Building an international and interdisciplinary community to develop immune digital twins for complex human pathologies

Anna Niarakis[1,2], Reinhard Laubenbacher[3], Gary An[4], Yaron Ilan[5], Jasmin Fisher[6], Åsmund Flobak[7,8,9], Kristin Reiche[10,11,12], María Rodríguez Martínez[13], Liesbet Geris[14,15,16], Luiz Ladeira[16], Lorenzo Veschini[17], Michael L. Blinov[18], Francesco Messina[19], Luis L. Fonseca[3], Sandra Ferreira[20], Arnau Montagud[21], Vincent Noël[22,23,24], Malvina Marku[25], Eirini Tsirvouli[26,27], Marcella Torres[28], Leonard A. Harris[29,30,31], T.J. Sego[3], Chase Cockrell[4], Amanda Shick[32], Hasan Balci[33], Albin Salazar[34], Kinza Rian[35], Ahmed Abdelmonem Hemedan[36], Marina Esteban-Medina[35], Bernard Staumont[16], Esteban Hernandez-Vargas[37], Shiny Martis B.[38], Alejandro Madrid-Valiente[21], Panagiotis Karampelesis[39], Luis Sordo Vieira[3], Pradyumna Harlapur[40], Alexander Kulesza[38], Niloofar Nikaein[41,42], Winston Garira[43,44,45], Rahuman S. Malik Sheriff[46], Juilee Thakar[47], Van Du T. Tran[48], Jose Carbonell-Caballero[21], Soroush Safaei[49,50], Alfonso Valencia[21, 51], Andrei Zinovyev[52], James A. Glazier[53]

Corresponding author: Anna Niarakis; anna.niarakis@inria.fr

Affiliations

1. Molecular, Cellular and Developmental Biology Unit (MCD), Centre de Biologie Integrative (CBI), University of Toulouse, UPS, CNRS, Toulouse, France
2. Lifeware Group, Inria, Saclay-île de France, Palaiseau, France
3. Department of Medicine, University of Florida, Gainesville, Florida, USA
4. Department of Surgery, University of Vermont Larner College of Medicine
5. Faculty of Medicine Hebrew University, Hadassah Medical Center, Jerusalem, Israel
6. UCL Cancer Institute, University College London, Paul O'Gorman Building, 72 Huntley Street, London, WC1E 6BT, UK
7. Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway
8. The Cancer Clinic, St Olav's University Hospital, Trondheim, Norway
9. Department of Biotechnology and Nanomedicine, SINTEF Industry, Trondheim, Norway
10. Department of Diagnostics, Fraunhofer Institute for Cell Therapy and Immunology, Leipzig, Germany
11. Institute of Clinical Immunology, Medical Faculty, University Hospital, University of Leipzig, Leipzig, Germany
12. Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany
13. IBM Zürich Research Laboratory, Zürich, Switzerland
14. Prometheus Division of Skeletal Tissue Engineering, KU Leuven, Leuven, Belgium
15. Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven, Leuven
16. Biomechanics Research Unit, GIGA In Silico Medicine, GIGA institute, University of Liège, Liège, Belgium
17. Faculty of Dentistry Oral & Craniofacial Sciences, King's College London
18. Center for Cell Analysis and Modeling, UConn Health, CT 06030 USA
19. Department of Epidemiology, Preclinical Research and Advanced Diagnostic, National Institutefor Infectious Diseases 'Lazzaro Spallanzani' - I.R.C.C.S., Rome, Italy
20. Mathematics Department and Center of Mathematics, University of Beira Interior, Portugal

21. Barcelona Supercomputing Center (BSC)
22. Institut Curie, Université PSL, F-75005, Paris, France
23. INSERM, U900, F-75005, Paris, France
24. Mines ParisTech, Université PSL, F-75005, Paris, France
25. Université de Toulouse, Inserm, CNRS, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France
26. Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
27. Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway
28. Department of Mathematics and Statistics, University of Richmond, Richmond, Virginia, United States
29. Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR, USA
30. Interdisciplinary Graduate Program in Cell and Molecular Biology, University of Arkansas, Fayetteville, AR, USA
31. Cancer Biology Program, Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR
32. Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida, USA
33. Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, The Netherlands
34. INRIA Paris/CNRS/École Normale Supérieure/PSL Research University.
35. Andalusian Platform for Computational Medicine, Andalusian Public
Foundation Progress and Health-FPS, Seville, Spain
36. Bioinformatics Core Unit, Luxembourg Centre of Systems Biomedicine LCSB, Luxembourg University
37. Department of Mathematical and Statistical Sciences, University of Idaho, Moscow, ID, USA
38. Novadiscovery, Lyon, France
39. Department of Electrical and Computer Engineering, University of Patras, Patras, Greece
40. Department of Bioengineering, Indian Insititue of Science, Bengaluru, India
41. School of Medical Sciences, Faculty of Medicine and Health, Örebro University, SE-70182 Örebro, Sweden
42. X-HiDE - Exploring Inflammation in Health and Disease Consortium, Örebro University, Sweden
43. Multiscale Mathematical Modelling of Living Systems program (M3-LSP)
44. Department of Mathematical Sciences, Sol Plaatje University
45. Private Bag X5008, Kimberley 8300, South Africa.
46. EMBL-EBI
47. Department of Microbiology & Immunology and Department of Biostatistics &Computational Biology, University of Rochester Medical Center, NY 14642, USA
48. Vital-IT Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland
49. Institute of Biomedical Engineering and Technology, Ghent University, Belgium
50. Auckland Bioengineering Institute, University of Auckland, New Zealand
51. ICREA, 23 Passeig Lluís Companys, 08010 Barcelona, Spain
52. In silico R&D, Evotec, 31400 Toulouse, France
53. Biocomplexity Institute and Department of Intelligent Systems Engineering, Indiana University, Bloomington, Indiana, 47408, USA

Abstract
Digital twins in medicine are computational models that represent the health state of individual patients over time, enabling optimal therapeutics and forecasting patient prognosis, representing a key technology for personalised care. Many health conditions involve the immune system as an essential component, and hence, it is crucial to include its key features across spatial and temporal scales in medical digital twins. The immune response is complex and heterogeneous across diseases and patients, and its modelling requires the collective expertise of the international clinical, immunology, and computational modelling communities. A 2023 three-week workshop on immune digital twins brought together almost 100 researchers from these communities into a consortium to promote interdisciplinary collaboration and develop a detailed roadmap for immune digital twin modelling and application to be pursued over the next two years. This paper outlines the initial progress on immune digital twins achieved during the workshop and the environment that enabled effective communication between these three communities. Future steps include developing a repository of existing computational models related to the human immune system and developing infrastructure to construct complex disease models, including immune system components.

Keywords: immune system, immune digital twin, computational model, community effort, interdisciplinary collaboration.

Building a sustainable interdisciplinary community of researchers focused on Immune Digital Twin (IDT) technology

A digital twin (DT) in biomedicine is a virtual representation of a patient, or a patient's state, that allows communication and data feedback from the actual patient to the virtual patient and vice versa. A recent report by the National Academies of Sciences, Engineering and Medicine in the US specified that in healthcare, this feedback loop might not be through (semi-)automated interactions but might require a human-in-the-middle (1). This interpretation aligns with the definition taken by the European Commission in the development of their Virtual Human Twin (VHT) initiative and the recommendation in the VHT roadmap (2,3) . While DT approaches in medicine are still in their infancy, a few biomedical applications close to the DT concept have already been implemented in oncology, radiology, and cardiology (4–18). DTs of large blood vessels could allow the early diagnosis of potential abnormalities and aid in designing interventions (19–21). Pancreatic DTs, representing an "artificial pancreas", can largely automate the decision algorithms for the administration of insulin, leading to control and reduction of long-term consequences of type I diabetes. The clinical success achieved with the artificial pancreas proves that the DT paradigm can profoundly change medical care and improve human health (22–25).

Several factors have limited the development and adoption of mechanistic simulations of the immune system to improve patient care directly. Importantly, we still need a complete understanding of the immune system's functions in health, disease and response to therapies. To progress, we must comprehensively leverage what we know and benefit from a wealth of data, tools, and algorithms to augment mechanism-based simulations (26). Such a complex endeavour can only be achieved through a coordinated, combined effort of clinicians, immunologists, experimental and computational biologists, computer scientists, bioinformaticians, and mathematical modellers (Figure 1).

A workshop on Building Immune Digital Twins (IDTs) was held at the Institut Pascal, University of Paris Saclay, on the outskirts of Paris, France, bringing together almost 100 scientists from 19 countries[1]. Over the course of three weeks, the workshop included six keynotes, twelve advanced talks, nine advanced tutorials on simulation software and platforms, and nine interventions from the industry, pharma, biotech, start up and bio-cluster sectors[2]. The participants had different backgrounds, including biology, medicine, immunology, computational biology, molecular and cellular biology, biotechnology, engineering, mathematics, computer science, biochemistry, physics, and different levels of seniority, spanning from masters level students to chairs and department directors. The participants came from academia and the private sector, including pharmaceutical and biotech companies and start-ups. The group worked together to eliminate communication barriers, create synergies, and lay the foundation for an active community working to bring a prototype of an IDT technology to life, using as a guide the roadmap on Building Digital Twins for the human immune system (27). Genopole[3], a prominent French bio-cluster, also supported the workshop which helped create links with all stakeholders needed for

[1] https://www.institut-pascal.universite-paris-saclay.fr/en/scientific-programs/building-immune-digital-twins
[2] https://indico.ijclab.in2p3.fr/event/9017/page/377-orateursspeakers
[3] https://www.genopole.com/

progress. As a first success, the Working Group was selected as a new Research Data Alliance working group[4] and will receive technical support to reach its first milestones in the next two years. The purpose of this article is to describe the workshop's outcome and the group's activities that laid the foundation for creating the international and interdisciplinary community of Immune Digital Twins.

The workshop was structured around lectures and keynotes on the state-of-the-art, round tables and panel discussions in the mornings, followed by extensive breakout sessions and group activities during the afternoons. In the following sections, we will present the highlights of the working documents produced during the three-week workshop. Information about participants and access to the presentation slides can be found on the workshop's website[5].

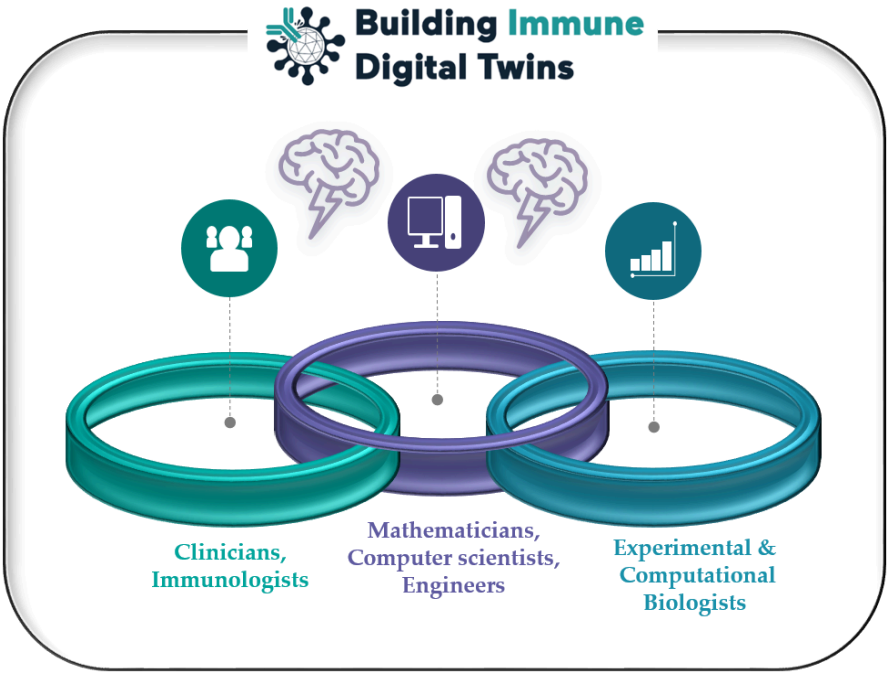

Figure 1. Bringing different stakeholders together to create an international and interdisciplinary community committed to developing and deploying Immune Digital Twins.

Basic Principles for Designing IDTs

An Immune Digital Twin is a digital twin for a particular medical application with a significant immune system component, e.g., a digital twin to treat pneumonia patients. Following industrial design practice, a DT should be a system that enables a two-way flow of information, designed to receive information from the patient (the physical twin), process it, and recalibrate the digital twin to improve the accuracy of the simulated dynamics, used to, in turn, forecast the prognosis and optimise the treatment of the physical twin. The flow of information may or may not proceed in real-time and should provide the necessary data to contextualise, recalibrate, and personalise the DT. The driving hypothesis is that a frequent feed of biomedical data related to different aspects of disease manifestations, combined with a robust computational environment, could give a significant advantage compared to

[4] https://www.rd-alliance.org/groups/building-immune-digital-twins-wg
[5] https://indico.ijclab.in2p3.fr/event/9017/

simulations of computational models not linked to the patient in a two-way relationship. This capability holds the potential to improve personalised care and patient-tailored treatments. However, implementing such technology may only be feasible for some pathologies.

In either case, data accessibility and integration for the IDT feed should be seamless and enabled in a protected and anonymised fashion to ensure patient privacy. This would require a federated database with harmonised and standardised multimodal data (clinical, omic, imaging, lifestyle, etc). In such a setup, statistical and machine learning (ML) analyses could be performed remotely (as in the successful cases of RHAPSODY for diabetes (28) or SOPHIA for obesity research (29) and streamed into the IDT without sensitive patient-level data disclosure. Ideally, the setup should allow updates whenever a new round of patient measurements is available. The personalised IDT could be considered part of the patient's health record.

Causal relationships and parameters within multiscale systems are usually inferred independently for each scale, often relying on experimental data from separate studies. To effectively implement future multiscale IDT, it is crucial to promote adoption of best practices in formulating multiscale experimental designs. These designs should involve interventions at multiple scales simultaneously, enabling the reliable establishment of causal relationships between them. However, implementing these interventions faces challenges within the personalised medicine approach of IDTs. It is not always feasible to acquire *in vivo* samples from actual patients. Humanised *in vitro* systems such as cell-based assays or organoids could be employed to produce relevant data and help parametrize the IDTs. However, these systems often lack a representative microenvironment for drawing meaningful conclusions.

Besides specific characteristics for the "internal design and content" of the IDT, there are also important general features that a successful IDT system should possess to comply with best practices and community guidelines regarding large-scale and multi-scale models. These features include the compliance with the FAIR Principles. The IDT should follow the FAIR principles (30–32) and be:

I. ***Findable***: The different IDT elements should be fully annotated and characterised by globally unique and persistent identifiers and stored in appropriate data and model repositories, facilitating their retrieval. Their metadata should also be indexed in a searchable resource.
II. ***Accessible***: The different IDT elements and their metadata should be retrievable in an open and accessible manner following a standard communication protocol. The metadata should be available even if the IDT elements are inaccessible.
III. ***Interoperable:*** IDTs should be interoperable and able to work together with other models and systems, regardless of the underlying infrastructure. Interoperability will allow the IDT to process information from multiple heterogeneous sources, ensuring a seamless flow of information. Standard input and output formats and the use of accessible programming languages and environments will help towards its adoption and usability. The IDT design requires a concerted effort by the systems biology community to adopt and implement suggested community standards, such as Systems Biology Markup

Language (SBML) (33) for mathematical model exchange, Systems Biology Graphical Notation (SBGN) (34) for model visualisation, Biological Pathway Exchange (BioPaX) (35) for pathway descriptions, and Simulation Experiment Description Markup Language (SED-ML) (36) for simulation specifications. IDTs will likely use various modelling platforms, including tools that support ODE, agent-based, discrete, stochastic, or data-driven models. Not all of them are currently supported by community standards, thus there is a critical need to create standards for model specification for a much broader class of models, through close collaboration and discussions with the COMBINE community (37). A useful resource is also the EDITH standards collection for Virtual Human Twins in Health (38).

IV. ***Reusable:*** All IDT elements should be described in detail, comprising multiple attributes using standard metadata structures. Naming and annotating DT elements should be orchestrated by existing ontologies and newly designed controlled vocabularies, where needed. Transparency and accuracy in description are necessary for maximising reusability. Special to the IDT field, in relation to reusability, are the aspects of scalability and modularity. The IDT should be:

a) **Modular:** An IDT should be designed to be modular. In this context, it is constructed from component models that capture elementary features of the immune system, such as a particular function of a particular immune cell type or cytokine signalling in a particular physiological context. A modular IDT architecture allows for the integration of different components and models. Each IDT module can be derived from previously built models designed to represent specific aspects of the immune system; the modules can be designed from scratch to match unanswered questions, enabling flexibility to mix and match models as needed. This architecture supports easy updates, replacements, or additions. It allows the integration of new data and new data types and formats as they are discovered or developed, ensuring that the IDT can adapt to new research findings and evolving medical and biological knowledge, keeping it always up-to-date (39). Although this modular approach represents a rapid way to move forward, it would need unambiguous and stable standards for model construction and input and output formats. In addition to that, the granularity of the modules and the level of mechanistic details should be well described to allow seamless assembly of the foundation modules, across the same or different scales. Ideally, the IDT design should be based on standardised, well-annotated modules that can be assembled into adaptable models. Note that the modelling community has long recognised (40) that constructing a model in a plug-and-play fashion is a natural approach to managing model complexity and offers additional opportunities, such as the potential to reuse model components. In particular, the SBML Model Composition package (SBMLcomp) (41) was developed to enable a modeller to include submodels within an enclosing model and edit, delete, or replace elements of that submodel. The concept of modularity in IDTs is similar to the use of containers and container libraries in bioinformatics frameworks (42).

b) **Scalable:** Scalability is indispensable for integrating the different computational modules accounting for different scales and the computational power demanded for the simulations. A successful IDT requires a clear multilevel and multiscale organisation of

the immune response that would allow for simpler surrogate models when complexity is unnecessary (43), (44)). Additionally, the IDT infrastructure should be able to respond to an increase in data, number of models, or size of models relative to the immune system and the pathological context under consideration. It should include connections to HPC and cloud computing (45). While supercomputers represent hardware-enhanced machines, HPC uses distributed resources to combine storage, applications, computational power, and network resources. Cloud computing refers to delivering computing services over the internet to facilitate access to resources and economies in scaling. Combination of HPC and cloud computing could accelerate simulations at a large scale, thus significantly reducing the time to market for an IDT prototype.

The FAIR principles are intended to foster collaboration, accelerate scientific discovery, and maximise the value and impact of the data for research and innovation. Various modelling communities have manifested the need for comprehensive, accessible, reusable, interoperable and reproducible computational models in systems biology. A key aspect is creating the model metadata and the model annotation in community-supported and standardised formats (46,47,48). Model repositories such as BioModels would need (49) to be extended to support FAIR dissemination of IDT and its components. A successful example of a large-scale community effort that leveraged knowledge assembly, platform and tools interoperability and FAIR implementations is the C19DMap project (50,51). The use of community standards for graphical representations, computational models and input and output formats allowed the building of an impressive ecosystem of interoperable software .

While the IDT should in principle be a two-way information system, ethical questions arise regarding the accessibility of the IDT predictions for patients (9). In the proposed schema, the decision is not directly accessible by the patient, and it implies the presence of a control point, where a clinician (expert) uses the IDT's *in silico* results to make an informed decision, that is then communicated to the patient The level of accessibility to the IDT's predictions should be controlled, and this can be addressed with different user categories having different types of rights. In Figure 2, we offer a conceptual design of an IDT implementation.

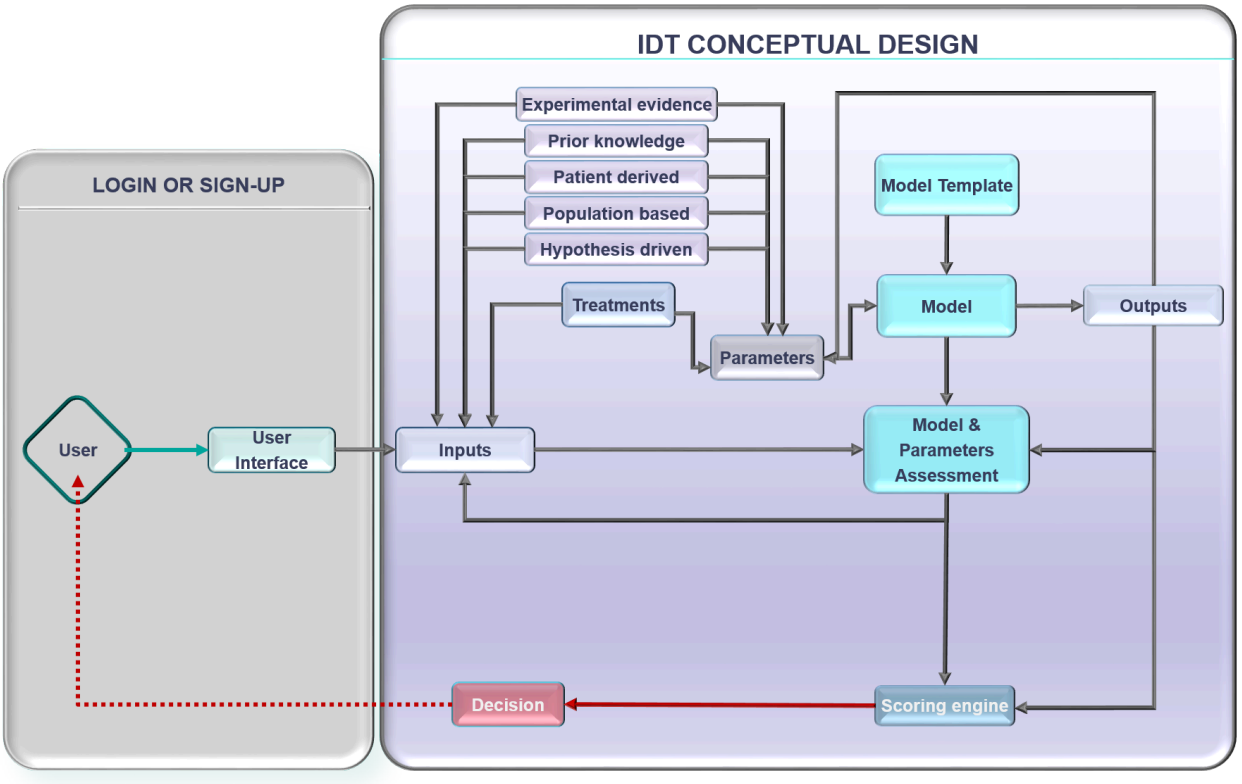

Figure 2. A minimalistic conceptual design of an IDT implementation. Producing a DT requires calibrating a computational model to data derived from a real-world patient. The connection to the real world is seen in the grey box to the left, where inputs of different types are generated for a particular individual and then passed to the virtual/computational model to personalise ("twin") that general model to the specific patient. The process of "twinning" involves parameterisation and a matching score to the real-world system by making predictions of how the real-world system propagates through time. This process iterates as new data becomes available and the DT is updated.

A stepwise, question-driven guide for the development and implementation of IDTs

A defining feature of the immune system is that it operates across scales, bridging molecules to organs' dynamics and spanning timescales from seconds to weeks, months or years. This implies that digital twins incorporating immune system functions must be multiscale by default. This is a critical feature for simulating disease progression and predicting prognostic and therapeutic outcomes, allowing it to dynamically capture macroscopic immune system behaviours and microscopic cellular and molecular interactions. Consequently, any generic model would need to take this feature into account. It must delineate scales from the intracellular to the cellular, tissue, organ, and organism levels. While multiscale modelling technology in biomedicine has made significant progress over the last decade (52), many theoretical problems will need to be solved, from software engineering to mathematics, including methodological challenges in sensitivity analysis or uncertainty quantification (53).

To begin designing an IDT, one could ask the following questions: is time-series data available? Does the disease progress within hours or years? Next, one should determine the level of granularity and complexity of the virtual representation that will be the core of the digital twin. In particular, one could ask, is there a real need for intracellular pathways, or

would tissue and organ crosstalks suffice? In addition, the appropriate formalism and modelling approach should be carefully chosen according to the available data and computational capabilities. Where quantitative data are limited, discrete, logic-based models can represent regulation and causal effects. However, if a quantitative answer is needed, for example, the highest dosage of an administered drug without the patient suffering severe side effects, then quantitative data is needed, and appropriate formalisms such as ordinary differential equations/partial differential equations (ODE/PDE) -based models could be employed (54).

Specific use cases will direct the degree of representation and detail required for the IDT. A significant endeavour is identifying the mechanism of interest and the core application ("fit-for-use"). Defining the specific use case can inform both the details required in the IDT and guide the necessary data linkages/types to personalise the IDT (including to help guide the development of new assay technologies to meet the desired goal of the IDT). This approach is consistent with the emphasis on modularity in IDTs (noted above) and allows the construction and deployment of clinically useful IDTs as new knowledge and technologies are developed. Mechanism-based simulations and subsequent experimental and clinical validation will allow for iterative improvement of models of the human immune system. These models will become increasingly more accurate and robust in their capacity to simulate the human immune system's reactivity against insults, and dysregulation in disease, and predict potential pharmacologic intervention points at different scales.

Integrating Artificial Intelligence (AI) models with mechanistic models for IDT construction

Data-driven solutions can bring valuable insights when the precise mechanism of interest is unknown but sufficient data is available. Deep learning (DL) models are playing increasingly pivotal roles in various domains (55, 56). However, their application still presents challenges, such as (i) the need for large amounts of data and computing resources; (ii) ethical and privacy issues, particularly concerning the potential misuse of AI models and the risk of perpetuating existing biases in patient data; and (iii) a complex regulatory landscape across countries and institutions, which severely limits the sharing of sensitive human data. Due to the latter, many algorithms are trained on small and homogeneous cohorts, leading to data overfitting and limited generalizability to new patient groups. Innovative ML methodologies are emerging to combat these challenges. For instance, generating synthetic data informed by mechanistic knowledge offers a way to augment datasets and mitigate data scarcity and imbalance (57). Furthermore, models can be contextualised to represent a wide range of demographics and conditions, enhancing the diversity in patient population representation (58). Foundational predictive AI models, built upon extensive multi-modal data encompassing scientific texts, molecular datasets, and biomedical knowledge graphs, are emerging in the field of biology and show great potential in facilitating all aspects of model engineering, from biocuration to training model parameters (59). Transfer learning (60), federated learning (61,62), and Explainable AI (XAI) (63) approaches lead to more reliable, safer and interpretable predictions of immune response upon perturbations or treatments. Finally, the current approach to 'Omics data analysis, especially in the single cell

data science, is undergoing a significant transformation with the advent of the new generation of generative (64) and causal AI methods (65). These methods are effectively bridging the gap between the data-driven approach (66,67) and the mechanistic modelling, marking a notable shift in the traditional distinction between them.

A promising avenue for future development is the integration of AI and Mechanism-based Multiscale Models (MSMs) (65). Mechanistic models excel in inferring causal relationships based on known biological mechanisms (68, 69), while AI models can help identify patterns and correlations within extensive datasets (20). Hybrid IDTs could combine the robustness and interpretability of mechanistic models with the capability of AI models for extracting information from large data sets. Furthermore, hybrid models can address data scarcity while enhancing the robustness of the model, as demonstrated in physics with physics-informed neural networks (PINNs) (70), i.e., neural networks constrained to comply with established physical principles. Constraining the model with prior knowledge typically boosts accuracy even with limited data and enhances its generalizability to new scenarios. However, applying similar approaches in biology is challenging, as biological systems are typically described in qualitative terms, e.g., using networks to describe intracellular signalling or transcriptional processes or statistical and probabilistic approaches to describe random interactions at the systems, organs, tissue and cellular levels. Despite the difficulty of integrating qualitative knowledge into deep learning models, proof of concept cases have already been demonstrated, such as with pathway-aware multi-layered hierarchical networks used to classify cancer patients (71) or visible neural networks that can reproduce the inner workings of eukaryotic cells (72).

While further research is essential to integrate AI and mechanistic models to replicate multiscale immunological processes, hybrid IDTs could be particularly effective in predicting and suggesting therapeutic interventions targeting specific mechanisms (73). In fact, immune system functions that have effects across several levels and their complex interactions are well suited to be addressed by hybrid IDTs and are problems that are currently not solved by either AI or MSMs. Hybrid IDTs could extract the most relevant information and identify relevant players across levels and provide robust and interpretable predictions on intervention targets of diseases.

Implementing Immune Digital Twins in the Study of Complex Human Pathologies

The human immune system is central to several classes of disease, such as infectious, autoimmune, cancer, and others, as a causal and modulating factor. This section will outline how an IDT could be implemented as a valuable tool in different pathologies with distinct and disease-specific characteristics. The examples described here will also be hands-on use cases for developing working IDT implementations.

To achieve tangible results, the collaboration of multiple stakeholders is needed. Moreover, a change of mindset is needed so that the computational biologist, modeller, or bioinformatician participates from the beginning in the experimental design and the study setup. Traditionally, clinicians, immunologists, and experimental biologists identified hallmarks of the disease, biomarkers and pathways affected, organ and body systemic manifestations, measurable factors used in the clinic, and experimental techniques that

could be employed in the short or long-term to enrich the molecular, genomic, metabolic and clinical profile of the patients. Bioinformaticians and computational biologists then used integrative methods to analyse the data available and provide coherent links and possible abstractions that could capture the essential characteristics of the system. However, an early inclusion of the computational and mathematical modellers in the study design could have a great impact in making sure that the minimal set of measurements for building a reliable model is factored in. Likewise, exchanges and discussions early on in a research project would allow for a maximum of comprehension of the disease mechanisms and questions at stake. Besides IDT design and implementation, bioengineers can help identify and manufacture critical biosensor technologies that could be implemented into the IDT computational ecosystem. Partnerships with startups could accelerate the production of prototypes, and the industry could contribute by providing infrastructure for the necessary scaling and support for bench-to-market pilot studies. The workshop "Building Immune Digital Twins" successfully brought together representatives of all stakeholder communities to advance toward IDT preclinical and clinical implementations for various complex human pathologies.

During the workshop, the participants worked on four different use cases, briefly presented in the following paragraphs. The basic steps across scientific fields that are required for a full-circle IDT implementation are shown briefly in Figure 3.

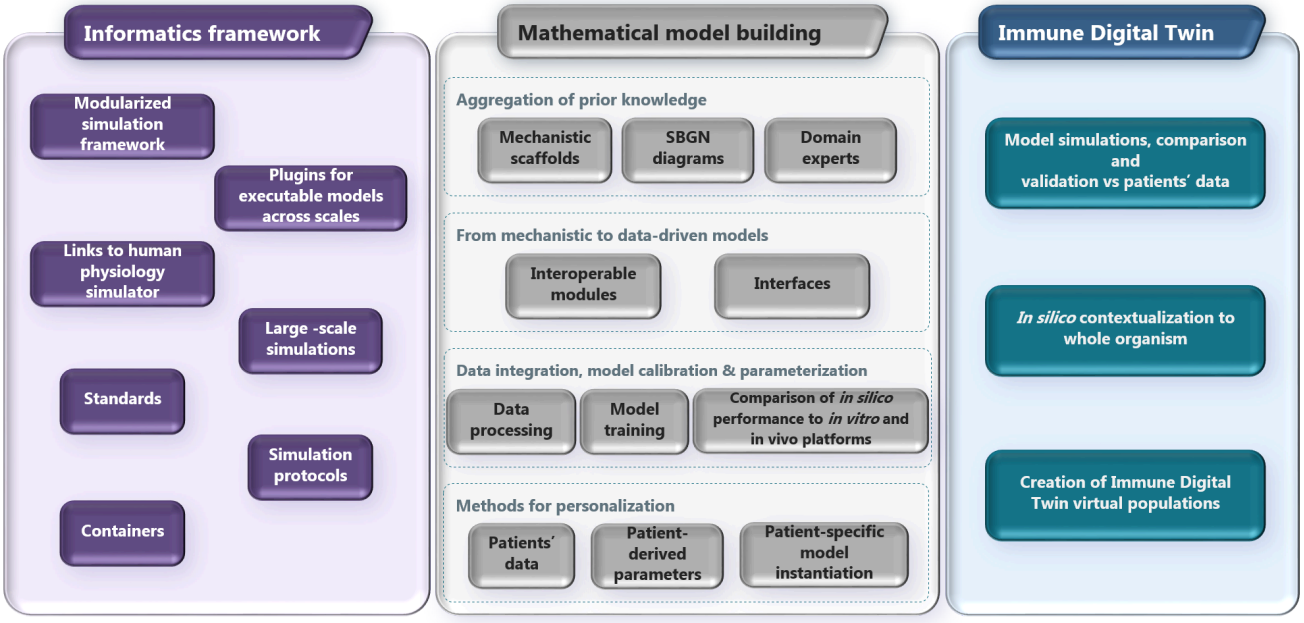

Figure 3. Different steps across scientific fields are required for a full-circle IDT implementation.

Infectious pneumonia Immune Digital Twin (IP-IDT) paradigm: Infectious pneumonia inflames the air sacs in one or both lungs, which can quickly become life-threatening. Pathogenic insults such as viruses (e.g., influenza, coronaviruses), fungi (*A. fumigatus),* or bacteria (*K. pneumoniae, S. aureus*) can cause the lungs' air sacs (alveoli) to become inflamed and filled up with fluid or pus. The infected host will then struggle to get oxygen into the bloodstream. Mounting a robust immune response is crucial for the clearance of the

pathogen and the resolution of inflammation. On the other hand, an overpowering response can lead to acute respiratory distress syndrome (ARDS), as exemplified in severe COVID-19 patients, and/or unresolved formation of scar tissue. The alveoli will be essential in IP-IDT. Alveolar macrophages are the sentinels of the alveoli; their functions are broad, e.g., phagocytosis, clearing debris, resolution of inflammatory responses, and tissue remodelling (74). Pulmonary macrophages are diverse, including tissue-resident alveolar macrophages that maintain immune balance and monocyte-derived alveolar macrophages that adapt to the microenvironment (75). Recently, alveolar epithelial cells were found to actively participate in innate immunity by directly communicating with alveolar macrophages, phagocytosis of pathogens, and/or recruiting other leukocytes to the injury site. If the pneumonia lasts several days, the importance of adaptive immunity, such as T-cells and B cells, must be included (76). A DT that captures the relevant lung biology and can be calibrated to individual patient characteristics, such as their immune profile or the extent of damage to the lung epithelium from the infection, could serve as a decision support tool for the ICU physician. Implementing an IP-IDT could include models at the cell and tissue level combined with a physiological model of oxygen exchange and blood flow. An early prototype of a computational model underlying such an IDT is published in (77). Other interventions in which physiological details will be necessary include the effects of prone positioning of patients in ICU beds or mechanical ventilator-induced injury.

The Rheumatoid Arthritis Immune Digital Twin paradigm (RA-IDT): Rheumatoid arthritis is an autoimmune complex disease that affects the articular joints of the human body. The disease is multifactorial, with genetic and environmental factors pivotal in the disease pathogenesis. RA's aetiology is unknown, and the treatment is primarily symptomatic. The disease affects the immune system, which mistakenly attacks the synovial lining of the joints, causing inflammation, cartilage destruction and bone erosion. The autoimmune component is central; however, other mechanisms, both immunologic and tissue-derived, clearly contribute to its onset and progression (78). In the first stages of the disease, leukocytes infiltrate the synovial compartment of the joint, secreting pro-inflammatory mediators that induce inflammatory cascades. Interactions between the joint's resident cells, such as fibroblast-like and macrophage-like synoviocytes, with the cells of the innate, like mastocytes, dendritic cells, etc., as well as cells of the adaptive immune system, such as T cells and B cells, contribute to the sustained inflammation and tissue damage. These conditions can also lead to a decrease in osteoblasts and an increase in osteoclasts and synoviocytes, leading to bone loss. If left untreated, the fulminant stage of the disease is described by a hyperplastic inflamed synovium, cartilage damage, bone erosion, and other systemic consequences (79).

The main objective of the RA-IDT would be to decipher the interplay between resident cells of the joint and immune cells in RA, which eventually leads to bone erosion, cartilage breakdown, and inflammation. A composable, multicellular, and multiscale model for RA could be a complex and sophisticated tool for studying the underlying mechanisms of RA and testing new treatments for the disease. In addition, it could provide valuable insights into the pathogenesis of the disease and help identify new targets for therapeutic

intervention. Recently, several models on the intra and inter-cellular level have been developed (80–85) that could serve as the core components of an RA-IDT. Moreover, given some shared characteristics, especially regarding bone erosion and cartilage destruction, between RA and Osteoarthritis (OA), OA models could also be contextualised and implemented in the RA-IDT (86,87). Modelling methods that couple signalling, gene regulation and metabolic fluxes are now available (80, 88) and can be combined with omics data technologies to create personalised instantiations. Hybrid modelling methods that allow for combinations of large-scale inter-cellular models with cell-level agent-based models could be employed to create a virtual joint. Biosensors that can measure matrix degradation and bone erosion could also be valuable tools in diagnosing and managing RA. Biosensor results and patient-reported outcomes and scores could be used to assess the patient's joint and bone health and monitor changes over time (89). The RA-IDT could also be integrated with other technologies, such as imaging techniques or wearable non-invasive sensors (smart watches, smartphone applications), to provide a more comprehensive picture of the patient's joint and overall health.

For a successful RA-IDT implementation, the collaboration between rheumatologists, computational and experimental biologists, and engineers is indispensable for tackling the multiple facets of this debilitating disease. The response rate to current therapies is estimated to be around 40% (90) demonstrating the pressing need for accelerating innovative and powerful technologies for personalised care.

The Sepsis-IDT paradigm for therapeutic discovery: Sepsis is a syndrome that often arises from severe infection but can also arise from severe trauma or burns, where a disordered immune response can lead to both early proinflammatory collateral tissue damage/organ dysfunction (“cytokine storm”) (91) and later immune incompetence due to a prolonged anti-inflammatory state, leading to increased susceptibility to nosocomial infections (92). In this context, several bacteria (*P. aeruginosa*, *A. baumannii*, and *Enterobacterales* Multi-Drug Resistant) represent the leading causes of nosocomial infections (93,94), even if such severe infection could be due to many other bacterial or fungal pathogens (i.e. *S. aureus* and *C. auris*, respectively) (95,96). The primary goal of a Sepsis IDT is the development of a high-fidelity computational approximation of a specific human, employing techniques from machine learning and artificial intelligence to personalise computational models (97) to be used for deriving precision therapeutic strategies to return the patient to a state of health and full immune/inflammatory functionality (98). As sepsis is a systemic disease that can lead to multiple organ dysfunction, a Sepsis IDT will necessarily represent and integrate those organs at risk: the immune system, lung, liver, kidney, gut, and cardiovascular system. Given the modular nature of IDTs, each organ system can be cast as a sub-compartment model of the entire IDT for an early example of such an architecture that represents the gut-lung axis of sepsis (99). The instantiation of an individualised sepsis DT requires that it be linked to a suite of clinical and laboratory measurements reflecting dynamic molecular profiling of the patient’s immune state and determinants of trajectories of organ function. For the latter, the Sequential Organ Failure Score (SOFA) (100) and its variants (101), which have been used to predict mortality and outcomes in the critically ill population, are generated by readily

obtainable clinical measurements (e.g., serum bilirubin or creatinine, blood oxygen concentration) can be sequentially measured to update the IDT. An initial example of dynamic multiscale molecule-to-organ integration can be seen in a Critical Illness Digital Twin that links plasma cytokine levels with respiratory SOFA scores (102). Data collection to establish the inflammatory state of the patient will ultimately require the development of more advanced sensors and assays; strides in this direction are being made as real-time blood-serum cytokine measurements (103) become more readily obtainable to capture the relatively rapid temporal dynamics (minutes to hours) of sepsis and be used to inform the molecular scale inputs of a Sepsis IDT (97). Numerous challenges must be overcome to develop a generalisable Sepsis IDT, from insufficient knowledge regarding protein interactions in the inflammatory/immune signalling network to technological developments in sensor technologies for real-time proteomic readouts of patient state. Furthermore, the deep interaction between host and bacterial pathogens in sepsis and the immune response in systemic infection should be considered in sepsis IDT to understand the molecular mechanisms, considering different pathogens, outcomes, and symptoms (104). In fact, the transcriptomics analysis at single cells, performed on blood immune cell types of patients with sepsis, allowed the identification of a cytological signature of the disease (i.e., CD14+ Monocytes) (105), as well as highlighted a differential molecular response to different pathogens in early stages of systemic infection (106).

The Onco-IDT paradigm: The ability to evade immune surveillance and destruction is now a well-recognised hallmark of cancer (107), and its targeting through immunotherapy approaches has already significantly improved cancer outcomes for several cancer types (108). During oncogenesis, the immune system activates a multifaceted response involving both the innate and adaptive immune systems. However, cancer cells and their environment progressively evade immune surveillance, for instance, by down-regulating major histocompatibility complexes (MHC) and upregulating immune checkpoint proteins (109), (110). Cancerous, immune, and stromal cells are the critical components of the tumour microenvironment (TME); therefore, their crosstalk should be represented in an Onco-IDT as the TME is crucial for carcinogenesis and acquiring malignant traits (111, 112). An ideal Onco-IDT should include elements like the TME, neoangiogenesis, the creation of pre-metastatic niches, and ultimately, system-level information like blood and lymphatic transport that together with cancer-intrinsic metastasis-enabling molecular programs underpin systemic disease.

Today, personalised therapy in oncology is making progress in identifying cancer-driver mutations for each patient (113) or cellular patterns linked with disease state/progression (114, 115, 116). Extensive cell phenotyping, genetic testing, or even sequencing of tumour material is possible but requires tissue obtained by biopsy. The biopsy-based molecular subtyping provides direct data on a tumour's current state and microenvironment, and single-cell techniques are powerful tools to capture natural and pharmacologically induced tumour immunity (117). However, cancer genomics-guided approaches harnessing or targeting the immune system are incomplete and still in their infancy, and, typically, surrogate markers for the immune system engagement are used. Also, since biopsies are

invasive, they are usually reserved for diagnostics, and only performed on rare occasions after that, which limits their ability to track tumour development over time and sample intra- and inter-tumoral heterogeneity.

On the other hand, data and measurements from non or semi-invasive interventions can be integrated with the biopsy data points. These types of data and measurements include electronic health records, radiological imaging, and serological or molecular data that could inform about the inflammatory state of the patient. In addition, blood samples can provide insights into drug pharmacokinetic (PK) relevant processes, while wearables can offer additional information regarding vital signs, body temperature and physical activity levels. Notably, progress has been made to provide detailed insight into the molecular features of the tumour from non-invasive techniques by investigating circulating tumour cells and circulating tumour DNA as surrogates for the biopsied tumours.

The analysis of each data type independently poses a challenge, and becomes almost insurmountable when numerous data types are considered together. Onco-IDTs are envisioned to provide a solution by integrating a multitude of data measurements across numerous data types and time points, and provide the means for a systems level approach to patient-centred medical insights, disease simulation for prognostics, and *in silico* experimentation of therapies.

An effective Onco-IDT will provide oncologists with a dynamic clinical decision support platform, aiding prognosis and disease management. More specifically, the Onco-IDT could contribute significantly in prognostic predictions regarding disease course, considering factors like metastatic capacity and patient survival. Additionally, it could offer actionable insights concerning therapeutic interventions, involving the selection of the most effective therapy that maximises benefits while minimising side effects and adverse outcomes. Therapeutic decisions might also be optimised to select effective monotherapies versus combination therapies, drug doses, and treatment schedules. Interestingly, the US National Cancer Institute and the US Department of Energy started to explore the development and implementation of predictive Cancer Patient Digital Twins for personalised treatment (118).

IDTs in drug discovery

Drug development is costly and slow. The costs include expenses from the early stages of research and discovery through clinical development, regulatory approval, and post-marketing surveillance. The majority of candidate targets and drugs experience failure in the early stages, contributing significantly to the overall cost of delivering more successful candidates. Therefore, optimising these earlier stages holds transformative potential in the pharmaceutical industry. A strong consensus among experts supports the opinion that the involvement of digital twins in this transformation can be essential (119).

The early stages of drug development increasingly involve the use of growing volumes of molecular, imaging, and clinical data. This trend surpasses the human capacity to provide a rational and holistic approach to decision-making in the identification of therapeutic targets and their subsequent triaging for further validation based on complex causal relationships. The introduction of a specialised form of drug development digital twins (DDDT) has the potential to be a game-changer (120). Major pharmaceutical companies

believe in the innovative and transformative potential of DTs in drug discovery and actively invest in their development (121). This approach is geared towards discovering new therapies, in contrast to DTs focusing on optimising the existing process for a known treatment. DDDTs, by design, are characterised by a specific set of challenging requirements, with one of the most crucial being the granularity of representation of molecular mechanisms (i.e., mechanisms of action of drugs and targets). Current mathematical models of biological processes are usually too abstract to be applicable to the task of the DDDT. For instance, a mathematical model of viral infection with an aggregated term for "pro-inflammation" cannot evaluate the effect of a specific drug because "pro-inflammation" is not a mechanistic target for a chemical compound. Consequently, DDDTs must be inherently multiscale, initially developed at the level of intracellular components and genes before expanding to the physiological scale (120). This is particularly crucial in the field of immunology.

Moreover, we can envision further specialisation of DDDTs based on various tasks in the early drug discovery process. These tasks include 1) identifying targets and their combinations, as well as determining the most promising treatment modalities (encompassing not only small chemical compounds but also antibody-drug conjugates, various types of biologics, and gene or cell therapies); deciding on the level at which the target should be affected (whether directly, through its RNA, or its involvement in protein-protein interactions), 2) experimental target and drug validation, aiding in identifying the most informative experimental systems (such as cell lines or organoids) and experiment designs, 3) repurposing drugs for alternative indications in case of a failure for the primary one, 4) delivering drugs by integrating pharmacokinetic models into the global *in silico* models of treatment and taking into account safety aspects early in the process, 5) finally, there might be flavours of DDDTs aimed at optimising the process of drug production, with notable examples like Sanofi exploring the use of digital twins for vaccine manufacturing (122).

All these DT specialisations require specific designs, functionalities, and connections to the existing wealth of public and proprietary data. Furthermore, the scope application of DTs goes beyond the scope of early drug discovery or production; e.g. virtual populations of patient DTs can be used for running *in silico* clinical trials that can accompany or be used for designing real-life trials (123). One recent example is the Universal Immune Simulator (UIS) (124). The European Medicines Agency (EMA) provided a letter of support for the use of the UIS as a simulation platform to predict how the circulating interferon gamma (IFNγ) changes over time as a function of the treatment dose in a cohort of virtual patients, to select the doses to be tested in escalating dose phase IIa trials of new therapeutic whole cell / fragmented based vaccines against a number of diseases (125). Whereas more work is required before qualification advice can be given, it does show that EMA believes this is a genuine possibility. Recently, a book was published focusing on best practices for the use of computational modeling and simulation in the regulatory process of biomedical products, showcasing the need to address policy and implementation early on in the DT design (126). Different types of challenges associated with the DT development and implementation are depicted in Figure 4.

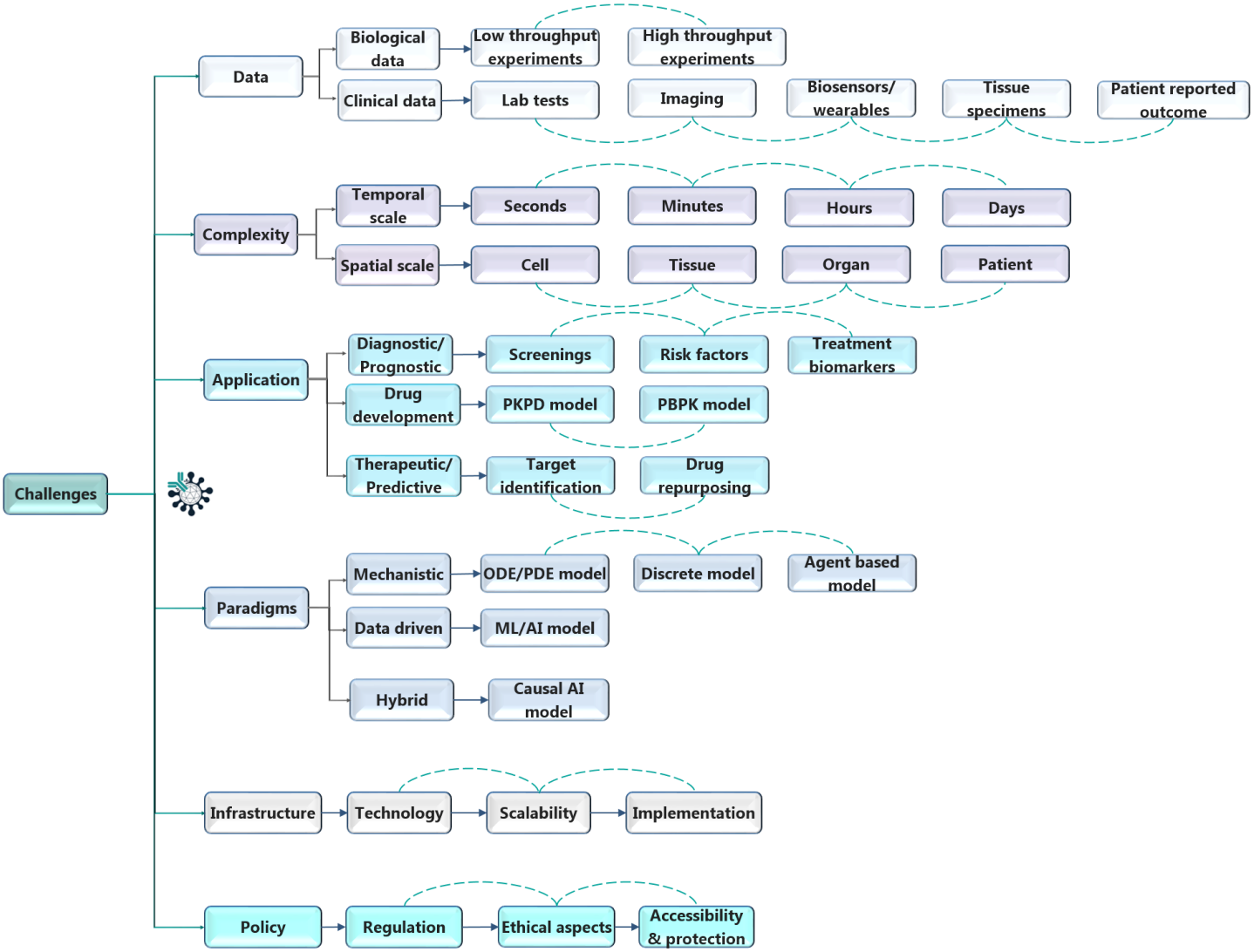

Figure 4. Key challenges in developing and implementing IDTs in pre-clinical and clinical settings .

Perspectives

The major challenge in building UDTs is bringing together clinicians, biologists, and mathematical modellers coming from academia and the biotech industry, from Europe, the U.S., and many other countries in deep, prolonged discussions and collaborations around DT technology and implementation. Very close collaboration and coordination among the different communities are required for success. Our workshop, together with similar efforts toward the same direction (127) show that the right intellectual environment and a well-articulated common purpose might help overcome barriers and set the basis of international multi-partner collaborations between academia and industry. The final success of these multi-partner collaborations crucially depends on the ability to have a common culture, vocabulary, and understanding of the potential and current limitations of DT technology. Over the next two years, the Working Group that emerged from the Institut Pascal workshop aims to create additional venues for such interdisciplinary interactions at different scales and move forward with the development and implementation of IDTs.

Acknowledgements

This work is mainly the result of the three-week workshop on Building Immune Digital Twins, which was made possible by the support of the Institut Pascal, University of Paris-Saclay, France, via the program *Investissements d'avenir, ANR-11-IDEX-0003-0*. We would also like to thank Genopole for the support of the workshop. Lastly, we thank the keynotes, and every participant and member of the community for the discussions and brainstorming sessions that helped shape some of the concepts described here.

Individual Funding

AN was supported by a public-private partnership grant (CIFRE contract, n° 2020/0766) with SANOFI-AVENTIS R&D. RL acknowledges financial support from the Defense Advanced Research Projects Agency (grant HR00112220038), and the National Institutes of Health (grants R01 GM127909, R01 AI135128, and R01 HL169974). LCM, BS and LG acknowledge support from the European Commission (grants H2020-SC1-BHC-11-2020 963845 and DigitalEurope EDITH-CSA 101083771). LLF acknowledges financial support from the Defense Advanced Research Projects Agency (grant HR00112220038). LH acknowledged support from NIH/NCI Transition Career Development Award to Promote Diversity (K22-CA237857). LSV acknowledges support from the National Institutes of Health (grants R01 AI135128, R01 HL169974, K25 AI175668).PH acknowledges support from Prime Ministers' Research Fellowship (PMRF). HB acknowledges the support of ZonMw (Grant No. 10430012010015) NN acknowledges support from the Knowledge Foundation (20200017) and The Swedish Fund for Research Without Animal Experiments. AM, JCC and AV acknowledge support from the European Commission (grants CREXDATA 101092749, PerMedCoE 951773 and EDITH-CSA 101083771). FM acknowledge support from the Italian Ministry of Health (grant “Ricerca Corrente” and “5 per 1000–2021” Grant No. 5M-2021-23683787) and European Commission with HORIZON programme (Grant No. 101046203—BY-COVID). JAG acknowledges support from the National Institutes of Health (grants U24 EB028887) and the National Science Foundation (grants NSF 2303695, NSF 2120200 and NSF 1720625)

Conflict of interest

The authors declare no conflict of interest.

Authors’ contributions

AN was the main organiser, and RL was the co-organiser of the workshop. AN and RL secured funding, designed the programme and AN coordinated activities. All authors participated in the working groups and writing activities during and/or after the workshop, either in person or remotely. AN and RL prepared the manuscript draft. All authors contributed to the scientific content of the meeting and participated in manuscript writing and editing. AN, SS, JG, LL, RSMS and BS prepared figures.

References

1. Committee on Foundational Research Gaps and Future Directions for Digital Twins, Board on Mathematical Sciences and Analytics, Committee on Applied and Theoretical Statistics, Computer Science and Telecommunications Board, Board on Life Sciences, Board on Atmospheric Sciences and Climate, et al. Foundational research gaps and future directions for digital twins. Washington, D.C.: National Academies Press; 2023.
2. Consortium E. EDITH CSA Deliverable 3.2: first draft of the VHT roadmap. Zenodo. 2023;
3. Viceconti M, De Vos M, Mellone S, Geris L. Position paper From the digital twins in healthcare to the Virtual Human Twin: a moon-shot project for digital health research. IEEE J Biomed Health Inform. 2023 Oct 11;PP.
4. Sahal R, Alsamhi SH, Brown KN. Personal Digital Twin: A Close Look into the Present and a Step towards the Future of Personalised Healthcare Industry. Sensors. 2022 Aug 8;22(15).
5. Björnsson B, Borrebaeck C, Elander N, Gasslander T, Gawel DR, Gustafsson M, et al. Digital twins to personalize medicine. Genome Med. 2019 Dec 31;12(1):4.
6. Chang H-C, Gitau AM, Kothapalli S, Welch DR, Sardiu ME, McCoy MD. Understanding the need for digital twins' data in patient advocacy and forecasting oncology. Front Artif Intell. 2023 Nov 10;6:1260361.
7. Chaudhuri A, Pash G, Hormuth DA, Lorenzo G, Kapteyn M, Wu C, et al. Predictive digital twin for optimizing patient-specific radiotherapy regimens under uncertainty in high-grade gliomas. Front Artif Intell. 2023 Oct 11;6:1222612.
8. Heudel PE, Renard F, Attye A. [Digital twins in cancer research and treatment: A future for personalized medicine]. Bull Cancer. 2023 Oct;110(10):1085–7.
9. Iqbal JD, Krauthammer M, Biller-Andorno N. The use and ethics of digital twins in medicine. J Law Med Ethics. 2022;50(3):583–96.
10. Rahmim A, Brosch-Lenz J, Fele-Paranj A, Yousefirizi F, Soltani M, Uribe C, et al. Theranostic digital twins for personalized radiopharmaceutical therapies: Reimagining theranostics via computational nuclear oncology. Front Oncol. 2022 Dec 15;12:1062592.
11. Pesapane F, Rotili A, Penco S, Nicosia L, Cassano E. Digital twins in radiology. J Clin Med. 2022 Nov 4;11(21).
12. Jung A, Gsell MAF, Augustin CM, Plank G. An Integrated Workflow for Building Digital Twins of Cardiac Electromechanics-A Multi-Fidelity Approach for Personalising Active Mechanics. Mathematics (Basel). 2022 Mar 1;10(5):823.
13. Corral-Acero J, Margara F, Marciniak M, Rodero C, Loncaric F, Feng Y, et al. The "Digital Twin" to enable the vision of precision cardiology. Eur Heart J. 2020 Dec 21;41(48):4556–64.
14. Abdollahi H, Saboury B, Soltani M, Shi K, Uribe C, Rahmim A. Radiopharmaceutical therapy on-a-chip: a perspective on microfluidic-driven digital twins towards personalized cancer therapies. Sci Bull (Beijing). 2023 Sep 30;68(18):1983–8.
15. Trayanova NA, Prakosa A. Up digital and personal: how heart digital twins can transform heart patient care. Heart Rhythm. 2023 Oct 21;
16. Zhang Y, Zhang K, Prakosa A, James C, Zimmerman SL, Carrick R, et al. Predicting Ventricular Tachycardia Circuits in Patients with Arrhythmogenic Right Ventricular Cardiomyopathy using Genotype-specific Heart Digital Twins. medRxiv. 2023 Aug 1;
17. Winter PD, Chico TJA. Using the Non-Adoption, Abandonment, Scale-Up, Spread, and Sustainability (NASSS) Framework to Identify Barriers and Facilitators for the Implementation of Digital Twins in Cardiovascular Medicine. Sensors. 2023 Jul 12;23(14).
18. Clarke MA, Fisher J. Executable cancer models: successes and challenges. Nat Rev Cancer. 2020 Jun;20(6):343–54.
19. Chakshu NK, Sazonov I, Nithiarasu P. Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis. Biomech Model Mechanobiol. 2021 Apr;20(2):449–65.
20. Coorey G, Figtree GA, Fletcher DF, Snelson VJ, Vernon ST, Winlaw D, et al. The health digital

twin to tackle cardiovascular disease-a review of an emerging interdisciplinary field. npj Digital Med. 2022 Aug 26;5(1):126.
21. Gillette K, Gsell MAF, Prassl AJ, Karabelas E, Reiter U, Reiter G, et al. A Framework for the generation of digital twins of cardiac electrophysiology from clinical 12-leads ECGs. Med Image Anal. 2021 Jul;71:102080.
22. Shamanna P, Joshi S, Shah L, Dharmalingam M, Saboo B, Mohammed J, et al. Type 2 diabetes reversal with digital twin technology-enabled precision nutrition and staging of reversal: a retrospective cohort study. Clin Diabetes Endocrinol. 2021 Nov 15;7(1):21.
23. Kovatchev B. The artificial pancreas in 2017: The year of transition from research to clinical practice. Nat Rev Endocrinol. 2018 Feb;14(2):74–6.
24. Brown SA, Kovatchev BP, Raghinaru D, Lum JW, Buckingham BA, Kudva YC, et al. Six-Month Randomized, Multicenter Trial of Closed-Loop Control in Type 1 Diabetes. N Engl J Med. 2019 Oct 31;381(18):1707–17.
25. Armeni P, Polat I, De Rossi LM, Diaferia L, Meregalli S, Gatti A. Digital Twins in Healthcare: Is It the Beginning of a New Era of Evidence-Based Medicine? A Critical Review. J Pers Med. 2022 Jul 30;12(8).
26. Handel A, La Gruta NL, Thomas PG. Simulation modelling for immunologists. Nat Rev Immunol. 2020 Mar;20(3):186–95.
27. Laubenbacher R, Niarakis A, Helikar T, An G, Shapiro B, Malik-Sheriff RS, et al. Building digital twins of the human immune system: toward a roadmap. npj Digital Med. 2022 May 20;5(1):64.
28. Li S, Dragan I, Tran VDT, Beulens J, Slieker R, Gerl MJ, et al. Multi-omics subgroups associated with glycaemic deterioration in type 2 diabetes: an IMI-RHAPSODY Study. Frontiers in Endocrinology. 2024 Feb 22;
29. Delfin C, Dragan I, Kuznetsov D, Tajes JF, Smit F, Coral DE, et al. A Federated Database for Obesity Research: An IMI-SOPHIA Study. Life (Basel). 2024 Feb 16;14(2).
30. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3:160018.
31. Wilkinson MD, Dumontier M, Jan Aalbersberg I, Appleton G, Axton M, Baak A, et al. Addendum: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2019 Mar 19;6(1):6.
32. Rocca-Serra P, Gu W, Ioannidis V, Abbassi-Daloii T, Capella-Gutierrez S, Chandramouliswaran I, et al. The FAIR Cookbook - the essential resource for and by FAIR doers. Sci Data. 2023 May 19;10(1):292.
33. Keating SM, Waltemath D, König M, Zhang F, Dräger A, Chaouiya C, et al. SBML Level 3: an extensible format for the exchange and reuse of biological models. Mol Syst Biol. 2020 Aug;16(8):e9110.
34. Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, et al. The systems biology graphical notation. Nat Biotechnol. 2009 Aug 7;27(8):735–41.
35. Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, et al. The BioPAX community standard for pathway data sharing. Nat Biotechnol. 2010 Sep 1;28(9):935–42.
36. Waltemath D, Adams R, Bergmann FT, Hucka M, Kolpakov F, Miller AK, et al. Reproducible computational biology experiments with SED-ML--the Simulation Experiment Description Markup Language. BMC Syst Biol. 2011 Dec 15;5:198.
37. Waltemath D, Golebiewski M, Blinov ML, Gleeson P, Hermjakob H, Hucka M, et al. The first 10 years of the international coordination network for standards in systems and synthetic biology (COMBINE). J Integr Bioinform. 2020 Jun 29;17(2–3).
38. FAIRsharing [Internet]. [cited 2024 Mar 4]. Available from: https://fairsharing.org/4787
39. Masison J, Beezley J, Mei Y, Ribeiro H, Knapp AC, Sordo Vieira L, et al. A modular computational framework for medical digital twins. Proc Natl Acad Sci USA. 2021 May 18;118(20).
40. Neal ML, Cooling MT, Smith LP, Thompson CT, Sauro HM, Carlson BE, et al. A reappraisal of how to build modular, reusable models of biological systems. PLoS Comput Biol. 2014 Oct

2;10(10):e1003849.
41. Smith LP, Hucka M, Hoops S, Finney A, Ginkel M, Myers CJ, et al. SBML Level 3 package: Hierarchical Model Composition, Version 1 Release 3. J Integr Bioinform. 2015 Sep 4;12(2):268.
42. Gruening B, Sallou O, Moreno P, da Veiga Leprevost F, Ménager H, Søndergaard D, et al. Recommendations for the packaging and containerizing of bioinformatics software. [version 2; peer review: 2 approved, 1 approved with reservations]. F1000Res. 2018 Jun 14;7.
43. Garira W. The research and development process for multiscale models of infectious disease systems. PLoS Comput Biol. 2020 Apr 2;16(4):e1007734.
44. Garira W. A complete categorization of multiscale models of infectious disease systems. J Biol Dyn. 2017 Dec;11(1):378–435.
45. Montagud A, Ponce de León M, Valencia A. Systems Biology at the giga-scale: large multi-scale models of complex, heterogeneous multicellular systems. Current Opinion in Systems Biology. 2021 Sep;100385.
46. Niarakis A, Waltemath D, Glazier J, Schreiber F, Keating SM, Nickerson D, et al. Addressing barriers in comprehensiveness, accessibility, reusability, interoperability and reproducibility of computational models in systems biology. Brief Bioinformatics. 2022 Jul 18;23(4).
47. Niarakis A, Kuiper M, Ostaszewski M, Malik Sheriff RS, Casals-Casas C, Thieffry D, et al. Setting the basis of best practices and standards for curation and annotation of logical models in biology-highlights of the [BC]2 2019 CoLoMoTo/SysMod Workshop. Brief Bioinformatics. 2021 Mar 22;22(2):1848–59.
48. Karr J, Malik-Sheriff RS, Osborne J, Gonzalez-Parra G, Forgoston E, Bowness R, et al. Model integration in computational biology: the role of reproducibility, credibility and utility. Front Syst Biol. 2022 Mar 7;2.
49. Malik-Sheriff RS, Glont M, Nguyen TVN, Tiwari K, Roberts MG, Xavier A, et al. BioModels-15 years of sharing computational models in life science. Nucleic Acids Res. 2020 Jan 8;48(D1):D407–15.
50. Niarakis A, Ostaszewski M, Mazein A, Kuperstein I, Kutmon M, Gillespie ME, et al. Drug-target identification in COVID-19 disease mechanisms using computational systems biology approaches. Front Immunol. 2024 Feb 13;14.
51. Ostaszewski M, Niarakis A, Mazein A, Kuperstein I, Phair R, Orta-Resendiz A, et al. COVID19 Disease Map, a computational knowledge repository of virus-host interaction mechanisms. Mol Syst Biol. 2021 Oct;17(10):e10387.
52. Kirschner DE, Hunt CA, Marino S, Fallahi-Sichani M, Linderman JJ. Tuneable resolution as a systems biology approach for multi-scale, multi-compartment computational models. Wiley Interdiscip Rev Syst Biol Med. 2014 May 9;6(4):289–309.
53. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008 Sep 7;254(1):178–96.
54. Le Novère N. Quantitative and logic modelling of molecular and gene networks. Nat Rev Genet. 2015 Mar;16(3):146–58.
55. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. Cell. 2019 May 16;177(5):1330-1345.e18.
56. Manica M, Oskooei A, Born J, Subramanian V, Sáez-Rodríguez J, Rodríguez Martínez M. Toward Explainable Anticancer Compound Sensitivity Prediction via Multimodal Attention-Based Convolutional Encoders. Mol Pharm. 2019 Dec 2;16(12):4797–806.
57. An G, Cockrell C. Generating synthetic multidimensional molecular time series data for machine learning: considerations. Front Syst Biol. 2023 Jul 25;3.
58. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. Nat Biomed Eng. 2022 Dec;6(12):1330–45.
59. Theodoris CV, Xiao L, Chopra A, Chaffin MD, Al Sayed ZR, Hill MC, et al. Transfer learning enables predictions in network biology. Nature. 2023 Jun;618(7965):616–24.
60. Weber A, Born J, Rodriguez Martínez M. TITAN: T-cell receptor specificity prediction with bimodal attention networks. Bioinformatics. 2021 Jul 12;37(Suppl_1):i237–44.

61. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. npj Digital Med. 2020 Sep 14;3:119.
62. Dayan I, Roth HR, Zhong A, Harouni A, Gentili A, Abidin AZ, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med. 2021 Oct;27(10):1735–43.
63. Novakovsky G, Dexter N, Libbrecht MW, Wasserman WW, Mostafavi S. Obtaining genetics insights from deep learning via explainable artificial intelligence. Nat Rev Genet. 2023 Feb;24(2):125–37.
64. Bordukova M, Makarov N, Rodriguez-Esteban R, Schmich F, Menden MP. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. Expert Opin Drug Discov. 2024 Jan 8;19(1):33–42.
65. Tejada-Lapuerta A, Bertin P, Bauer S, Aliee H, Bengio Y, Theis FJ. Causal machine learning for single-cell genomics. arXiv. 2023;
66. Dong M, Wang B, Wei J, de O Fonseca AH, Perry CJ, Frey A, et al. Causal identification of single-cell experimental perturbation effects with CINEMA-OT. Nat Methods. 2023 Nov 2;20(11):1769–79.
67. Qiu X, Zhang Y, Martin-Rufino JD, Weng C, Hosseinzadeh S, Yang D, et al. Mapping transcriptomic vector fields of single cells. Cell. 2022 Feb 17;185(4):690-711.e45.
68. Garira W, Muzhinji K. Application of the replication-transmission relativity theory in the development of multiscale models of infectious disease dynamics. J Biol Dyn. 2023 Dec;17(1):2255066.
69. Garira W, Muzhinji K. The universal theory for multiscale modelling of infectious disease dynamics. Mathematics. 2023 Sep 11;11(18):3874.
70. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019 Feb;378:686–707.
71. Elmarakeby HA, Hwang J, Arafeh R, Crowdis J, Gang S, Liu D, et al. Biologically informed deep neural network for prostate cancer discovery. Nature. 2021 Oct;598(7880):348–52.
72. Ma J, Yu MK, Fong S, Ono K, Sage E, Demchak B, et al. Using deep learning to model the hierarchical structure and function of a cell. Nat Methods. 2018 Apr;15(4):290–8.
73. Martínez MR, Barberis M, Niarakis A. Computational modelling of immunological mechanisms: From statistical approaches to interpretable machine learning. ImmunoInformatics. 2023 Jul;100029.
74. Rodriguez-Rodriguez L, Gillet L, Machiels B. Shaping of the alveolar landscape by respiratory infections and long-term consequences for lung immunity. Front Immunol. 2023 Apr 4;14:1149015.
75. Hou F, Xiao K, Tang L, Xie L. Diversity of macrophages in lung homeostasis and diseases. Front Immunol. 2021 Sep 24;12:753940.
76. Kumar V. Pulmonary Innate Immune Response Determines the Outcome of Inflammation During Pneumonia and Sepsis-Associated Acute Lung Injury. Front Immunol. 2020 Aug 4;11:1722.
77. Ribeiro HA, Vieira LS, Scindia Y, Adhikari B, Wheeler M, Knapp A, et al. Multi-scale mechanistic modelling of the host defence in invasive aspergillosis reveals leucocyte activation and iron acquisition as drivers of infection outcome. J R Soc Interface. 2022 Apr 13;19(189):20210806.
78. Firestein GS, McInnes IB. Immunopathogenesis of rheumatoid arthritis. Immunity. 2017 Feb 21;46(2):183–96.
79. Guo Q, Wang Y, Xu D, Nossent J, Pavlos NJ, Xu J. Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. Bone Res. 2018 Apr 27;6:15.
80. Aghakhani S, Soliman S, Niarakis A. Metabolic reprogramming in Rheumatoid Arthritis Synovial Fibroblasts: A hybrid modeling approach. PLoS Comput Biol. 2022 Dec 12;18(12):e1010408.
81. Zerrouk N, Alcraft R, Hall BA, Augé F, Niarakis A. Large-scale computational modelling of the M1 and M2 synovial macrophages in rheumatoid arthritis. NPJ Syst Biol Appl. 2024 Jan 26;10(1):10.

82. Miagoux Q, Singh V, de Mézquita D, Chaudru V, Elati M, Petit-Teixeira E, et al. Inference of an Integrative, Executable Network for Rheumatoid Arthritis Combining Data-Driven Machine Learning Approaches and a State-of-the-Art Mechanistic Disease Map. J Pers Med. 2021 Aug 12;11(8).
83. Singh V, Naldi A, Soliman S, Niarakis A. A large-scale Boolean model of the rheumatoid arthritis fibroblast-like synoviocytes predicts drug synergies in the arthritic joint. NPJ Syst Biol Appl. 2023 Jul 15;9(1):33.
84. Macfarlane FR, Chaplain MAJ, Eftimie R. Quantitative predictive modelling approaches to understanding rheumatoid arthritis: A brief review. Cells. 2019 Dec 27;9(1).
85. Moise N, Friedman A. Rheumatoid arthritis - a mathematical model. J Theor Biol. 2019 Jan 14;461:17–33.
86. Kerkhofs J, Leijten J, Bolander J, Luyten FP, Post JN, Geris L. A qualitative model of the differentiation network in chondrocyte maturation: A holistic view of chondrocyte hypertrophy. PLoS ONE. 2016 Aug 31;11(8):e0162052.
87. Lesage R, Ferrao Blanco MN, Narcisi R, Welting T, van Osch GJVM, Geris L. An integrated in silico-in vitro approach for identifying therapeutic targets against osteoarthritis. BMC Biol. 2022 Nov 9;20(1):253.
88. Aghakhani S, Niarakis A, Soliman S. MetaLo: metabolic analysis of Logical models extracted from molecular interaction maps. J Integr Bioinform. 2024 Feb 6;
89. Creagh AP, Hamy V, Yuan H, Mertes G, Tomlinson R, Chen W-H, et al. Digital health technologies and machine learning augment patient reported outcomes to remotely characterise rheumatoid arthritis. npj Digital Med. 2024 Feb 12;7(1):33.
90. Humby F, Durez P, Buch MH, Lewis MJ, Rizvi H, Rivellese F, et al. Rituximab versus tocilizumab in anti-TNF inadequate responder patients with rheumatoid arthritis (R4RA): 16-week outcomes of a stratified, biopsy-driven, multicentre, open-label, phase 4 randomised controlled trial. Lancet. 2021 Jan 23;397(10271):305–17.
91. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 Feb 23;315(8):762–74.
92. Schefold JC, Hasper D, Volk HD, Reinke P. Sepsis: time has come to focus on the later stages. Med Hypotheses. 2008 Aug;71(2):203–8.
93. Antimicrobial resistance surveillance in Europe.
94. WHO publishes list of bacteria for which new antibiotics are urgently needed [Internet]. [cited 2024 Feb 25]. Available from: https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed
95. About Candida auris (C. auris) | Candida auris | Fungal Diseases | CDC [Internet]. [cited 2024 Feb 25]. Available from: https://www.cdc.gov/fungal/candida-auris/candida-auris-qanda.html
96. Staphylococcus aureus in Healthcare Settings | HAI | CDC [Internet]. [cited 2024 Feb 25]. Available from: https://www.cdc.gov/hai/organisms/staph.html
97. Cockrell C, An G. Utilizing the Heterogeneity of Clinical Data for Model Refinement and Rule Discovery Through the Application of Genetic Algorithms to Calibrate a High-Dimensional Agent-Based Model of Systemic Inflammation. Front Physiol. 2021 May 19;12:662845.
98. Cockrell C, Larie D, An G. Preparing for the next pandemic: Simulation-based deep reinforcement learning to discover and test multimodal control of systemic inflammation using repurposed immunomodulatory agents. Front Immunol. 2022 Nov 21;13:995395.
99. An G. Introduction of an agent-based multi-scale modular architecture for dynamic knowledge representation of acute inflammation. Theor Biol Med Model. 2008 May 27;5:11.
100. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA. 2001 Oct 10;286(14):1754–8.
101. Raith EP, Udy AA, Bailey M, McGloughlin S, MacIsaac C, Bellomo R, et al. Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit. JAMA. 2017 Jan

17;317(3):290–300.
102. Cockrell C, Schobel-McHugh S, Lisboa F, Vodovotz Y, An G. Generating synthetic data with a mechanism-based Critical Illness Digital Twin: Demonstration for Post Traumatic Acute Respiratory Distress Syndrome. BioRxiv. 2022 Nov 24;
103. Bonavia AS, Samuelsen A, Liang M, Hanson J, McKeone D, Chroneos ZC, et al. Comparison of whole blood cytokine immunoassays for rapid, functional immune phenotyping in critically ill patients with sepsis. ICMx. 2023 Oct 13;11(1):70.
104. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. Lancet Respir Med. 2016 Apr;4(4):259–71.
105. Reyes M, Filbin MR, Bhattacharyya RP, Billman K, Eisenhaure T, Hung DT, et al. An immune-cell signature of bacterial sepsis. Nat Med. 2020 Mar;26(3):333–40.
106. Oelen R, de Vries DH, Brugge H, Gordon MG, Vochteloo M, single-cell eQTLGen consortium, et al. Single-cell RNA-sequencing of peripheral blood mononuclear cells reveals widespread, context-specific gene expression regulation upon pathogenic exposure. Nat Commun. 2022 Jun 7;13(1):3267.
107. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646–74.
108. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat Rev Immunol. 2020 Nov;20(11):651–68.
109. Kumar AR, Devan AR, Nair B, Vinod BS, Nath LR. Harnessing the immune system against cancer: current immunotherapy approaches and therapeutic targets. Mol Biol Rep. 2021 Dec;48(12):8075–95.
110. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012 Mar 22;12(4):252–64.
111. Yuan S, Stewart KS, Yang Y, Abdusselamoglu MD, Parigi SM, Feinberg TY, et al. Ras drives malignancy through stem cell crosstalk with the microenvironment. Nature. 2022 Dec;612(7940):555–63.
112. Andre F, Berrada N, Desmedt C. Implication of tumor microenvironment in the resistance to chemotherapy in breast cancer patients. Curr Opin Oncol. 2010 Nov;22(6):547–51.
113. Luchini C, Lawlor RT, Milella M, Scarpa A. Molecular tumor boards in clinical practice. Trends Cancer. 2020 Sep;6(9):738–44.
114. Patel VG, Qin Q, Mellgard G, Parikh AB, Wang B, Alerasool P, et al. Characterizing patterns of disease progression in patients with genitourinary cancers treated with immune checkpoint inhibitors. JCO. 2020 Feb 20;38(6_suppl):482–482.
115. Hue J, Valinciute Z, Thavaraj S, Veschini L. Multifactorial estimation of clinical outcome in HPV-associated oropharyngeal squamous cell carcinoma via automated image analysis of routine diagnostic H&E slides and neural network modelling. Oral Oncol. 2023 Jun;141:106399.
116. Rade M, Kreuz M, Borkowetz A, Sommer U, Blumert C, Füssel S, et al. A reliable transcriptomic risk-score applicable to formalin-fixed paraffin-embedded biopsies improves outcome prediction in localized prostate cancer. Mol Med. 2024 Feb 1;30(1):19.
117. Chuah S, Chew V. High-dimensional immune-profiling in cancer: implications for immunotherapy. J Immunother Cancer. 2020 Feb;8(1).
118. Stahlberg EA, Abdel-Rahman M, Aguilar B, Asadpoure A, Beckman RA, Borkon LL, et al. Exploring approaches for predictive cancer patient digital twins: Opportunities for collaboration and innovation. Front Digit Health. 2022 Oct 6;4:1007784.
119. Meet Your Digital Twin: The Coming Revolution In Drug Development [Internet]. [cited 2024 Feb 25]. Available from: https://www.forbes.com/sites/ganeskesari/2021/09/29/meet-your-digital-twin-the-coming-revolution-in-drug-development/?sh=39e3b552745f
120. An G, Cockrell C. Drug development digital twins for drug discovery, testing and repurposing: A schema for requirements and development. Front Syst Biol. 2022 Jun 20;2.
121. The Powerful Potential of Digital Twin Technology to Improve Drug Discovery, Development,

Manufacturing, and More [Internet]. [cited 2024 Feb 25]. Available from: https://www.pharmasalmanac.com/articles/the-powerful-potential-of-digital-twin-technology-to-improve-drug-discovery-development-manufacturing-and-more
122. Sanofi taps digital twin tech to design its new vaccine plants [Internet]. [cited 2024 Feb 25]. Available from: https://www.fiercepharma.com/manufacturing/sanofi-harnesses-dassaults-digital-twin-technology-optimize-production-future-vaccine
123. Moingeon P, Chenel M, Rousseau C, Voisin E, Guedj M. Virtual patients, digital twins and causal disease models: Paving the ground for in silico clinical trials. Drug Discov Today. 2023 Jul;28(7):103605.
124. Russo G, Sgroi G, Parasiliti Palumbo GA, Pennisi M, Juarez MA, Cardona P-J, et al. Moving forward through the in silico modeling of tuberculosis: a further step with UISS-TB. BMC Bioinformatics. 2020 Dec 14;21(Suppl 17):458.
125. Letter of support for UISS-TB-DR.
126. Viceconti M, Emili L, editors. Toward good simulation practice: best practices for the use of computational modelling and simulation in the regulatory process of biomedical products. Cham: Springer Nature Switzerland; 2024.
127. Laubenbacher R, Adler F, An G, Castiglione F, Eubank S, Fonseca LL, et al. Forum on immune digital twins: a meeting report. NPJ Syst Biol Appl. 2024 Feb 16;10(1):19.