



# Immune digital twins for complex human pathologies: applications, limitations, and challenges



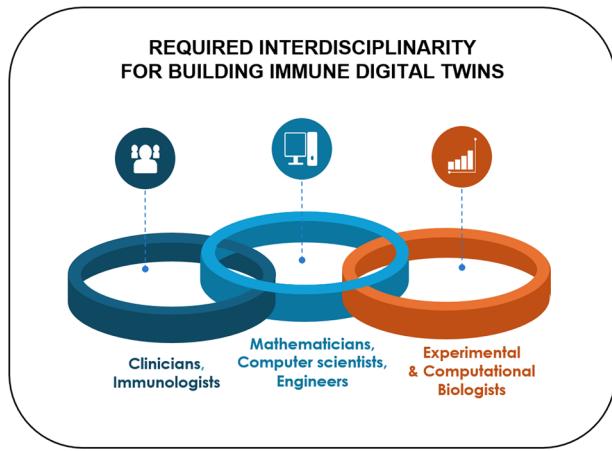
Anna Niarakis <sup>1,2</sup>✉, Reinhard Laubenbacher <sup>3</sup>, Gary An <sup>4</sup>, Yaron Ilan <sup>5</sup>, Jasmin Fisher <sup>6</sup>, Åsmund Flobak <sup>7,8,9</sup>, Kristin Reiche <sup>10,11,12</sup>, María Rodríguez Martínez <sup>13</sup>, Liesbet Geris <sup>14,15,16</sup>, Luiz Ladeira <sup>16</sup>, Lorenzo Veschini <sup>17,18</sup>, Michael L. Blinov <sup>19</sup>, Francesco Messina <sup>20</sup>, Luis L. Fonseca <sup>3</sup>, Sandra Ferreira <sup>21</sup>, Arnau Montagud <sup>22,23</sup>, Vincent Noël <sup>24,25,26</sup>, Malvina Marku <sup>27</sup>, Eirini Tsirvouli <sup>7,28</sup>, Marcella M. Torres <sup>29</sup>, Leonard A. Harris <sup>30,31,32</sup>, T. J. Sego <sup>3</sup>, Chase Cockrell <sup>4</sup>, Amanda E. Shick <sup>33</sup>, Hasan Balci <sup>34</sup>, Albin Salazar <sup>35</sup>, Kinza Rian <sup>36</sup>, Ahmed Abdelmonem Hemedan <sup>37</sup>, Marina Esteban-Medina <sup>36</sup>, Bernard Staumont <sup>16</sup>, Esteban Hernandez-Vargas <sup>38</sup>, Shiny Martis B <sup>39</sup>, Alejandro Madrid-Valiente <sup>22</sup>, Panagiotis Karampeleisis <sup>40</sup>, Luis Sordo Vieira <sup>3</sup>, Pradyumna Harlapur <sup>41</sup>, Alexander Kulesza <sup>39</sup>, Niloofar Nikaein <sup>42,43</sup>, Winston Garira <sup>44,45,46</sup>, Rahuman S. Malik Sheriff <sup>47,48</sup>, Juilee Thakar <sup>49</sup>, Van Du T. Tran <sup>50</sup>, Jose Carbonell-Caballero <sup>22</sup>, Soroush Safaei <sup>51,52</sup>, Alfonso Valencia <sup>22,53</sup>, Andrei Zinovyev <sup>54</sup> & James A. Glazier <sup>18</sup>

Digital twins represent a key technology for precision health. Medical digital twins consist of computational models that represent the health state of individual patients over time, enabling optimal therapeutics and forecasting patient prognosis. Many health conditions involve the immune system, so it is crucial to include its key features when designing medical digital twins. The immune response is complex and varies across diseases and patients, and its modelling requires the collective expertise of the clinical, immunology, and computational modelling communities. This review outlines the initial progress on immune digital twins and the various initiatives to facilitate communication between interdisciplinary communities. We also outline the crucial aspects of an immune digital twin design and the prerequisites for its implementation in the clinic. We propose some initial use cases that could serve as “*proof of concept*” regarding the utility of immune digital technology, focusing on diseases with a very different immune response across spatial and temporal scales (minutes, days, months, years). Lastly, we discuss the use of digital twins in drug discovery and point out emerging challenges that the scientific community needs to collectively overcome to make immune digital twins a reality.

## 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. This capability holds the potential to improve personalised care and patient-tailored treatments. However, implementing such a technology may only be feasible for some pathologies. A recent report by the National Academies of Sciences, Engineering, and Medicine (NASEM) 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<sup>1</sup>. This interpretation aligns with the definition taken by the European Commission in developing their Virtual Human Twin (VHT) initiative and the recommendations in the VHT roadmap<sup>2,3</sup>. 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<sup>4–18</sup>. DTs of large blood vessels could allow the early diagnosis of potential abnormalities and aid in designing interventions<sup>19–21</sup>. Pancreatic DTs, representing an “artificial pancreas”, can largely automate the decision algorithms for insulin



**Fig. 1 | Interdisciplinarity as a key factor in Building Immune Digital Twins.** Illustration of the collaboration among diverse stakeholders to establish an international and interdisciplinary community dedicated to the development and deployment of Immune Digital Twins. Modified template from <https://youexec.com/>.

administration, 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<sup>22–25</sup>.

Several factors have limited the development and adoption of in silico simulations of the immune system to improve patient care directly. We still need a complete understanding of the immune system's health, disease, and therapy response functions. To progress, we must comprehensively leverage what we know and benefit from a wealth of data, tools, and algorithms to augment AI-enabled and mechanism-based simulations<sup>26</sup>.

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 (Fig. 1). Initiatives to bring interdisciplinary and international consortia together are multiplying (*Building Immune Digital Twins; Forum On Precision Immunology: Immune Digital Twins* showcasing the need to bring together stakeholders from industry, pharma, biotech, start-ups, and bio-cluster sectors to form an active community on Building Digital Twins for the human immune system<sup>27</sup>). Bringing these communities together is one of the major challenges we face.

## Basic principles for designing IDTs

An IDT is a digital twin for a particular medical application with a significant immune system component. Besides specific characteristics for the “internal design and content”, the IDT should follow the FAIR principles<sup>28–30</sup> to comply with best practices and community guidelines regarding large-scale and multi-scale models and be:

### Findable

The different IDT elements should be fully annotated and characterised by globally unique and persistent identifiers. They should be stored in appropriate data and model repositories, facilitating their retrieval. Their metadata should also be indexed in a searchable resource.

### 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.

### Interoperable

Interoperability will allow the IDT to process information from multiple heterogeneous sources, ensuring a seamless flow of information. Standard

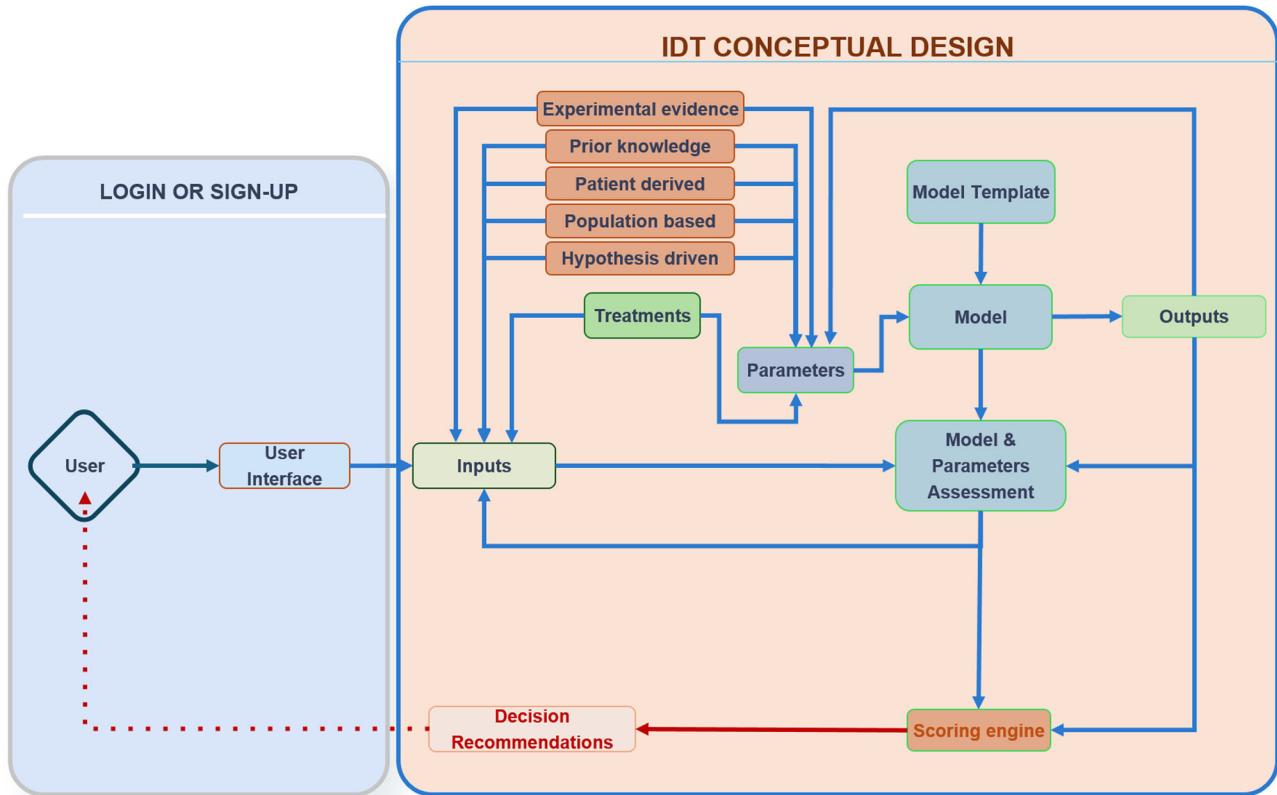
input and output formats and the use of accessible programming languages and environments will help towards its adoption and usability. The IDT design requires a concerted effort by the systems biology community to adopt and implement suggested community standards, such as Systems Biology Markup Language (SBML)<sup>31</sup> for mathematical model exchange, Systems Biology Graphical Notation (SBGN)<sup>32</sup> for model visualisation, Biological Pathway Exchange (BioPAx)<sup>33</sup> for pathway descriptions, and Simulation Experiment Description Markup Language (SED-ML)<sup>34</sup> 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<sup>35</sup>. A helpful resource is also the EDITH standards collection for Virtual Human Twins in Health<sup>36</sup>.

### 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 the description are necessary for maximising reusability. The aspects of scalability and modularity are particular to the IDT field of reusability. The IDT should be:

**Modular.** 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<sup>37</sup>. 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<sup>38</sup> 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)<sup>39</sup> 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 using containers and container libraries in bioinformatics frameworks<sup>40</sup>.

**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, allowing for simpler surrogate models when complexity is unnecessary<sup>41,42</sup> or computing resources are scarce. 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<sup>43</sup>. While supercomputers represent hardware-enhanced machines, HPC uses distributed resources to combine storage, applications, computational power, and network resources. Cloud computing refers to delivering services over the internet to facilitate access to resources and scaling. The combination of HPC and cloud computing could accelerate simulations at a large scale, thus significantly reducing the time to market for an IDT prototype. Furthermore, scalability, modularity, and interoperability should allow researchers to use IDTs at different complexity levels and with different computation resource needs. Alternatively, surrogate models of highly complex IDTs<sup>41,42</sup>,



**Fig. 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 “twinning” process 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.

methods like model order reduction<sup>44</sup>, or the use of precomputed scenarios might be helpful to simulate IDTs in resource-limited settings. In addition, as a best practice in this context, model developers are encouraged to conduct reproducible comparative tests at different levels of scalability. This approach accurately establishes the relationship between the accuracy and quality of the results and the computational resources used in each test. By providing comprehensive metrics, model developers enable end users to determine the minimum computational resources required for their specific use cases while achieving results comparable to those initially reported.

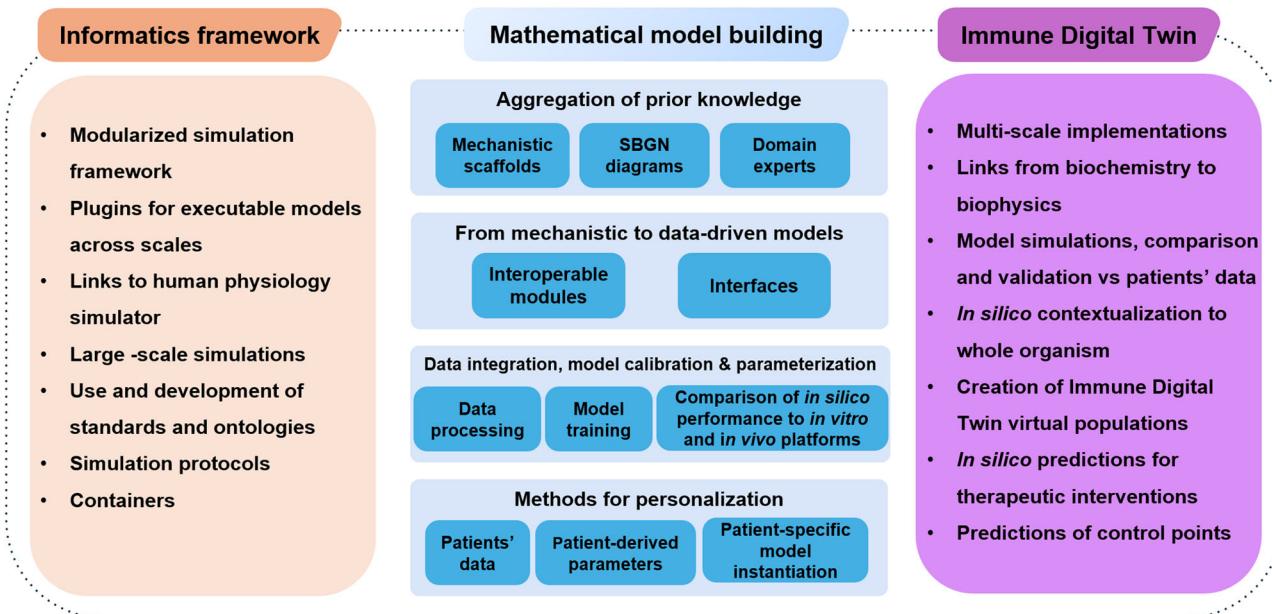
While the IDT should in principle be a two-way information system, ethical questions arise regarding the accessibility of the IDT predictions for patients<sup>45</sup>. In the proposed schema, the decision is not directly accessible by the patient, and it implies the presence of a control point, where 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 various types of rights. In Fig. 2, we offer a conceptual design of an IDT implementation.

### 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<sup>45,46</sup>. 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; (iii) a complex regulatory landscape across countries and institutions, which severely limits

the sharing of sensitive human data; and (iv) the need for interpretability, especially in the context of the GDPR EU law, which highlights the *right to explanation* and the necessity of transparency and accountability in automated decision-making processes. Due to the latter, many algorithms are trained on small and homogeneous cohorts, leading to data overfitting and limited generalisability 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<sup>47</sup>.

Furthermore, models can be contextualised to represent a wide range of demographics and conditions, enhancing the diversity in patient population representation<sup>48</sup>. Foundational predictive AI models, built upon extensive multi-modal data encompassing scientific texts, molecular datasets, and biomedical knowledge graphs, are emerging in biology and show great potential in facilitating all aspects of model engineering, from biocuration to training model parameters<sup>49</sup>. Transfer learning, which involves pretraining models on larger datasets before fine-tuning with specific data, has been successfully implemented in data-scarce applications<sup>50</sup>. Federated learning, a collaborative learning approach that tackles data sharing and privacy issues by keeping data localised while enabling collective model training<sup>51,52</sup>. Lastly, the field of Explainable AI (XAI) is advancing solutions to identify data and algorithmic biases and to derive new insights from black-box models<sup>53</sup>, facilitating more reliable, safer and interpretable predictions of the immune response upon perturbations or treatments. Finally, the current approach to omics data analysis, especially in single-cell data science, is undergoing a significant transformation with the new generation of generative<sup>54</sup> and causal AI methods<sup>55</sup>. These methods effectively bridge the gap between the data-driven approach<sup>56,57</sup> and the mechanistic modelling, marking a notable shift in the traditional distinction between them.



**Fig. 3 |** Different steps across scientific fields for a full-circle IDT implementation. Stepwise and domain specific steps for the implementation of IDTs. The building blocks of each scientific domain can be developed independently. However,

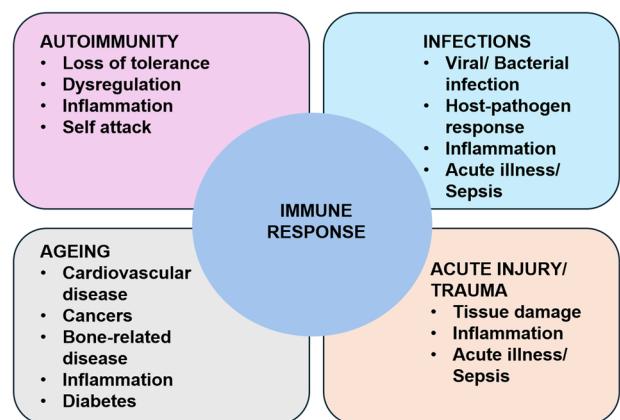
attention to interoperability, use of common standards, and compliance with the FAIR principles will accelerate the building of IDTs that cover most of the technical/methodological needs.

A promising avenue for future development is the integration of AI and multiscale mechanistic models<sup>55</sup>. Mechanistic models excel in inferring causal relationships based on known biological mechanisms<sup>58,59</sup>, while AI models can help identify patterns and correlations within extensive datasets<sup>20</sup>. 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 exemplified in physics with physics-informed neural networks (PINNs), which showed that constraining neural networks with prior knowledge improves accuracy and generalisability even in data-limited scenarios<sup>60</sup>. However, biological systems are typically described in qualitative terms, and how to effectively integrate qualitative prior knowledge with quantitative and mathematical models requires further investigation. Despite the difficulty of integrating qualitative knowledge into deep learning models, proof of concept cases have already been published, such as pathway-aware multi-layered hierarchical networks for cancer patient classification<sup>61</sup> or visible neural networks that can reproduce the inner workings of eukaryotic cells<sup>62</sup>. While further research is essential, hybrid IDTs could be particularly effective in predicting and suggesting therapeutic interventions targeting specific mechanisms<sup>63</sup>.

To summarise, efficiently building Digital Twins of the human immune system requires advancements in informatics and mathematics to accommodate the emergent needs that follow such an endeavour (Fig. 3). To bring IDTs to reality, one would need to leverage the progress in computational biology, AI, and the development of sophisticated integrative methods for low and high-throughput biological data spanning many biological layers to produce a robust ecosystem where data analysis and modelling could be directly linked to patients' data, at both clinical and biological levels. Several building blocks can be developed independently, ensuring that standards, FAIR principles, and interoperability are factored in as we move towards the bigger picture.

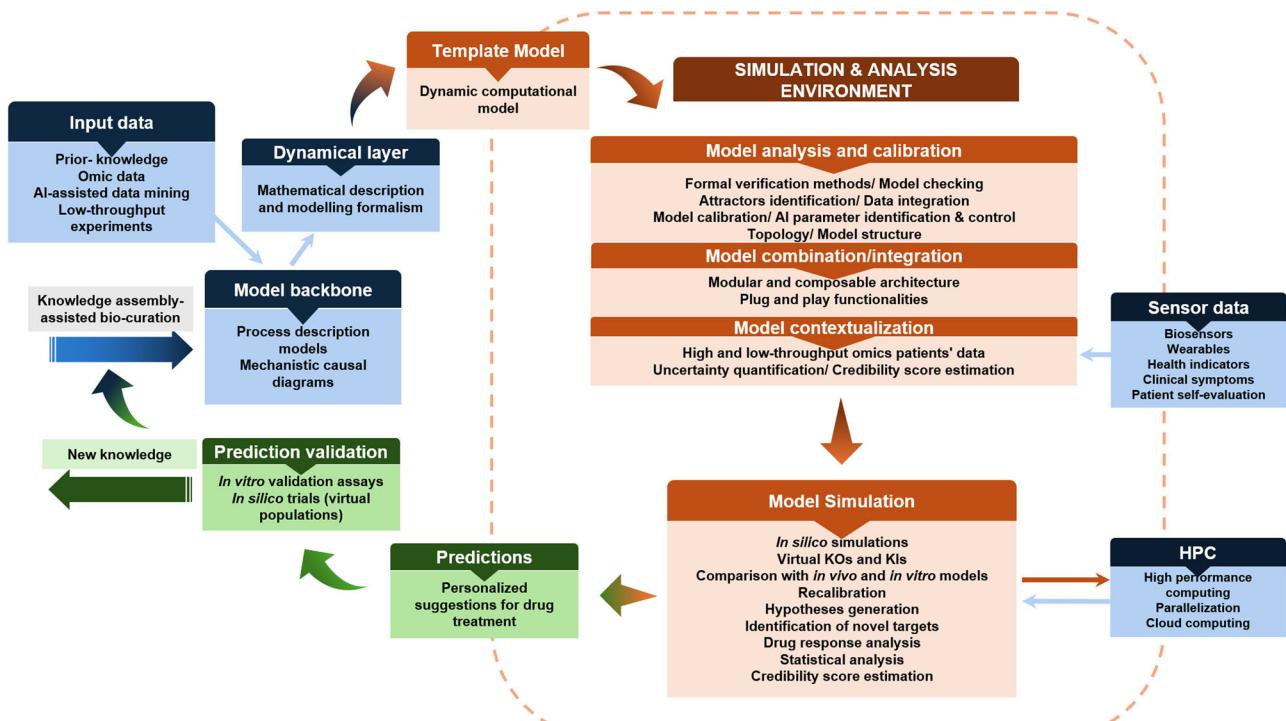
### Implementing Immune Digital Twins in the study of complex human pathologies

As a causal and modulating factor, the human immune system is central to several disease classes, including infectious, autoimmune, and cancer (See Fig. 4).



**Fig. 4 |** Human immune response in various pathologies. Immune/Inflammatory functions and their relationship to various classes of diseases: Autoimmune diseases in which failure in non-self-recognition or negative feedback control of proinflammation leads to persistent inflammation and long-term tissue damage; Infections in which the immune response is responding to various types of microbes (viruses, bacteria and fungi); Ageing, where changes in the function of the immune response can lead to a host of diseases; Acute Illness, where a host of perturbations rapidly activates the immune response. Immune pathophysiological processes range in time scale from hours for acute illness and sepsis to years and decades in autoimmune diseases and cancer. We propose that nearly every disease process and its potential resolution involves to some degree, inflammation and immunity.

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. The ubiquity of the immune system across a host of pathophysiological processes and the multiscale nature of how that pathophysiology manifests point to two classes of challenges when developing the infrastructure for implementing IDTs. The first is related to representing the relevant biology to be incorporated in a specific "fit-for-use" IDT while taking advantage of a shared knowledge base regarding the various components of the immune system. While having a



**Fig. 5 | A bioinformatics ecosystem for data analysis, integration and modelling in IDT implementations.** Stepwise process to obtain a DT implementation that accommodates different types of input data, integrative methods, and modelling formalisms to template model building, simulation, and analysis to create

personalised instantiations for therapeutic interventions. The predictions can be tested using in vitro assays using humanised cellular systems (like organoids) and in silico population trials.

comprehensive computational representation of the immune system that can directly model across different disease processes is an aspirational goal, in initial implementations of IDTs there will be specific choices made regarding what parts of the immune system will be incorporated into a specific IDT; the examples below will provide some insight into how this might be handled. The second challenge is technical: how should such IDTs be implemented? While multiscale modelling technology in biomedicine has made significant progress over the last decade<sup>64</sup>, many problems need to be solved, from software engineering to mathematics, including methodological challenges in sensitivity analysis or uncertainty quantification<sup>65</sup>. Specific use cases will direct the degree of representation and detail required for the IDT. A significant endeavour is identifying the mechanisms of interest and the core application (“fit-for-use”). The basic steps required for a full-circle IDT implementation are shown briefly in Fig. 5, and the examples of four fit-for-use IDTs are discussed afterward. In short, regardless of the pathology at stake, building a Digital Twin needs to address at least two things: pathology-specific events and the immune response to these events. From a more practical point of view, there are several distinct steps that one could follow to reach a mechanistic DT implementation, with adjustments and adaptations where possible to fit the purpose.

The following section presents four paradigms of potential DT implementations for four characteristic pathologies, including intra- and inter-cellular interactions expanding to organ levels and their communication with the immune system. All paradigms are based on existing efforts, published models, and/or disease approaches for DT, extended to account for multiple scales and frameworks. We call them paradigms to stress the goals and approaches, not the specific implementation. These paradigms include infectious pneumonia, rheumatoid arthritis, sepsis, and cancer.

#### Example 1: Infectious pneumonia Immune Digital Twin (IP-IDT) paradigm

**Background.** Pneumonia is among the most common human diseases and the 4th leading cause of death. Infectious pneumonia inflames the air

sacs in one or both lungs, which can quickly become life-threatening. Pathogenic insults such as viruses, fungi, or bacteria can cause the lungs’ air sacs (alveoli) to become inflamed and filled up with fluid or pus. A robust immune response is crucial for the clearance of the pathogen and the resolution of inflammation. However, an overpowering response can lead to acute respiratory distress syndrome (ARDS). Alveolar macrophages are the sentinels of the alveoli; their functions are broad, e.g., phagocytosis, clearing debris, resolution of inflammatory responses, and tissue remodelling<sup>66</sup>. Pulmonary macrophages are diverse, including tissue-resident alveolar macrophages that maintain immune balance and monocyte-derived alveolar macrophages that adapt to the microenvironment<sup>67</sup>. 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 leucocytes to the injury site. If the pneumonia lasts several days, adaptive immunity, such as T-cells and B cells, must be considered<sup>68</sup>.

**Objective.** Current ICU risk calculators accurately predict the likelihood of death but do not provide actionable information about therapeutic interventions in a particular patient. This paradigm aims to develop a computational model that encodes disease-relevant *biological mechanisms* and is *dynamically recalibrated* as new patient data becomes available in the clinic. The pneumonia DT can be used at any given time to simulate different interventions and help the critical care doctor optimise personalised interventions, such as timing and duration of antibiotic treatment based on patient status.

**Implementation suggestion.** The computational model underlying the IP-IDT needs to include mechanisms from the intracellular to the organism scale, as well as features of the pathogen. The epithelial cells lining the alveoli are involved in signalling events and coordinate the early immune response. They also sustain major damage as the infection develops. Immune cells, such as macrophages, dendritic cells, and NK

cells, are recruited and secrete defensive substances. Fungal pneumonia starts in a highly localised fashion in immunocompromised patients, and the infection progresses slowly compared to bacterial pneumonia. To simulate pharmaceutical interventions, the intracellular and tissue scales are particularly important. Other interventions, such as prone positioning of patients in ICU beds or mechanical ventilator assistance, require lung physiology representation. A suitable implementation of an IP-IDT could consist of Ordinary Differential Equations (ODE) or discrete models at the intracellular scale, agent-based models (ABM) at the tissue scale and ODE models at the whole-organ scale.

**Input data.** In an ICU setting, routinely collected data include immune cell counts, cytokine levels, and regular blood draws. One can periodically obtain data from intubated patients through bronchoalveolar lavage characterising the lung environment, such as cell counts, infection severity, and epithelial damage. X-ray images are collected regularly, as well as occasional CT scans. Intracellular data are not likely to be collected, requiring the construction of surrogate models.

**Potential impact.** When fully developed, a pneumonia DT could provide a critical care doctor with a decision-support tool that provides actionable, personalised recommendations for a given patient at a given time.

#### **Example 2: The rheumatoid arthritis Immune Digital Twin paradigm (RA-IDT)**

**Background.** Rheumatoid arthritis is an autoimmune complex disease affecting the human body's articular joints. 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, mistakenly attacking 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<sup>69</sup>. If left untreated, the fulminant stage of the disease is described by a hyperplastic inflamed synovium, cartilage damage, bone erosion, and other systemic consequences<sup>70</sup>. The response rate to current therapies is around 40%<sup>71</sup>, demonstrating the pressing need for accelerating innovative and powerful technologies for personalised care.

**Objective.** Many cells and their cytokines play critical roles in the development of RA. The synovial compartment is infiltrated by leucocytes and the synovial fluid is inundated with pro-inflammatory mediators that are produced to induce an inflammatory cascade, which is characterised by interactions of fibroblast-like synoviocytes with the cells of the innate immune system, including monocytes, macrophages, mast cells, dendritic cells, and so on, as well as cells of adaptive immune system such as T cells and B cells. Endothelial cells contribute to the extensive angiogenesis. The fulminant stage contains hyperplastic synovium, cartilage damage, bone erosion, and systemic consequence. The destruction of the subchondral bone can eventually result in the degeneration of the articular cartilage as the result of a decrease in osteoblasts and an increase in osteoclasts and synoviocytes<sup>70</sup>. The main aim of the RA-IDT would be to digitally represent the interplay between resident cells of the joint and immune cells in RA, which leads to bone erosion, cartilage breakdown, and inflammation, to the level where the RA-IDTs can propose personalised therapy interventions. One major unmet need in the field is the understanding of response to therapy, especially for the non-responders to traditional medication.

**Implementation suggestion.** Hybrid modelling methods that allow for combinations of large-scale inter-cellular models with cell-level agent-based models could create a virtual joint. Recently, disease and cell-specific pathway models<sup>72,73</sup> and dynamic models on the intra and inter-cellular level have been developed<sup>74–79</sup> 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<sup>80,81</sup>. Modelling methods that couple signalling, gene regulation, and metabolic fluxes are now available<sup>74,82</sup> and can be combined with omics data technologies to create personalised instantiations. For the RA-IDT, a combination of large-scale Boolean models that govern cellular behaviour, with agent based models (ABM) for accounting for the interactions of multiple cellular types, could be a suitable implementation. In ABM, agents may receive signals and input from the environment and their neighbouring agents, provide output to the environment and their neighbours, and make 'decisions' based on the input from around them and their internal, sub-cellular decision making rules. An agent may grow, proliferate, enter a quiescent state, express cytokines/chemokines or undergo apoptosis or necrosis in response to surrounding environmental conditions. A first attempt to link Boolean models with ABMs was done during the COVID-19 Disease Map initiative<sup>83</sup>.

**Input data.** Biosensors that can measure matrix degradation and bone erosion along with patient-reported outcomes and scores, could be used to assess the patient's joint and bone health and monitor changes over time<sup>84</sup>. 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.

**Potential impact.** Complex diseases are associated with a heavy societal burden that stems from patients' disabilities and health conditions and the economic costs which come with it. The RA-IDT could help in identifying novel therapeutic combinations tailored to the needs of individual patients, especially beneficial for the ones not responding to traditionally administered medication. In addition, it could provide valuable insights into the disease pathogenesis and help identify new drug targets.

#### **Example 3: The sepsis-IDT paradigm**

**Background.** Sepsis is a syndrome in which a disordered immune response to severe infection or injury leads to early proinflammatory collateral tissue damage/organ dysfunction<sup>85</sup> and later immune incompetence, leading to increased susceptibility to nosocomial infections<sup>86</sup>. As sepsis is a systemic disease, a sepsis-IDT will necessarily include the organs at risk: the immune system, lung, liver, kidney, gut, and cardiovascular system.

**Objective.** The primary goal of the Sepsis-IDT is to provide the capability to treat sepsis by multimodal adaptive modulation of a patient's underlying cytokine milieu ("fit-for-purpose" as per the NASEM report). It will account for sepsis's heterogeneity and dynamic complexity by having an ongoing data link between the virtual and real twin and informing control/guiding therapy<sup>1</sup>.

**Implementation suggestion.** The clinical time scale of sepsis (hours to weeks) focuses on the acute aspects of the innate immune response and its initial transition to adaptive immunity. An example of an existing computational model for this can be seen in ref. 87. Furthermore, sepsis manifests in dysfunction of multiple organs, and thus, each organ system can be cast as a module of the system-level Sepsis-IDT. An early example of this approach simulates the gut-lung axis in sepsis. It consists of modular agent-based models of tissue/organ-specific epithelial cells interacting with and connected to an agent-based model of endothelial cells and circulating immune cells.<sup>88</sup> These existing examples of dynamic multiscale molecule-to-organ integration form the basis of an initial Critical Illness Digital Twin (CIDT)<sup>89,90</sup>. This CIDT would be used to train (off-line) an artificial intelligence controller (AIC) offline, as described in refs. 90,91. This digital twin-trained AIC would be the "brain" of an

integrated cyber-physical system that monitors real-time plasma cytokine/mediator levels (using existing technologies as described in ref. 92 and uses the AIC to guide the administration of different amounts of pro- and anti-inflammatory mediators/monoclonal antibodies to steer an individual patient back to a state of health.

**Input data.** The ongoing data link between the virtual and real world, per the NASEM DT definition, can be accomplished for the Sepsis-IDT via a suite of clinical and laboratory measurements providing dynamic molecular profiling of the patient's immune state as reflected by circulating inflammatory mediators<sup>92</sup> and determinants of trajectories of organ function. For the latter, the Sequential Organ Failure Score (SOFA)<sup>93</sup> and its variants<sup>94</sup> can be generated by readily obtainable clinical measurements and sequentially measured to update the IDT. The current version of the CIDT is poised for testing in a sufficiently complex animal model of sepsis, but should the performance of this cyber-physical system be insufficient, the development loop can both refine the Sepsis-IDT and guide sensor/assay developments as per guidance from the NASEM report.

**Potential impact.** Given that the current inability to effectively modulate the inflammatory/immune dynamics in sepsis is due in significant part to the heterogeneity and dynamic complexity of the disease process<sup>95</sup>, the personalised and adaptive control capabilities offered by the digital twin paradigm point to the potential therapeutic benefits of a successfully implemented Sepsis-IDT.

#### Example 4: The onco-IDT paradigm

**Background.** The ability to evade immune surveillance and destruction is a hallmark of cancer<sup>96</sup>. Immunotherapy approaches have significantly improved cancer outcomes for several cancer types<sup>97</sup>. During oncogenesis, the immune system activates a multifaceted response involving innate and adaptive immune systems. However, cancer cells and their environment progressively evade immune surveillance by down-regulating major histocompatibility complexes (MHC) and upregulating immune checkpoint proteins<sup>98,99</sup>. Cancerous, immune, and stromal cells are the critical components of the tumour microenvironment (TME), and their interplay represents a complex system. The US National Cancer Institute and the US Department of Energy have started to explore the development and implementation of predictive Cancer Patient Digital Twins for personalised treatment<sup>100</sup>.

**Objective.** The main aim of Onco-IDTs would be to provide a computational representation of the patient's cancer disease, including the interactions between cancer cells, tumour-associated cells, and the immune system. An ideal Onco-IDT should have the ability to be calibrated to patient's data and used for clinical decision support. An Onco-IDT should include elements like the TME, neo-angiogenesis, pre-metastatic and metastatic conditions, and system-level information like blood and lymphatic transport. Today, personalised therapy in oncology is making progress in identifying cancer-driver mutations for each patient<sup>101</sup> or cellular patterns linked with disease state/progression<sup>102–104</sup>. Successful implementations of *in silico* mechanistic models led to the identification of optimal treatments with minimal toxicity in melanoma and breast cancer<sup>105,106</sup>. Recently, patient-specific Boolean models of signalling networks were used to guide personalised treatments<sup>107</sup>.

**Implementation suggestion.** A possible implementation of an Onco-IDT can draw inspiration from the multiscale model of the different modes of cancer cell invasion described in ref. 108. The framework includes agent-based modelling and continuous time Markov processes applied on Boolean network models. The model is focused on cell migration considering not only spatial information obtained from the agent-based simulation but also intracellular regulation obtained from the Boolean model. It could be expanded to account for more complex

phenotypes and interactions of tumour cells with cells from TME and immune cells, and also adapted to include clinical features and patient-derived characteristics. An Onco-IDT could be employed to optimise treatments and treatment combinations and potentially predict response to a particular treatment for individual patients.

**Input data.** Extensive cell phenotyping, genetic testing, and 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<sup>109</sup>. However, cancer genomics-guided approaches harnessing or targeting the immune system are still in their infancy, and surrogate markers are used. Also, biopsies are invasive and usually reserved for diagnostics, limiting their ability to track tumour development over time and sample intra- and inter-tumoral heterogeneity. Data and measurements from non or semi-invasive interventions, such as electronic health records, radiological imaging, and serological (circulating tumour cells and circulating tumour DNA) or molecular data that could inform about the patient's inflammatory state, can be integrated with the biopsy data points. In addition, blood samples can provide insights into relevant drug pharmacokinetic (PK) processes, while wearables can offer additional information regarding vital signs, body temperature, and physical activity levels.

**Potential impact.** 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 to prognostic predictions regarding disease course, considering factors like metastatic capacity and patient survival. Additionally, it could offer actionable insights concerning therapeutic interventions, involving selecting 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.

These paradigms outline four scenarios where digital twins could be deployed for treatment optimisation. The sepsis and pneumonia paradigms take place on very short time scales, days or weeks, whereas the oncology and autoimmune paradigms focus on time scales of months or years. Yet, the implementation of IDTs represents a key enabling technology for the personalisation and optimisation of treatment. While the pathological characteristics are very different, the technical modelling challenges are similar. Besides data challenges, from obtaining appropriate data to interpreting them and turning them into model parameterisations, the modelling scaffolds and implementations could follow shared reasoning. For example, for molecular level simulations: one could use networks, Boolean models, ODE-based models, discrete logic-based models; for cellular level: networks, Boolean models, ODE-based models, discrete logic-based models, multiscale models, hybrid models; for cell-cell communication and tissular level: Agent-based models, multicellular models, Boolean models, ODE-based models, discrete logic-based models; for organ level: structural, biophysical, biomechanical models, ODE models, and for system level: multiscale models, biophysical models, hybrid models, human physiology engines. In Table 1 we list available software and platforms for different modelling types and scales.

#### IDTs in drug discovery

Drug development is costly and slow. The costs include early research and discovery costs through clinical development, regulatory approval, and post-marketing surveillance. Most 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<sup>110</sup>.

**Table 1 | Computational modelling software and simulation platforms that can be employed to build models of different scales**

Computational modelling		
Scale	Tool	Type
Single scale (Molecular/ Cellular)	Aeon	Boolean
	Bio Model Analyzer	Logic-based, multivalued
	Biocham	Boolean
	BioNetGen	Rule-based/ODE/ABM
	BooleSim	Boolean
	BoolNetR	Boolean
	Cell Collective	Boolean
	CellDesigner	ODE
	CellNetAnalyzer	Boolean
	CellNOpt	Boolean
	COMSOL Multiphysics	PDE
	COPASI	ODE/SDE/Gillespie/steady-state solvers
	GINsim	Boolean - genetic regulatory networks
	JWS online	ODE
	Kappa	Rule-based/ABM
	MaBoSS	Boolean, stochastic
	MATLAB toolkit	ODE
	M-cell	ABM cellular spatial
	NERDSS	ABM spatial
	PhysiCell	ABM cellular spatial
	PyBoolNet	Boolean
	Smoldyn	ABM spatial
	SpringSaLaD	ABM spatial
	Tellurium	ODE
	Virtual Cell (VCell)	ODE/PDE/SSA/ABM
	WebMaBoSS	Boolean models
Metabolic	CellNetAnalyzer	Flux Balance Analysis
	COBREXA	Flux Balance and Flux Variability Analysis
	Escher	Flux Balance Analysis
	KBase	Flux Balance Analysis
	Metabolizer	Flux Balance Analysis
	MetaLo	Metabolic analysis of logical models extracted from maps
	RAVEN	Genome scale metabolic modelling
Multiscale	Chaste	Molecular + Cellular, supports CellML
	CompuCell3D	Molecular + Cellular
	Morpheus	Molecular + Cellular
	NetLogo	Molecular + Cellular
	PhysiBoSS	Molecular + Cellular
	Tissue Forge	Molecular + Cellular
	Vivarium	Merging multiple scales
System level	BioGears	Human Physiology
	Physiome Project	Virtual Physiological Human
	Pulse	Human Physiology

**Table 1 (continued) | Computational modelling software and simulation platforms that can be employed to build models of different scales**

Computational modelling		
Scale	Tool	Type
	Universal Immune System Simulator	Generic Immune System simulator

Additional useful resources, such as pathway editors, databases, visualisation software, simulation environments and repositories, as well as ML/AI tools, that can be used in various steps of the IDT building can be found in Supplementary Data 1.

The introduction of a specialised form of drug development digital twins (DDDT) has the potential to be a game-changer<sup>111</sup>. Moreover, we can envision further specialisation of DDDT's 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 DDDT's aimed at optimising the process of drug production, with notable examples like Sanofi exploring the use of digital twins for vaccine manufacturing<sup>112</sup>.

All these DT specialisations require specific designs, functionalities, and connections to the existing wealth of public and proprietary data. Furthermore, virtual populations of patient DTs can be used to run in silico clinical trials that can accompany or be used to design real-life trials<sup>113</sup>. One recent example is the Universal Immune System Simulator (UISS)<sup>114</sup>. The European Medicines Agency (EMA) provided a letter of support for the use of the UISS as a simulation platform to predict how the circulating interferon-gamma (IFN $\gamma$ ) 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 several diseases<sup>115</sup>. 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 using computational modelling and simulation in the regulatory process of biomedical products, showcasing the need to address policy and implementation early on in the DT design<sup>116</sup>.

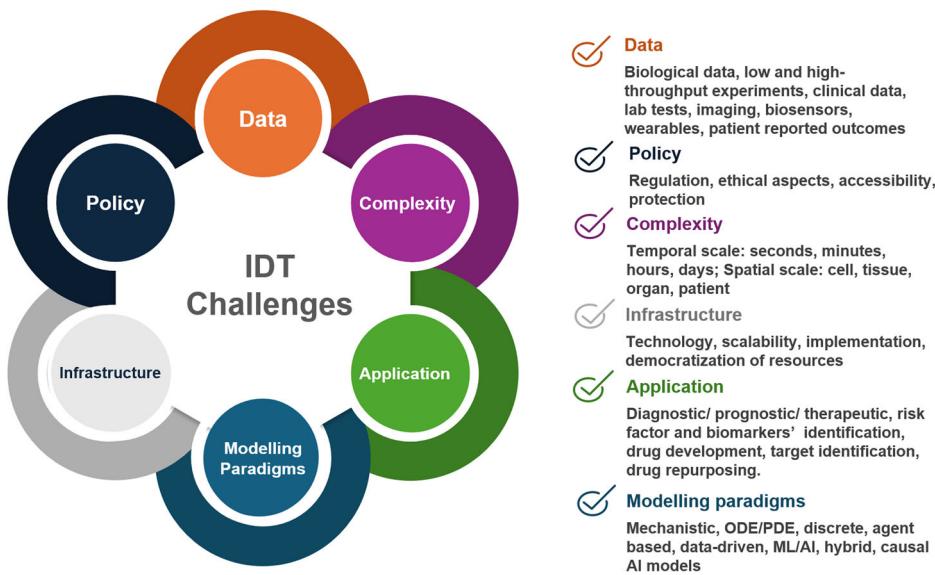
## Regulatory and ethical aspects of IDTs

Regulatory stakeholders like the Food and Drug Administration (FDA), European Medicines Agency (EMA), and other national agencies have expressed their acceptance of simulation-derived results, as evidenced in the submission dossiers. FDA started the Model-Informed Drug Development (MIDD) pilot program about a decade ago<sup>117</sup>. Owing to the success it found in advancing medicinal product development, the FDA has now established MIDD meeting formats for the fiscal years 2023-2027, welcoming modelling and simulation applications at various stages of product development, especially in the domain of rare/orphan and paediatric indications<sup>118,119</sup>. EMA has also evaluated the modelling and simulation approaches earlier through the Modeling and Working Party to increase awareness across European national authorities.

Moreover, National funding agencies in Europe and the US encourage the use of innovative digital technologies as alternative methods for animal experiments.

**Fig. 6 | Key challenges in developing and implementing IDTs in pre-clinical and clinical settings.** The implementation of IDTs require a community-driven approach to tackle challenges in data acquisition, analysis and integration, policy and data protection, methodological aspects to address complexity, dedicated infrastructure development. Tailor-made solutions are also needed to address specific unmet needs in different fields of application, and, lastly, robust and credible scalable modelling approaches for complex human pathologies that involve characteristic immune responses.

Modified template from <https://youexec.com/>.



If IDTs are to form an intrinsic part of the medical decision processes, they must be classified as software devices for medical use. Accordingly, they must undergo a software development process that complies with the IEC 62304 standard, among others<sup>120</sup>. The FDA has also published guidelines regarding the credibility assessment of computational modelling and simulations in medical device submissions<sup>121</sup>. For a smooth transition from research-use-only software to medical software devices, meeting these requirements as early as possible in the development process and at a reasonable cost will be essential. Following established guidelines from the systems biology community<sup>122–125</sup> for model specification, documentation, data file formats, etc., is the critical first step in ensuring compliance with regulatory requirements.

Digital twin technology and the AI era also raise ethical questions regarding the protection of personal data, their use, and their legacy, as well as the democratisation of such technology for the benefit of humanity<sup>126</sup>. The success of these technologies in health care depends on addressing three main ethical challenges: the right and exhaustive training of AI models; healthcare data management and respecting the privacy of patients; and encouraging patient trust in clinicians who use AI-based tools<sup>127</sup>.

Questions regarding the accessibility of the template and personalised models, the use of clinical data to train AI algorithms that could be commercialised, employers' and insurance companies' access to digital twins of employees, and the fate of a digital twin after the passing of the actual patient are open to discussion and debate<sup>128</sup>.

## Limitations and challenges

Digital twin technology for healthcare is advancing at a fast pace. The numerous scientific articles published in the last two years have tried pinpointing the multiple challenges the scientific community faces to bring DTs into the preclinical and clinical settings<sup>4–18</sup>. One of the most common issues is the reconciliation of the different scales in biology. While some fields have been progressing fast, especially in the cardiology domain<sup>20</sup>, the twin is built on the organ level (i.e. modelling the heart), focusing primarily on biophysical, structural, and biomechanical characteristics derived from imaging sources, electrocardiogram (ECG) databases and, in some cases, clinical data, including data from biosensors and wearable technology<sup>129,130</sup>. However, pharmaceutical treatments operate on the molecular and cellular level, and the absence of causal molecular interaction networks underlying the biophysical processes creates a critical gap. While not all DTs should span all layers and scales, building the technology and creating reproducible, scalable, and interoperable frameworks to link these scales when needed is a necessary step for moving forward.

Similarly, digital twins on the cellular or tissue scale often fail to scale up and provide links to full-body manifestations and clinical measurements as they focus on a few molecular or cellular processes. Conclusions extrapolated to an organ or patient level are often made based on a limited number of biomarkers or phenotypes. In this case, these sophisticated computational models are complex to implement in the clinical or pre-clinical setting, as they operate on a different level regarding routine clinical measurements and patient assessment.

Building Digital Twins of the human immune system comes with some additional hardship. The immune system is inherently complex and operates on multiple scales, including organs and cells. Moreover, it is difficult to establish a “baseline” modus operandi that fits a general population. Focusing on specific pathological conditions, which have distinct localised and systemic manifestations of the dysregulation, and where the interplay between resident and immune cells is more straightforward to measure and quantify, might be the best approach for the first “proof of concept” IDT implementations. Moreover, as inflammation seems omnipresent in most disease settings, it could be seen as the core immune response that could be built and modelled in an adaptable way to fit most cases. Figure 6 summarises the challenges associated with DT development and implementation, especially those related to data, policy, complexity, infrastructure, application, and modelling paradigms.

## Perspectives

A change of mindset is needed to achieve tangible results in IDT technology. Traditionally, clinicians, immunologists, and experimental biologists identify hallmarks of the disease, disease biomarkers, and affected pathways, organs, and systemic manifestations. They also include in the study measurable factors used in the clinic and employ available experimental techniques to enrich the molecular, genomic, metabolic, and clinical profile of the patients. Bioinformaticians and computational biologists then 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 bioinformaticians and modellers in the study design could ensure 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 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.

Mechanism-based simulations, AI-enabled data integration, and subsequent experimental and clinical validation will allow for the iterative improvement of human immune system models. 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.

Received: 27 May 2024; Accepted: 27 September 2024;

Published online: 30 November 2024

## References

- 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* (National Academies Press, 2023).
- Consortium E. Edith C. S. A. Deliverable 3.2: first draft of the VHT roadmap. *Zenodo* (2023).
- 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.* **28**, 491–501 (2023).
- Sahal, R., Alsamhi, S. H. & Brown, K. N. Personal digital twin: a close look into the present and a step towards the future of personalised healthcare industry. *Sensors* **22**, 5918 (2022).
- Björnsson, B. et al. Digital twins to personalize medicine. *Genome Med.* **12**, 4 (2019).
- Chang, H.-C. et al. Understanding the need for digital twins' data in patient advocacy and forecasting oncology. *Front. Artif. Intell.* **6**, 1260361 (2023).
- Chaudhuri, A. et al. Predictive digital twin for optimizing patient-specific radiotherapy regimens under uncertainty in high-grade gliomas. *Front. Artif. Intell.* **6**, 1222612 (2023).
- Heudel, P. E., Renard, F. & Attye, A. [Digital twins in cancer research and treatment: a future for personalized medicine]. *Bull. Cancer* **110**, 1085–1087 (2023).
- Iqbal, J. D., Krauthammer, M. & Biller-Andorno, N. The use and ethics of digital twins in medicine. *J. Law Med. Ethics* **50**, 583–596 (2022).
- Rahmim, A. et al. Theranostic digital twins for personalized radiopharmaceutical therapies: reimagining theranostics via computational nuclear oncology. *Front. Oncol.* **12**, 1062592 (2022).
- Pesapane, F., Rotili, A., Penco, S., Nicosia, L. & Cassano, E. Digital twins in radiology. *J. Clin. Med.* **11**, 6553 (2022).
- Jung, A., Gsell, M. A. F., Augustin, C. M. & Plank, G. An integrated workflow for building digital twins of cardiac electromechanics—a multi-fidelity approach for personalising active mechanics. *Mathematics (Basel)* **10**, 823 (2022).
- Corral-Aceró, J. et al. The “Digital Twin” to enable the vision of precision cardiology. *Eur. Heart J.* **41**, 4556–4564 (2020).
- Abdollahi, H. et al. Radiopharmaceutical therapy on-a-chip: a perspective on microfluidic-driven digital twins towards personalized cancer therapies. *Sci. Bull. (Beijing)* **68**, 1983–1988 (2023).
- Trayanova, N. A. & Prakosa, A. Up digital and personal: how heart digital twins can transform heart patient care. *Heart Rhythm* **21**, 89–99 (2023).
- Zhang, Y. et al. Predicting ventricular tachycardia circuits in patients with arrhythmogenic right ventricular cardiomyopathy using genotype-specific heart digital twins. *medRxiv* <https://doi.org/10.1101/2023.05.31.23290587> (2023).
- Winter, P. D. & Chico, T. J. A. 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* **23**, 6333 (2023).
- Clarke, M. A. & Fisher, J. Executable cancer models: successes and challenges. *Nat. Rev. Cancer* **20**, 343–354 (2020).
- Chakshu, N. K., Sazonov, I. & Nithiarasu, P. Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis. *Biomech. Model. Mechanobiol.* **20**, 449–465 (2021).
- Coorey, G. et al. The health digital twin to tackle cardiovascular disease—a review of an emerging interdisciplinary field. *npj Digital Med.* **5**, 126 (2022).
- Gillette, K. et al. A Framework for the generation of digital twins of cardiac electrophysiology from clinical 12-leads ECGs. *Med. Image Anal.* **71**, 102080 (2021).
- Shamanna, P. et al. Type 2 diabetes reversal with digital twin technology-enabled precision nutrition and staging of reversal: a retrospective cohort study. *Clin. Diabetes Endocrinol.* **7**, 21 (2021).
- Kovatchev, B. The artificial pancreas in 2017: the year of transition from research to clinical practice. *Nat. Rev. Endocrinol.* **14**, 74–76 (2018).
- Brown, S. A. et al. Six-month randomized, multicenter trial of closed-loop control in type 1 diabetes. *N. Engl. J. Med.* **381**, 1707–1717 (2019).
- Armeni, P. et al. Digital twins in healthcare: is it the beginning of a new era of evidence-based medicine? a critical review. *J. Pers. Med.* **12**, 1255 (2022).
- Handel, A., La Gruta, N. L. & Thomas, P. G. Simulation modelling for immunologists. *Nat. Rev. Immunol.* **20**, 186–195 (2020).
- Laubenbacher, R. et al. Building digital twins of the human immune system: toward a roadmap. *npj Digital Med.* **5**, 64 (2022).
- Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
- Wilkinson, M. D. et al. Addendum: the FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **6**, 6 (2019).
- Rocca-Serra, P. et al. The FAIR Cookbook—the essential resource for and by FAIR doers. *Sci. Data* **10**, 292 (2023).
- Keating, S. M. et al. SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol. Syst. Biol.* **16**, e9110 (2020).
- Le Novère, N. et al. The systems biology graphical notation. *Nat. Biotechnol.* **27**, 735–741 (2009).
- Demir, E. et al. The BioPAX community standard for pathway data sharing. *Nat. Biotechnol.* **28**, 935–942 (2010).
- Waltemath, D. et al. Reproducible computational biology experiments with SED-ML—the simulation experiment description markup language. *BMC Syst. Biol.* **5**, 198 (2011).
- Waltemath, D. et al. The first 10 years of the international coordination network for standards in systems and synthetic biology (COMBINE). *J. Integr. Bioinform* **17**, 20200005 (2020).
- FAIRsharing [Internet]. Available from <https://fairsharing.org/4787>.
- Masison, J. et al. A modular computational framework for medical digital twins. *Proc. Natl. Acad. Sci. USA* **118**, e2024287118 (2021).
- Neal, M. L. et al. A reappraisal of how to build modular, reusable models of biological systems. *PLoS Comput. Biol.* **10**, e1003849 (2014).
- Smith, L. P. et al. SBML level 3 package: hierarchical model composition, version 1 release 3. *J. Integr. Bioinform.* **12**, 268 (2015).
- Gruening, B. et al. Recommendations for the packaging and containerizing of bioinformatics software. [version 2; peer review: 2 approved, 1 approved with reservations]. *F1000Res* **7**, ISCB Comm J-742 (2018).
- Garira, W. The research and development process for multiscale models of infectious disease systems. *PLoS Comput. Biol.* **16**, e1007734 (2020).
- Garira, W. A complete categorization of multiscale models of infectious disease systems. *J. Biol. Dyn.* **11**, 378–435 (2017).

43. Montagud, A., Ponce de León, M. & Valencia, A. Systems biology at the giga-scale: large multi-scale models of complex, heterogeneous multicellular systems. *Curr. Opin. Syst. Biol.* **28**, 100385 (2021).

44. Kumar, R. & Sikander A. *Review and Analysis of Model Order Reduction Techniques for High-dimensional Complex Systems* (Microsystem Technologies, 2024).

45. Wagner, J. et al. A single-cell atlas of the tumor and immune ecosystem of human breast. *Cancer Cell* **177**, 1330–1345.e18 (2019).

46. Manica, M. et al. Toward explainable anticancer compound sensitivity prediction via multimodal attention-based convolutional encoders. *Mol. Pharm.* **16**, 4797–4806 (2019).

47. An, G. & Cockrell, C. Generating synthetic multidimensional molecular time series data for machine learning: considerations. *Front Syst. Biol.* **3** (2023).

48. Zhang, A., Xing, L., Zou, J. & Wu, J. C. Shifting machine learning for healthcare from development to deployment and from models to data. *Nat. Biomed. Eng.* **6**, 1330–1345 (2022).

49. Theodoris, C. V. et al. Transfer learning enables predictions in network biology. *Nature* **618**, 616–624 (2023).

50. Weber, A., Born, J. & Rodriguez Martínez, M. TITAN: T-cell receptor specificity prediction with bimodal attention networks. *Bioinformatics* **37**, i237–i244 (2021).

51. Rieke, N. et al. The future of digital health with federated learning. *npj Digital Med.* **3**, 119 (2020).

52. Dayan, I. et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **27**, 1735–1743 (2021).

53. Novakovsky, G., Dexter, N., Libbrecht, M. W., Wasserman, W. W. & Mostafavi, S. Obtaining genetics insights from deep learning via explainable artificial intelligence. *Nat. Rev. Genet.* **24**, 125–137 (2023).

54. Bordukova, M., Makarov, N., Rodriguez-Esteban, R., Schmich, F. & Menden, M. P. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. *Expert Opin. Drug Discov.* **19**, 33–42 (2024).

55. Tejada-Lapuerta, A. et al. Causal machine learning for single-cell genomics. Preprint at <https://arxiv.org/abs/2310.14935> (2023).

56. Dong, M. et al. Causal identification of single-cell experimental perturbation effects with CINEMA-OT. *Nat. Methods* **20**, 1769–1779 (2023).

57. Qiu, X. et al. Mapping transcriptomic vector fields of single cells. *Cell* **185**, 690–711.e45 (2022).

58. Garira, W. & Muzhini, K. Application of the replication-transmission relativity theory in the development of multiscale models of infectious disease dynamics. *J. Biol. Dyn.* **17**, 2255066 (2023).

59. Garira, W. & Muzhini, K. The universal theory for multiscale modelling of infectious disease dynamics. *Mathematics* **11**, 3874 (2023).

60. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).

61. Elmarakeby, H. A. et al. Biologically informed deep neural network for prostate cancer discovery. *Nature* **598**, 348–352 (2021).

62. Ma, J. et al. Using deep learning to model the hierarchical structure and function of a cell. *Nat. Methods* **15**, 290–298 (2018).

63. Martínez, M. R., Barberis, M. & Niarakis, A. Computational modelling of immunological mechanisms: from statistical approaches to interpretable machine learning. *Immunol Informatics* **12**, 100029 (2023).

64. Kirschner, D. E., Hunt, C. A., Marino, S., Fallahi-Sichani, M. & Linderman, J. J. Tuneable resolution as a systems biology approach for multi-scale, multi-compartment computational models. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **6**, 289–309 (2014).

65. Marino, S., Hogue, I. B., Ray, C. J. & Kirschner, D. E. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* **254**, 178–196 (2008).

66. Rodriguez-Rodriguez, L., Gillet, L. & Machiels, B. Shaping of the alveolar landscape by respiratory infections and long-term consequences for lung immunity. *Front Immunol.* **14**, 1149015 (2023).

67. Hou, F., Xiao, K., Tang, L. & Xie, L. Diversity of macrophages in lung homeostasis and diseases. *Front Immunol.* **12**, 753940 (2021).

68. Kumar, V. Pulmonary innate immune response determines the outcome of inflammation during pneumonia and sepsis-associated acute lung injury. *Front Immunol.* **4**, 1722 (2020).

69. Firestein, G. S. & McInnes, I. B. Immunopathogenesis of rheumatoid arthritis. *Immunity* **46**, 183–196 (2017).

70. Guo, Q. et al. Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. *Bone Res* **6**, 15 (2018).

71. Humby, 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* **397**, 305–317 (2021).

72. Zerrouk, N., Aghakhani, S., Singh, V., Augé, F. & Niarakis, A. A mechanistic cellular atlas of the rheumatic joint. *Front. Syst. Biol.* **2** (2022).

73. Singh, V. et al. RA-map: building a state-of-the-art interactive knowledge base for rheumatoid arthritis. *Database (Oxford)* **2020**, baaa017 (2020).

74. Aghakhani, S., Soliman, S. & Niarakis, A. Metabolic reprogramming in rheumatoid arthritis synovial fibroblasts: a hybrid modeling approach. *PLoS Comput. Biol.* **18**, e1010408 (2022).

75. Zerrouk, N., Alcraft, R., Hall, B. A., Augé, F. & Niarakis, A. Large-scale computational modelling of the M1 and M2 synovial macrophages in rheumatoid arthritis. *NPJ Syst. Biol. Appl.* **10**, 10 (2024).

76. Miagoux, Q. 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.* **11**, 785 (2021).

77. 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.* **9**, 33 (2023).

78. Macfarlane, F. R., Chaplain, M. A. J. & Eftimie, R. Quantitative predictive modelling approaches to understanding rheumatoid arthritis: a brief review. *Cells* **9**, 74 (2019).

79. Moise, N. & Friedman, A. Rheumatoid arthritis—a mathematical model. *J. Theor. Biol.* **461**, 17–33 (2019).

80. Kerkhofs, J. et al. A qualitative model of the differentiation network in chondrocyte maturation: a holistic view of chondrocyte hypertrophy. *PLoS ONE* **11**, e0162052 (2016).

81. Lesage, R. et al. An integrated in silico-in vitro approach for identifying therapeutic targets against osteoarthritis. *BMC Biol.* **20**, 253 (2022).

82. Aghakhani, S., Niarakis, A. & Soliman, S. MetaLo: metabolic analysis of Logical models extracted from molecular interaction maps. *J. Integr. Bioinform.* **21**, 20230048 (2024).

83. Niarakis, A. et al. Drug-target identification in COVID-19 disease mechanisms using computational systems biology approaches. *Front Immunol.* **14**, 1282859 (2023).

84. Creagh, A. P. et al. Digital health technologies and machine learning augment patient reported outcomes to remotely characterise rheumatoid arthritis. *npj Digital Med.* **7**, 33 (2024).

85. Seymour, C. W. et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* **315**, 762–774 (2016).

86. Schefold, J. C., Hasper, D., Volk, H. D. & Reinke, P. Sepsis: time has come to focus on the later stages. *Med. Hypotheses* **71**, 203–208 (2008).

87. An, G. In silico experiments of existing and hypothetical cytokine-directed clinical trials using agent-based modeling. *Crit. Care Med.* **32**, 2050–2060 (2004).

88. An, G. Introduction of an agent-based multi-scale modular architecture for dynamic knowledge representation of acute inflammation. *Theor. Biol. Med Model* **5**, 11 (2008).

89. 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* <https://doi.org/10.1101/2022.11.22.517524> (2022).

90. An, G. & Cockrell C. A design specification for Critical Illness Digital Twins to cure sepsis: responding to the National Academies of Sciences, Engineering and Medicine Report: Foundational Research Gaps and Future Directions for Digital Twins. Preprint at <https://arxiv.org/abs/2405.05301> (2024).

91. 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.* **13**, 995395 (2022).

92. Aldo, P., Marusov, G., Svacara, D., David, J. & Mor, G. Simple Plex<sup>TM</sup>: a novel multi-analyte, automated microfluidic immunoassay platform for the detection of human and mouse cytokines and chemokines. *Am. J. Reprod. Immunol.* **75**, 678–693 (2016).

93. Ferreira, F. L., Bota, D. P., Bross, A., Mélot, C. & Vincent, J. L. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA* **286**, 1754–1758 (2001).

94. Raith, E. P. 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* **317**, 290–300 (2017).

95. Cockrell, C. & An, G. Sepsis reconsidered: Identifying novel metrics for behavioral landscape characterization with a high-performance computing implementation of an agent-based model. *J. Theor. Biol.* **430**, 157–168 (2017).

96. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).

97. Waldman, A. D., Fritz, J. M. & Lenardo, M. J. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat. Rev. Immunol.* **20**, 651–668 (2020).

98. Kumar, A. R., Devan, A. R., Nair, B., Vinod, B. S. & Nath, L. R. Harnessing the immune system against cancer: current immunotherapy approaches and therapeutic targets. *Mol. Biol. Rep.* **48**, 8075–8095 (2021).

99. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–264 (2012).

100. Stahlberg, E. A. et al. Exploring approaches for predictive cancer patient digital twins: opportunities for collaboration and innovation. *Front Digit Health* **4**, 1007784 (2022).

101. Luchini, C., Lawlor, R. T., Milella, M. & Scarpa, A. Molecular tumor boards in clinical practice. *Trends Cancer* **6**, 738–744 (2020). Sep.

102. Patel, V. G., et al. Characterizing patterns of disease progression in patients with genitourinary cancers treated with immune checkpoint inhibitors. *J. Clin. Oncol.* **38**, 482–482 (2020).

103. Hue, J., Valinciute, Z., Thavaraj, S. & Veschiini, 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.* **141**, 106399 (2023).

104. Rade, M. et al. A reliable transcriptomic risk-score applicable to formalin-fixed paraffin-embedded biopsies improves outcome prediction in localized prostate cancer. *Mol. Med.* **30**, 19 (2024).

105. Howell, R. et al. Localized immune surveillance of primary melanoma in the skin deciphered through executable modeling. *Sci. Adv.* **9**, eadd1992 (2023).

106. Kreuzaler, P. et al. Heterogeneity of Myc expression in breast cancer exposes pharmacological vulnerabilities revealed through executable mechanistic modeling. *Proc. Natl Acad. Sci. USA* **116**, 22399–22408 (2019).

107. Montagud, A. et al. Patient-specific Boolean models of signalling networks guide personalised treatments. *eLife* **11**, e72626 (2022).

108. Ruscone, M. et al. Multiscale model of the different modes of cancer cell invasion. *Bioinformatics* **39**, btad374 (2023).

109. Chuah, S. & Chew, V. High-dimensional immune-profiling in cancer: implications for immunotherapy. *J. Immunother. Cancer* **8**, e000363 (2020).

110. 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>.

111. An, G. & Cockrell, C. Drug development digital twins for drug discovery, testing and repurposing: A schema for requirements and development. *Front Syst. Biol.* **2**, 928387 (2022).

112. 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>.

113. 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* **28**, 103605 (2023).

114. Russo, G. et al. Moving forward through the in silico modeling of tuberculosis: a further step with UISS-TB. *BMC Bioinform.* **21**, 458 (2020).

115. Letter of support for UISS-TB-DR. [https://www.ema.europa.eu/en/documents/other/letter-support-universal-immune-system-simulator-tuberculosis-disease-model-uiss-tb-dr\\_en.pdf](https://www.ema.europa.eu/en/documents/other/letter-support-universal-immune-system-simulator-tuberculosis-disease-model-uiss-tb-dr_en.pdf).

116. Viceconti, M. & Emili, L., eds. *toward Good Simulation Practice: Best Practices for the Use of Computational Modelling and Simulation in the Regulatory Process of Biomedical Products* (Springer Nature Switzerland, 2024).

117. Madabushi, R. et al. The US Food and Drug Administration's Model-Informed Drug Development Paired Meeting Pilot Program: early experience and impact. *Clin. Pharm. Ther.* **106**, 74–78 (2019).

118. Wang, Y. et al. Model-Informed Drug Development: current US regulatory practice and future considerations. *Clin. Pharm. Ther.* **105**, 899–911 (2019).

119. Marshall, S. et al. Model-informed drug discovery and development: current industry good practice and regulatory expectations and future perspectives. *CPT Pharmacomet. Syst. Pharm.* **8**, 87–96 (2019).

120. Väri, A., Kranz-Zuppan, P. & de la Cruz, R. IEC 62304 ed. 2: software life cycle standard for health software. *Stud. Health Technol. Inf.* **264**, 868–872 (2019).

121. Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions | FDA [Internet]. [cited 2024 Aug 5]. Available from <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/assessing-credibility-computational-modeling-and-simulation-medical-device-submissions>.

122. Niarakis, A. et al. Addressing barriers in comprehensiveness, accessibility, reusability, interoperability and reproducibility of computational models in systems biology. *Brief. Bioinform.* **23**, bbac212 (2022).

123. Sinaci, A. A. et al. From raw data to FAIR data: the fairification workflow for health research. *Methods Inf. Med.* **59**, e21–e32 (2020).

124. Tiwari, K. et al. Reproducibility in systems biology modelling. *Mol. Syst. Biol.* **17**, e9982 (2021).

125. Bergmann, F. T. et al. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinform.* **15**, 369 (2014).

126. Huang, P.-H., Kim, K.-H. & Schermer, M. Ethical issues of digital twins for personalized health care service: preliminary mapping study. *J. Med. Internet Res.* **24**, e33081 (2022).
127. Reddy, S., Allan, S., Coghlan, S. & Cooper, P. A governance model for the application of AI in health care. *J. Am. Med Inf. Assoc.* **27**, 491–497 (2020).
128. Bruynseels, K., Santoni de Sio, F. & van den Hoven, J. Digital twins in health care: ethical implications of an emerging engineering paradigm. *Front. Genet.* **9**, 31 (2018).
129. Semakova, A. & Zvartau, N. Data-driven identification of hypertensive patient profiles for patient population simulation. *Procedia Comput. Sci.* **136**, 433–442 (2018).
130. Martinez-Velazquez, R., Gamez, R. & Saddik, A. E. "Cardio Twin: A Digital Twin of the human heart running on the edge," 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Istanbul, Turkey, 2019, pp. 1–6, <https://doi.org/10.1109/MeMeA.2019.8802162>.

## Acknowledgements

This work is based on work during 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 supporting the workshop. Lastly, the authors thank Dr Laurence Calzone for her insightful comments. The authors also acknowledge the following support: A.N. was supported by a public-private partnership grant (CIFRE contract, no. 2020/0766) with SANOFI-AVENTIS R&D. R.L. 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). L.C.M., B.S., and L.G. acknowledge support from the European Commission (grants H2020-SC1-BHC-11-2020 963845 and DigitalEurope EDITH-CSA 101083771). M.L.B. acknowledges support from the National Institutes of Health (grants R24 GM137787 and P41 EB023912). L.L.F. acknowledges financial support from the Defense Advanced Research Projects Agency (grant HR00112220038). L.H. acknowledged support from NIH/NCI Transition Career Development Award to Promote Diversity (K22-CA237857). L.S.V. acknowledges support from the National Institutes of Health (grants R01 AI135128, R01 HL169974, K25 AI175668). P.H. acknowledges support from the Prime Ministers' Research Fellowship (PMRF). H.B. acknowledges the support of ZonMw (Grant No. 10430012010015). N.N. acknowledges support from the Knowledge Foundation (20200017) and The Swedish Fund for Research Without Animal Experiments. A.M., J.C.C., and A.V. acknowledge support from the European Commission (grants CREXDATA 101092749, PerMedCoE 951773, and EDITH-CSA 101083771). F.M. acknowledges support from the Italian Ministry of Health (grant "Ricerca Corrente" Linea 4 Project 5 and "5 per 1000–2021" Grant No. 5M-2021-23683787) and the European Commission with HORIZON programme (Grant No. 101046203—BY-COVID). J.A.G. acknowledges support from the National Institutes of Health (grants U24 EB028887) and the National Science Foundation (grants NSF 2303695, NSF 2120200, and NSF 1720625). G.A. and C.C. acknowledge support from the National Institutes of Health Award U01EB025825 and the Defense Advanced Research Projects Agency through Cooperative Agreement D20AC00002 awarded by the U.S. Department of the Interior (DOI), Interior Business Center. K.R. was supported by the BMBF (Federal Ministry of Education and Research) in DAAD project 57616814 (SECAI, School of Embedded and Composite AI) as part of the program Konrad Zuse Schools of Excellence in Artificial Intelligence.

Further, KR was supported by the imSAVAR and the CERTAINTY project. imSAVAR received funding from the Innovative Medicine Initiative 2 Joint Undertaking (JU) under grant agreement No 853988. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and JDRF INTERNATIONAL. The CERTAINTY project is funded by the European Union (Grant Agreement 101136379).

## Author contributions

Manuscript assembly and first draft: A.N. Figure preparation: A.N. with input from S.S., L.V., A.Z., K.R., L.L., J.A.G., R.S.M.S., A.Z., B.S., E.T., A.F., A.Z. and G.A. Table preparation: A.N., M.M., M.L.B., M.E.M., A.M., M.R.M., V.N. Scientific content and text contributions: A.N., R.L., G.A., Y.I., J.F., A.F., K.R., M.R.M., L.G., L.L., L.V., M.L.B., F.M., L.L.F., S.F., A.M., V.N., M.M., E.T., M.M.T., L.A.H., T.J.S., C.C., A.E.S., H.B., A.S., K.R., A.A.H., M.E.M., B.S., E.H.V., S.M.B., A.M.V., P.K., L.S.V., P.H., A.K., N.N., W.G., R.S.M.S., J.T., V.D.T.T., J.C.C., S.S., A.V., A.Z., J.A.G. Manuscript writing: A.N., R.L., G.A., J.F., A.F., K.R., M.R.M., L.L., L.V., M.L.B., F.M., L.L.F., S.F., A.M., V.N., M.M., E.T., M.M.T., L.A.H., T.J.S., C.C., A.S., A.E.S., K.R., M.E.M., B.S., E.H.V., S.M.B., A.M.V., P.K., L.S.V., P.H., A.K., R.S.M.S., V.D.T.T., S.S., A.Z. Editing: A.N., R.L., J.T., A.V., R.L., Y.I., J.F., M.R.M., K.R., A.F., E.T., L.L., V.D.T.T., A.Z. All authors agree with the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41540-024-00450-5>.

**Correspondence** and requests for materials should be addressed to Anna Niarakis.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

<sup>1</sup>Molecular, Cellular and Developmental Biology Unit (MCD), Centre de Biologie Intégrative (CBI), University of Toulouse, UPS, CNRS, Toulouse, France. <sup>2</sup>Lifeware Group, Inria, Saclay-Île de France, Palaiseau, France. <sup>3</sup>Department of Medicine, University of Florida, Gainesville, FL, USA. <sup>4</sup>Department of Surgery, University of Vermont Larner College of Medicine, Vermont, USA. <sup>5</sup>Faculty of Medicine Hebrew University, Hadassah Medical Center, Jerusalem, Israel. <sup>6</sup>UCL Cancer Institute, University College London, Paul O'Gorman Building, 72 Huntley Street, London WC1E 6BT, UK. <sup>7</sup>Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway. <sup>8</sup>The Cancer Clinic, St Olav's University Hospital, Trondheim, Norway. <sup>9</sup>Department of Biotechnology and Nanomedicine, SINTEF

Industry, Trondheim, Norway. <sup>10</sup>Department of Diagnostics, Fraunhofer Institute for Cell Therapy and Immunology, Leipzig, Germany. <sup>11</sup>Institute of Clinical Immunology, Medical Faculty, University Hospital, University of Leipzig, Leipzig, Germany. <sup>12</sup>Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany. <sup>13</sup>Department of Biomedical Informatics & Data Science, Yale School of Medicine, New Haven, CT, USA. <sup>14</sup>Prometheus Division of Skeletal Tissue Engineering, KU Leuven, Leuven, Belgium. <sup>15</sup>Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven, Leuven, Belgium. <sup>16</sup>Biomechanics Research Unit, GIGA Molecular and Computational Biology, University of Liège, Liège, Belgium. <sup>17</sup>Faculty of Dentistry Oral & Craniofacial Sciences, King's College London, London, UK. <sup>18</sup>Biocomplexity Institute and Department of Intelligent Systems Engineering, Indiana University, Bloomington, Indiana 47408, USA. <sup>19</sup>Center for Cell Analysis and Modeling, UConn Health, Farmington, CT 06030, USA. <sup>20</sup>Department of Epidemiology, Preclinical Research and Advanced Diagnostic, National Institute for Infectious Diseases 'Lazzaro Spallanzani' - I.R.C.C.S., Rome, Italy. <sup>21</sup>Mathematics Department and Center of Mathematics, University of Beira Interior, Covilhã, Portugal. <sup>22</sup>Barcelona Supercomputing Center (BSC), Barcelone, Spain. <sup>23</sup>Institute for Integrative Systems Biology (I2SysBio), CSIC-UV, Valencia, Spain. <sup>24</sup>Institut Curie, Université PSL, F-75005 Paris, France. <sup>25</sup>INSERM, U900, F-75005 Paris, France. <sup>26</sup>Mines ParisTech, Université PSL, F-75005 Paris, France. <sup>27</sup>Université de Toulouse, Inserm, CNRS, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France. <sup>28</sup>Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway. <sup>29</sup>Department of Mathematics and Statistics, University of Richmond, Richmond, VA, USA. <sup>30</sup>Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR, USA. <sup>31</sup>Interdisciplinary Graduate Program in Cell and Molecular Biology, University of Arkansas, Fayetteville, AR, USA. <sup>32</sup>Cancer Biology Program, Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR, USA. <sup>33</sup>Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA. <sup>34</sup>Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, Maastricht, The Netherlands. <sup>35</sup>INRIA Paris/CNRS/École Normale Supérieure/PSL Research University, Paris, France. <sup>36</sup>Andalusian Platform for Computational Medicine, Andalusian Public Foundation Progress and Health-FPS, Seville, Spain. <sup>37</sup>Bioinformatics Core Unit, Luxembourg Centre of Systems Biomedicine LCSB, Luxembourg University, Esch-sur-Alzette, Luxembourg. <sup>38</sup>Department of Mathematics and Statistical Science, University of Idaho, Moscow, ID 83844-1103, USA. <sup>39</sup>Novadiscovery, Lyon, France. <sup>40</sup>Department of Electrical and Computer Engineering, University of Patras, Patras, Greece. <sup>41</sup>Department of Bioengineering, Indian Institute of Science, Bengaluru, India. <sup>42</sup>School of Medical Sciences, Faculty of Medicine and Health, Örebro University, SE-70182 Örebro, Sweden. <sup>43</sup>X-HiDE - Exploring Inflammation in Health and Disease Consortium, Örebro University, Örebro, Sweden. <sup>44</sup>Multiscale Mathematical Modelling of Living Systems program (M3-LSP), Kimberley, South Africa. <sup>45</sup>Department of Mathematical Sciences, Sol Plaatje University, Kimberley, South Africa. <sup>46</sup>Private Bag X5008, Kimberley 8300, South Africa. <sup>47</sup>European Bioinformatics Institute, European Molecular Biology Laboratory (EMBL-EBI), Hinxton, Cambridge, UK. <sup>48</sup>Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK. <sup>49</sup>Department of Microbiology & Immunology and Department of Biostatistics & Computational Biology, University of Rochester Medical Center, Rochester, NY 14642, USA. <sup>50</sup>Vital-IT Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland. <sup>51</sup>Institute of Biomedical Engineering and Technology, Ghent University, Gent, Belgium. <sup>52</sup>Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand. <sup>53</sup>ICREA, 23 Passeig Lluís Companys, 08010 Barcelona, Spain. <sup>54</sup>In silico R&D, Evotec, 31400 Toulouse, France.

✉ e-mail: [anna.niaraki@univ-tlse3.fr](mailto:anna.niaraki@univ-tlse3.fr)