. L. Smith, K. L. Wiggins, A. Doumatey, D. Shriner, G. Veldre, M. Viigimaa, S. Kinra, D. Prabhakaran, V. Tripathy, C. D. Langefeld, A. Rosengren, D. S. Thelle, A. M. Corsi, A. Singleton, T. Forrester, G. Hilton, C. A. McKenzie, T. Salako, N. Iwai, Y. Kita, T. Ogihara, T. Ohkubo, T. Okamura, H. Ueshima, S. Umemura, S. Eyheramendy, T. Meitinger, H. E. Wichmann, Y. S. Cho, H. L. Kim, J. Y. Lee, J. Scott, J. S. Sehmi, W. Zhang, B. Hedblad, P. Nilsson, G. D. Smith, A. Wong, N. Narisu, A. Stancakova, L. J. Raffel, J. Yao, S. Kathiresan, C. J. O'Donnell, S. M. Schwartz, M. A. Ikram, W. T. Longstreth, Jr., T. H. Mosley, S. Seshadri, N. R. Shrine, L. V. Wain, M. A. Morken, A. J. Swift, J. Laitinen, I. Prokopenko, P. Zitting, J. A. Cooper, S. E. Humphries, J. Danesh, A. Rasheed, A. Goel, A. Hamsten, H. Watkins, S. J. Bakker, W. H. van Gilst, C. S. Janipalli, K. R. Mani, C. S. Yajnik, A. Hofman, F. U. Mattace-Raso, B. A. Oostra, A. Demirkan, A. Isaacs, F. Rivadeneira, E. G. Lakatta, M. Orru, A. Scuteri, M. Ala-Korpela, A. J. Kangas, L. P. Lyytikainen, P. Soininen, T. Tukiainen, P. Wurtz, R. T. Ong, M. Dorr, H. K. Kroemer, U. Volker, H. Volzke, P. Galan, S. Hercberg, M. Lathrop, D. Zelenika, P. Deloukas, M. Mangino, T. D. Spector, G. Zhai, J. F. Meschia, M. A. Nalls, P. Sharma, J. Terzic, M. V. Kumar, M. Denniff, E. Zukowska-Szczechowska, L. E. Wagenknecht, F. G. Fowkes, F. J. Charchar, P. E. Schwarz, C. Hayward, X. Guo, C. Rotimi, M. L. Bots, E. Brand, N. J. Samani, O. Polasek, P. J. Talmud, F. Nyberg, D. Kuh, M. Laan, K. Hveem, L. J. Palmer, Y. T. van der Schouw, J. P. Casas, K. L. Mohlke, P. Vineis, O. Raitakari, S. K. Ganesh, T. Y. Wong, E. S. Tai, R. S. Cooper, M. Laakso, D. C. Rao, T. B. Harris, R. W. Morris, A. F. Dominiczak, M. Kivimaki, M. G. Marmot, T. Miki, D. Saleheen, G. R. Chandak, J. Coresh, G. Navis, V. Salomaa, B. G. Han, X. Zhu, J. S. Kooner, O. Melander, P. M. Ridker, S. Bandinelli, U. B. Gyllensten, A. F. Wright, J. F. Wilson, L. Ferrucci, M. Farrall, J. Tuomilehto, P. P. Pramstaller, R. Elosua, N. Soranzo, E. J. Sijbrands, D. Altshuler, R. J. Loos, A. R. Shuldiner, C. Gieger, P. Meneton, A. G. Uitterlinden, N. J. Wareham, V. Gudnason, J. I. Rotter, R. Rettig, M. Uda, D. P. Strachan, J. C. Witteman, A. L. Hartikainen, J. S. Beckmann, E. Boerwinkle, R. S. Vasan, M. Boehnke, M. G. Larson, M. R. Jarvelin, B. M. Psaty, G. R. Abecasis, A. Chakravarti, P. Elliott, C. M. van Duijn, C. Newton-Cheh, D. Levy, M. J. Caulfield, and T. Johnson. 2011. 'Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk', *Nature*, 478: 103-9.

International Schizophrenia, Consortium, S. M. Purcell, N. R. Wray, J. L. Stone, P. M. Visscher, M. C. O'Donovan, P. F. Sullivan, and P. Sklar. 2009. 'Common polygenic variation contributes to risk of schizophrenia and bipolar disorder', *Nature*, 460: 748-52.

Kalra, D., T. Beale, and S. Heard. 2005. 'The openEHR Foundation', *Stud Health Technol Inform*, 115: 153-73.

Kemp, B., A. Varri, A. C. Rosa, K. D. Nielsen, and J. Gade. 1992. 'A simple format for exchange of digitized polygraphic recordings', *Electroencephalogr Clin Neurophysiol*, 82: 391-3.

Khera, A. V., M. Chaffin, K. G. Aragam, M. E. Haas, C. Roselli, S. H. Choi, P. Natarajan, E. S. Lander, S. A. Lubitz, P. T. Ellinor, and S. Kathiresan. 2018. 'Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations', *Nature Genetics*, 50: 1219-24.

Komorowski, M., L. A. Celi, O. Badawi, A. C. Gordon, and A. A. Faisal. 2018. 'The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care', *Nat Med*, 24: 1716-20.

Kuepfer, L., C. Niederalt, T. Wendl, J. F. Schlender, S. Willmann, J. Lippert, M. Block, T. Eissing, and D. Teutonico. 2016. 'Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model', *CPT Pharmacometrics Syst Pharmacol*, 5: 516-31.

Lappalainen, T., M. Sammeth, M. R. Friedlander, P. A. t Hoen, J. Monlong, M. A. Rivas, M. Gonzalez-Porta, N. Kurbatova, T. Griebel, P. G. Ferreira, M. Barann, T. Wieland, L. Greger, M. van Iterson, J. Almlof, P. Ribeca, I. Pulyakhina, D. Esser, T. Giger, A. Tikhonov, M. Sultan, G. Bertier, D. G. MacArthur, M. Lek, E. Lizano, H. P. Buermans, I. Padioleau, T. Schwarzmayr, O. Karlberg, H. Ongen, H. Kilpinen, S. Beltran, M. Gut, K. Kahlem, V. Amstislavskiy, O. Stegle, M. Pirinen, S. B. Montgomery, P. Donnelly, M. I. McCarthy, P. Flicek, T. M. Strom, Consortium Geuvadis, H. Lehrach, S. Schreiber, R. Sudbrak, A. Carracedo, S. E. Antonarakis, R. Hasler, A. C. Syvanen, G. J. van Ommen, A. Brazma, T. Meitinger, P. Rosenstiel, R. Guigo, I. G. Gut, X. Estivill, and E. T. Dermitzakis. 2013. 'Transcriptome and genome sequencing uncovers functional variation in humans', *Nature*, 501: 506-11.

Le Novere, N., A. Finney, M. Hucka, U. S. Bhalla, F. Campagne, J. Collado-Vides, E. J. Crampin, M. Halstead, E. Klipp, P. Mendes, P. Nielsen, H. Sauro, B. Shapiro, J. L. Snoep, H. D. Spence, and B. L. Wanner. 2005. 'Minimum information requested in the annotation of biochemical models (MIRIAM)', *Nat Biotechnol*, 23: 1509-15.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25: 2078-9.

Lipman, DJ, and WR Pearson. 1985. 'Rapid and sensitive protein similarity searches', *Science*, 227: 1435-41.

Lippert, J., M. Brosch, O. von Kampen, M. Meyer, H. U. Siegmund, C. Schafmayer, T. Becker, B. Laffert, L. Gorlitz, S. Schreiber, P. J. Neuvonen, M. Niemi, J. Hampe, and L. Kuepfer. 2012. 'A mechanistic, model-based approach to safety assessment in clinical development', *CPT Pharmacometrics Syst Pharmacol*, 1: e13.

Lloyd, C. M., M. D. Halstead, and P. F. Nielsen. 2004. 'CellML: its future, present and past', *Prog Biophys Mol Biol*, 85: 433-50.

Madsen, C., A. G. Moreno, P. Umesh, Z. Palchick, N. Roehner, C. Atallah, B. Bartley, K. Choi, R. S. Cox, T. Corochowski, R. K. Grunberg, C. Macklin, J. McLaughlin, X. W. Meng, T. Nguyen, M. Pocock, M. Samineni, J. Scott-Brown, Y. Tarter, M. Zhang, Z. Zhang, Z. Zundel, J. Beal, M. Bissell, K. Clancy, J. H. Gennari, G. Misirli, C. Myers, E. Oberortner, H. Sauro, and A. Wipat. 2019. 'Synthetic Biology Open Language (SBOL) Version 2.3', *Journal of Integrative Bioinformatics*, 16.

Maemondo, M., A. Inoue, K. Kobayashi, S. Sugawara, S. Oizumi, H. Isobe, A. Gemma, M. Harada, H. Yoshizawa, I. Kinoshita, Y. Fujita, S. Okinaga, H. Hirano, K. Yoshimori, T. Harada, T. Ogura, M. Ando, H. Miyazawa, T. Tanaka, Y. Saijo, K. Hagiwara, S. Morita, T. Nukiwa, and Group North-East Japan Study. 2010. 'Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR', *N Engl J Med*, 362: 2380-8.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. Mackay, S. A. McCarroll, and P. M. Visscher. 2009. 'Finding the missing heritability of complex diseases', *Nature*, 461: 747-53.

Marquez-Luna, C., P. R. Loh, Consortium South Asian Type 2 Diabetes, Sigma Type 2 Diabetes Consortium, and A. L. Price. 2017. 'Multiethnic polygenic risk scores improve risk prediction in diverse populations', *Genet Epidemiol*, 41: 811-23.

Martens, L., M. Chambers, M. Sturm, D. Kessner, F. Levander, J. Shofstahl, W. H. Tang, A. Rompp, S. Neumann, A. D. Pizarro, L. Montecchi-Palazzi, N. Tasman, M. Coleman, F. Reisinger, P. Souda, H. Hermjakob, P. A. Binz, and E. W. Deutsch. 2011. 'mzML--a community standard for mass spectrometry data', *Mol Cell Proteomics*, 10: R110 000133.

Mok, T. S., Y. L. Wu, S. Thongprasert, C. H. Yang, D. T. Chu, N. Saijo, P. Sunpaweravong, B. Han, B. Margono, Y. Ichinose, Y. Nishiwaki, Y. Ohe, J. J. Yang, B. Chewaskulyong, H. Jiang, E. L. Duffield,

C. L. Watkins, A. A. Armour, and M. Fukuoka. 2009. 'Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma', *N Engl J Med*, 361: 947-57.

Morrison, Tina M., Pras Pathmanathan, Mariam Adwan, and Edward Margerrison. 2018. 'Advancing Regulatory Science With Computational Modeling for Medical Devices at the FDA's Office of Science and Engineering Laboratories', *Frontiers in medicine*, 5: 1-11.

Moseley, P. L., and S. Brunak. 2019. 'Identifying Sepsis Phenotypes', *JAMA*, 322: 1416-17.

Muller, F. J., and A. Schuppert. 2011. 'Few inputs can reprogram biological networks', *Nature*, 478: E4; discussion E4-5.

Murad, M. H., N. Asi, M. Alsawas, and F. Alahdab. 2016. 'New evidence pyramid', *Evid Based Med*, 21: 125-7.

Musuamba, F. T., R. Bursi, E. Manolis, K. Karlsson, A. Kulesza, E. Courcelles, J. P. Boissel, R. Lesage, C. Crozatier, E. M. Voisin, C. F. Rousseau, T. Marchal, R. Alessandrello, and L. Geris. 2020. 'Verifying and Validating Quantitative Systems Pharmacology and In Silico Models in Drug Development: Current Needs, Gaps, and Challenges', *CPT Pharmacometrics Syst Pharmacol*, 9: 195-97.

Myers, C. J., G. Bader, P. Gleeson, M. Golebiewski, M. Hucka, N. Le Novere, D. P. Nickerson, F. Schreiber, and D. Waltemath. 2017. 'A Brief History of Combine', *2017 Winter Simulation Conference (Wsc)*: 884-95.

Neal, M. L., M. Konig, D. Nickerson, G. Misirli, R. Kalbasi, A. Drager, K. Atalag, V. Chelliah, M. T. Cooling, D. L. Cook, S. Crook, M. de Alba, S. H. Friedman, A. Garny, J. H. Gennari, P. Gleeson, M. Golebiewski, M. Hucka, N. Juty, C. Myers, B. G. Olivier, H. M. Sauro, M. Scharm, J. L. Snoep, V. Toure, A. Wipat, O. Wolkenhauer, and D. Waltemath. 2019. 'Harmonizing semantic annotations for computational models in biology', *Brief Bioinform*, 20: 540-50.

Nickerson, D., K. Atalag, B. de Bono, J. Geiger, C. Goble, S. Hollmann, J. Lonien, W. Muller, B. Regierer, N. J. Stanford, M. Golebiewski, and P. Hunter. 2016. 'The Human Physiome: how standards, software and innovative service infrastructures are providing the building blocks to make it achievable', *Interface Focus*, 6: 20150103.

Nielsen, Annelaura B., Hans-Christian Thorsen-Meyer, Kirstine Belling, Anna P. Nielsen, Cecilia E. Thomas, Piotr J. Chmura, Mette Lademann, Pope L. Moseley, Marc Heimann, Lars Dybdahl, Lasse Spangsege, Patrick Hulsen, Anders Perner, and Søren Brunak. 2019. 'Survival prediction in intensive-care units based on aggregation of long-term disease history and acute physiology: a retrospective study of the Danish National Patient Registry and electronic patient records', *The Lancet Digital Health*, 1: e78-e89.

Njolstad, P. R., O. A. Andreassen, S. Brunak, A. D. Borglum, J. Dillner, T. Esko, P. W. Franks, N. Freimer, L. Groop, H. Heimer, D. M. Hougaard, E. Hovig, K. Hveem, A. Jalanko, J. Kaprio, G. P. Knudsen, M. Melbye, A. Metspalu, P. B. Mortensen, J. Palmgren, A. Palotie, W. Reed, H. Stefansson, N. O. Stitziel, P. F. Sullivan, U. Thorsteinsdottir, M. Vaudel, E. Vuorio, T. Werge, C. Stoltenberg, and K. Stefansson. 2019. 'Roadmap for a precision-medicine initiative in the Nordic region', *Nature Genetics*, 51: 924-30.

Novère, Nicolas Le, Andrew Finney, Michael Hucka, Upinder S. Bhalla, Fabien Campagne, Julio Collado-Vides, Edmund J. Crampin, Matt Halstead, Edda Klipp, Pedro Mendes, Poul Nielsen, Herbert Sauro, Bruce Shapiro, Jacky L. Snoep, Hugh D. Spence, and Barry L. Wanner. 2005. 'Minimum information requested in the annotation of biochemical models (MIRIAM)', *Nat Biotechnol*, 23: 1509-15.

Ó Cathaoir K., Gefenas E., Hartlev M., Mourby M., Lukaseviciene V. 2020. "Legal and ethical review of in silico modelling." In. www.eu-stands4pm.eu.

Okser, S., T. Pahikkala, A. Airola, T. Salakoski, S. Ripatti, and T. Aittokallio. 2014. 'Regularized machine learning in the genetic prediction of complex traits', *PLoS Genet*, 10: e1004754.

Pare, G., S. Mao, and W. Q. Deng. 2017. 'A machine-learning heuristic to improve gene score prediction of polygenic traits', *Sci Rep*, 7: 12665.

Pearson, W R, and D J Lipman. 1988. 'Improved tools for biological sequence comparison', *Proceedings of the National Academy of Sciences*, 85: 2444-48.

Pérez-Urizar, José, Vinicio Granados-Soto, Francisco J. Flores-Murrieta, and Gilberto Castañeda-Hernández. 2000. 'Pharmacokinetic-Pharmacodynamic Modeling: Why?', *Archives of Medical Research*, 31: 539-45.

Pritchard, Jonathan K., and Nancy J. Cox. 2002. 'The allelic architecture of human disease genes: common disease–common variant… or not?', *Human Molecular Genetics*, 11: 2417-23.

Raue, A., C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. 2009. 'Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood', *Bioinformatics*, 25: 1923-9.

Rosse, Cornelius, and José L. V. Mejino. 2003. 'A reference ontology for biomedical informatics: the Foundational Model of Anatomy', *Journal of Biomedical Informatics*, 36: 478-500.

Saez-Rodriguez, Julio, and Nils Blüthgen. 2020. 'Personalized signaling models for personalized treatments', *Molecular Systems Biology*, 16: e9042.

Schreiber, F., G. D. Bader, P. Gleeson, M. Golebiewski, M. Hucka, N. Le Novere, C. Myers, D. Nickerson, B. Sommer, and D. Waltemath. 2016. 'Specifications of Standards in Systems and Synthetic Biology: Status and Developments in 2016', *Journal of Integrative Bioinformatics*, 13.

Schreiber, F., B. Sommer, G. D. Bader, P. Gleeson, M. Golebiewski, M. Hucka, S. M. Keating, M. Konig, C. Myers, D. Nickerson, and D. Waltemath. 2019. 'Specifications of Standards in Systems and Synthetic Biology: Status and Developments in 2019', *J Integr Bioinform*, 16: 1-5.

Schultze, J. L., Syscid consortium, and P. Rosenstiel. 2018. 'Systems Medicine in Chronic Inflammatory Diseases', *Immunity*, 48: 608-13.

Shefchek, K. A., N. L. Harris, M. Gargano, N. Matentzoglu, D. Unni, M. Brush, D. Keith, T. Conlin, N. Vasilevsky, X. A. Zhang, J. P. Balhoff, L. Babb, S. M. Bello, H. Blau, Y. Bradford, S. Carbon, L. Carmody, L. E. Chan, V. Cipriani, A. Cuzick, M. D. Rocca, N. Dunn, S. Essaid, P. Fey, C. Grove, J. P. Gourdine, A. Hamosh, M. Harris, I. Helbig, M. Hoatlin, M. Joachimiak, S. Jupp, K. B. Lett, S. E. Lewis, C. McNamara, Z. M. Pendlington, C. Pilgrim, T. Putman, V. Ravanmehr, J. Reese, E. Riggs, S. Robb, P. Roncaglia, J. Seager, E. Segerdell, M. Similuk, A. L. Storm, C. Thaxon, A. Thessen, J. O. B. Jacobsen, J. A. McMurry, T. Groza, S. Kohler, D. Smedley, P. N. Robinson, C. J. Mungall, M. A. Haendel, M. C. Munoz-Torres, and D. Osumi-Sutherland. 2020. 'The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species', *Nucleic Acids Res*, 48: D704-D15.

Simeoni, Monica, Paolo Magni, Cristiano Cammia, Giuseppe De Nicolao, Valter Croci, Enrico Pesenti, Massimiliano Germani, Italo Poggesi, and Maurizio Rocchetti. 2004. 'Predictive Pharmacokinetic-Pharmacodynamic Modeling of Tumor Growth Kinetics in Xenograft Models after Administration of Anticancer Agents', *Cancer Research*, 64: 1094-101.

Stamatakos, G., D. Dionysiou, A. Lunzer, R. Belleman, E. Kolokotroni, E. Georgiadi, M. Erdt, J. Pukacki, S. Rueping, S. Giatili, A. d' Onofrio, S. Sfakianakis, K. Marias, C. Desmedt, M. Tsiknakis, and N. Graf. 2014. 'The Technologically Integrated Oncosimulator: Combining Multiscale Cancer Modeling With Information Technology in the In Silico Oncology Context', *Ieee Journal of Biomedical and Health Informatics*, 18: 840-54.

t Hoen, P. A., M. R. Friedlander, J. Almlof, M. Sammeth, I. Pulyakhina, S. Y. Anvar, J. F. Laros, H. P. Buermans, O. Karlberg, M. Brannvall, Geuvadis Consortium, J. T. den Dunnen, G. J. van Ommen, I. G. Gut, R. Guigo, X. Estivill, A. C. Syvanen, E. T. Dermitzakis, and T. Lappalainen. 2013. 'Reproducibility of high-throughput mRNA and small RNA sequencing across laboratories', *Nat Biotechnol*, 31: 1015-22.

Teutonico, D., F. Musuamba, H. J. Maas, A. Facius, S. Yang, M. Danhof, and O. Della Pasqua. 2015. 'Generating Virtual Patients by Multivariate and Discrete Re-Sampling Techniques', *Pharm Res*, 32: 3228-37.

van der Graaf, P. H., and N. Benson. 2011. 'Systems pharmacology: bridging systems biology and pharmacokinetics-pharmacodynamics (PKPD) in drug discovery and development', *Pharm Res*, 28: 1460-4.

van der Wijst, M. G. P., D. H. de Vries, H. Brugge, H. J. Westra, and L. Franke. 2018. 'An integrative approach for building personalized gene regulatory networks for precision medicine', *Genome Med*, 10: 96.

Vidal, M., M. E. Cusick, and A. L. Barabasi. 2011. 'Interactome networks and human disease', *Cell*, 144: 986-98.

Vilhjalmsson, B. J., J. Yang, H. K. Finucane, A. Gusev, S. Lindstrom, S. Ripke, G. Genovese, P. R. Loh, G. Bhatia, R. Do, T. Hayeck, H. H. Won, Discovery Biology Schizophrenia Working Group of the Psychiatric Genomics Consortium, study Risk of Inherited Variants in Breast Cancer, S. Kathiresan, M. Pato, C. Pato, R. Tamimi, E. Stahl, N. Zaitlen, B. Pasaniuc, G. Belbin, E. E. Kenny, M. H. Schierup, P. De Jager, N. A. Patsopoulos, S. McCarroll, M. Daly, S. Purcell, D. Chasman, B. Neale, M. Goddard, P. M. Visscher, P. Kraft, N. Patterson, and A. L. Price. 2015. 'Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores', *Am J Hum Genet*, 97: 576-92.

Waltemath, D., R. Adams, F. T. Bergmann, M. Hucka, F. Kolpakov, A. K. Miller, Moraru, II, D. Nickerson, S. Sahle, J. L. Snoep, and N. Le Novere. 2011. 'Reproducible computational biology experiments with SED-ML--the Simulation Experiment Description Markup Language', *BMC Syst Biol*, 5: 198.

Westergaard, D., P. Moseley, F. K. H. Sorup, P. Baldi, and S. Brunak. 2019. 'Population-wide analysis of differences in disease progression patterns in men and women', *Nat Commun*, 10: 666.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S. A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons. 2016. 'The FAIR Guiding Principles for scientific data management and stewardship', *Sci Data*, 3: 160018.

Wolkenhauer, O., C. Auffray, O. Brass, J. Clairambault, A. Deutsch, D. Drasdo, F. Gervasio, L. Preziosi, P. Maini, A. Marciniak-Czochra, C. Kossow, L. Kuepfer, K. Rateitschak, I. Ramis-Conde, B. Ribba, A. Schuppert, R. Smallwood, G. Stamatakos, F. Winter, and H. Byrne. 2014. 'Enabling multiscale modeling in systems medicine', *Genome Med*, 6.

Wolkenhauer, O., C. Auffray, R. Jaster, G. Steinhoff, and O. Dammann. 2013. 'The road from systems biology to systems medicine', *Pediatr Res*, 73: 502-7.

Wray, N. R., M. E. Goddard, and P. M. Visscher. 2007. 'Prediction of individual genetic risk to disease from genome-wide association studies', *Genome Res*, 17: 1520-8.

Wray, N. R., J. Yang, B. J. Hayes, A. L. Price, M. E. Goddard, and P. M. Visscher. 2013. 'Pitfalls of predicting complex traits from SNPs', *Nat Rev Genet*, 14: 507-15.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. 'Common SNPs explain a large proportion of the heritability for human height', *Nature Genetics*, 42: 565-9.

Zhou, X., J. Menche, A. L. Barabasi, and A. Sharma. 2014. 'Human symptoms-disease network', *Nat Commun*, 5: 4212.

Acknowledgements

EU-STANDS4PM is funded by the European Union Horizon2020 framework programme of the European Commission Directorate-General for Research and Innovation under Grant Agreement # 825843.

The current document is part of work package 2 of EU-STANDS4PM "Integrative data analysis and *in silico* models in personalised medicine". As such it combines deliverables D2.1 "EU-wide mapping report on big data harmonization and integration methodologies for *in silico* modelling", D2.2 "EU-wide mapping report on methodologies of risk prediction and progression patterns of common diseases" and D2.3 "Recommendations for standards that enable the development of predictive *in silico* models for the interpretation of health data"


EU-STANDS4PM
standards for in silico models
for personalised medicine