

Dynamic Digital Twin

Diagnosis, Treatment, Prediction, and Prevention of Disease During the Life Course

Mulder, Skander Tahar; Omidvari, Amir-Houshang ; Rueten-Budde, Anja J.; Hai, Rihan; Akgün, Can; Tax, David M.J.; Reinders, M.J.T.; Reinders, Marcel; Visch, Valentijn; More Authors

DOI

[10.2196/35675](https://doi.org/10.2196/35675)

Publication date

2022

Document Version

Final published version

Published in

Journal of Medical Internet Research

Citation (APA)

Mulder, S. T., Omidvari, A.-H., Rueten-Budde, A. J., Hai, R., Akgün, C., Tax, D. M. J., Reinders, M. J. T., Reinders, M., Visch, V., & More Authors (2022). Dynamic Digital Twin: Diagnosis, Treatment, Prediction, and Prevention of Disease During the Life Course. *Journal of Medical Internet Research*, 24(9), Article e35675. <https://doi.org/10.2196/35675>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Viewpoint

Dynamic Digital Twin: Diagnosis, Treatment, Prediction, and Prevention of Disease During the Life Course

Skander Tahar Mulder^{1*}, PhD; Amir-Houshang Omidvari^{2,3*}, MD, PhD; Anja J Rueten-Budde⁴, PhD; Pei-Hua Huang⁵, PhD; Ki-Hun Kim⁶, PhD; Babette Bais⁷, PhD; Melek Rousian⁷, MD, PhD; Rihan Hai⁸, PhD; Can Akgun^{8,9}, PhD; Jeanine Roeters van Lennep¹⁰, MD, PhD; Sten Willemsen⁴, PhD; Peter R Rijnbeek¹¹, PhD; David MJ Tax¹, PhD; Marcel Reinders¹, IR, PhD; Eric Boersma², PhD; Dimitris Rizopoulos⁴, PhD; Valentijn Visch¹², PhD; Régine Steegers-Theunissen^{7*}, MD, PhD

¹Pattern Recognition Lab, Mathematics and Computer Science, Technical University Delft, Delft, Netherlands

²Department of Cardiology, Erasmus Medical Center, Rotterdam, Netherlands

³Department of Public Health, Erasmus Medical Center, Rotterdam, Netherlands

⁴Department of Biostatistics, Erasmus Medical Center, Rotterdam, Netherlands

⁵Department of Medical Ethics and Philosophy, Erasmus Medical Center, Rotterdam, Netherlands

⁶Department of Industrial Engineering, Pusan National University, Busan, Republic of Korea

⁷Obstetrics and Gynaecology, Erasmus Medical Center, Rotterdam, Netherlands

⁸Web Information Systems Group, Mathematics and Computer Science, Technical University of Delft, Delft, Netherlands

⁹Bioelectronics Section, Department of Microelectronics, Faculty of Electrical Engineering, Technical University Delft, Delft, Netherlands

¹⁰Department of Internal Medicine, Erasmus Medical Center, Rotterdam, Netherlands

¹¹Department of Medical Informatics, Erasmus Medical Center, Rotterdam, Netherlands

¹²Industrial Design, Mathematics and Computer Science, Technical University Delft, Delft, Netherlands

* these authors contributed equally

Corresponding Author:

Régine Steegers-Theunissen, MD, PhD

Obstetrics and Gynaecology

Erasmus Medical Center

Dr. Molewaterplein 40

Rotterdam, 3015GD

Netherlands

Phone: 31 10 7038256

Email: r.steegers@erasmusmc.nl

Abstract

A digital twin (DT), originally defined as a virtual representation of a physical asset, system, or process, is a new concept in health care. A DT in health care is not a single technology but a domain-adapted multimodal modeling approach incorporating the acquisition, management, analysis, prediction, and interpretation of data, aiming to improve medical decision-making. However, there are many challenges and barriers that must be overcome before a DT can be used in health care. In this viewpoint paper, we build on the current literature, address these challenges, and describe a dynamic DT in health care for optimizing individual patient health care journeys, specifically for women at risk for cardiovascular complications in the preconception and pregnancy periods and across the life course. We describe how we can commit multiple domains to developing this DT. With our cross-domain definition of the DT, we aim to define future goals, trade-offs, and methods that will guide the development of the dynamic DT and implementation strategies in health care.

(*J Med Internet Res* 2022;24(9):e35675) doi: [10.2196/35675](https://doi.org/10.2196/35675)

KEYWORDS

digital health; digital twin; machine learning; artificial intelligence; obstetrics; cardiovascular; disease; health

Background

Overview of the Concept

Interest has been growing worldwide in the virtual representation of a physical asset, process, or system to model and simulate a real-world event. This representation, called a digital twin (DT), can represent the real-time performance or failure incidence of a deterministic system (eg, a factory production line) [1]. Since the introduction of the concept of a DT in 2003, DTs have been developed and used in areas such as construction, power, and oil and gas industries. In the aforementioned industries, the DT served mainly as an umbrella term for managing data and models of a closed system, and these models then guided actions taken in the system. However, for health care, DT is a new concept in need of a working definition. The DT in health care is not a single technology but a domain-adapted multimodal modeling approach incorporating methodologies for the acquisition, management, analysis, prediction, and interpretation of health-related data, aiming to improve medical decision-making and patient lifestyle choices.

Health care providers strive to obtain and use all relevant information on patients for personalized decision-making in clinical practice, considering the available evidence, clinical guidelines, and patient preferences. For example, wearable technologies such as smartwatches have enabled individuals to record their health data continuously, which can form a part of personal health records [2]. They can be very useful for clinical decision-making in practice, and they are currently in use for detection or monitoring of some disorders like atrial fibrillation; however, they are not yet completely in use in practice [3]. Unsolved challenges such as human information overload [4]; variable quality of routinely collected data from medical, lifestyle, mental, societal, and environmental sources; and limited interoperability of digital systems in health care are barriers to use these scattered large data sets, also referred to as Big Data. In addition, the current methodological approaches in evidence-based medicine are not able to use all this information for medical decision-making, as the population generating these data is heterogeneous, and previously discovered relationships between predictor and outcomes might not always hold for subpopulations. To alleviate these limitations, the underlying methods and workflow for data use need to be adapted. Furthermore, currently available analytical models in health care, such as decision aids using risk prediction models (eg, the U-prevent software) [5,6], are often developed based on limited data and the defined outcomes of interest of a small number of health care professional. Combining high-quality, scattered data sets for inference by applying a comprehensive multimodal data management approach in health care such as DT is deemed necessary to design applications that allow for diagnosis, treatment, prediction, and prevention of disease. Moreover, successfully implemented DT has the potential to improve health care by optimizing individual health outcomes from the earliest moment in the life course by offering personalized medicine in primary, secondary, and tertiary health care [7].

Since 2015, publication of papers describing the concept of applying a DT to health care to solve health challenges (eg, reducing adverse outcomes in certain patient populations or understanding important factors such as dietary factors or biomarkers) [8-13] has increased. However, literature descriptions indicate open theoretical and practical challenges that need to be addressed before we can work toward efficiently implementing a DT in health care. Challenges are related to (1) redefining a target population and matching data set to develop a DT or (2) implementing a physical or data-driven approach with limited data and follow-up to learn causal personal patterns. In addition, we suggest a convergence of experts in the medical, technical, scientific, and ethical domains, which is required to design a DT that meets requirements from several theoretical backgrounds. In our vision, DT development starts with the identification and description of a medical problem such as the prediction of cardiovascular complications in a patient who experienced hypertension during periconception (Figure 1). Solving such a medical problem involves striving for the best health- and value-based outcomes and addressing ethical value goals such as health equality (Figure 2). With these outcomes and value goals in place, a technological system incorporating best methodological practices such as reliability and reproducibility can be developed. After the initial system is developed, many feedback loops between domains occur in order to optimize requirements across domains. This digital health system is then studied to develop best practices using the scientific domain and keeping the healthy patient life course journey in mind. In an open system such as the life course, as opposed to a closed physical system, causal drivers for change in health conditions might not be known; therefore, the DT uses algorithms that learn drivers of dynamic data such as user experiences, medical data fields, medical scans, etc. To serve that purpose, the DT acquires time-series data and updates predictions in an online and data-driven manner. Additional challenges are defining medical applications of a DT in health care and considering ethical values.

In this viewpoint paper, we address the aforementioned challenges and put forward a vision for a dynamic DT in health care for optimizing individual patient health journeys characterized by healthy outcomes and positive experiences. Our cross-domain, which has the ethical, medical, technical, and scientific definition of the DT, will define future goals, trade-offs, and methods to guide the implementation strategies and iterative development of a dynamic DT in health care. In our paper, we propose a dynamic DT for health care that applies to the management of dynamic patient data and models. These models of health and disease are dynamic because they are developed, trained, tested, and updated to meet the dynamic value goals stemming from ethical, medical, and technological domains. Additionally, the effect of these interventions (such as personal lifestyle advice) guided by models is scientifically evaluated and updated to continually strive for better health outcomes (Figure 2).

Figure 1. Digital twin in clinical practice is modeled for hypertension starting from the periconception period until later stages of life. There are four discrete steps in which the digital twin can bring additional value by using patient data to recommend interventions optimized for the relevant values and outcomes of interest. CVD: cardiovascular disease; DT: digital twin.

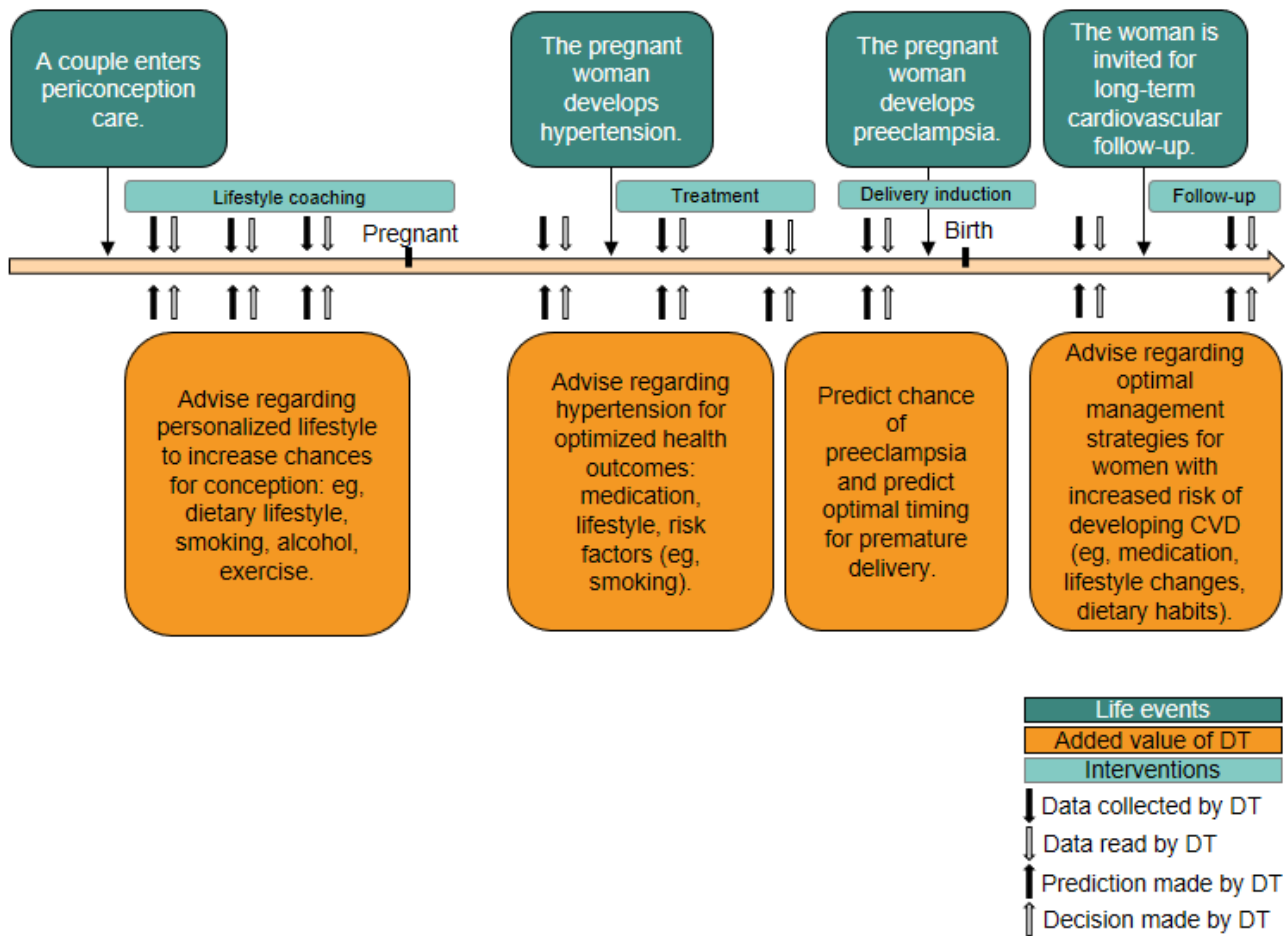
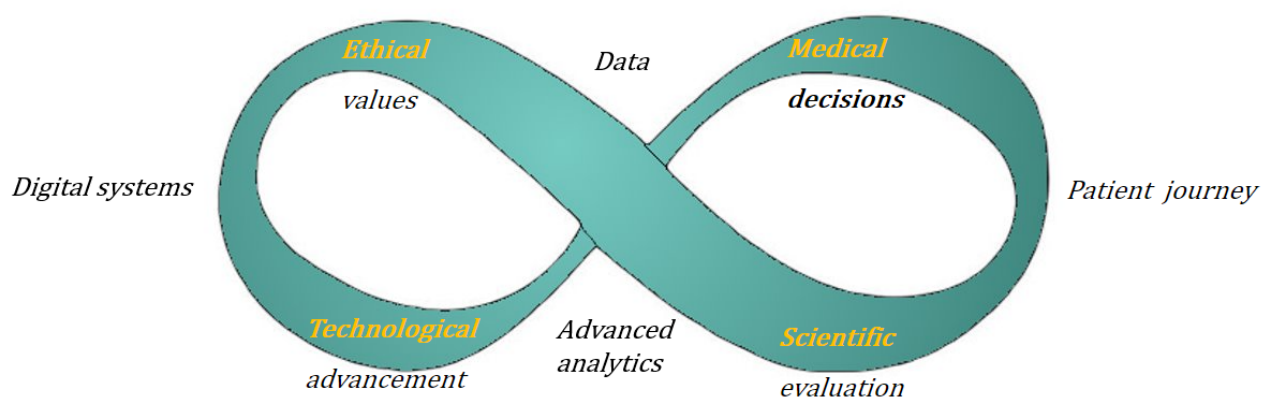


Figure 2. Digital systems and the patient health journey are improved by a continuous feedback loop across domains interacting to develop a digital twin.



Potential Future Application of a DT in Health Care

Currently, the translation from data to evidence relating to disease ontology, causation, and effectiveness of treatments into clinical practice is a slow and partially data-driven process [14], and even for some of the recommendations in the clinical

practice guidelines, there is no robust evidence available [15]. The dynamic DT can support a health journey by providing easy access to comprehensive patient data for the patient and health care providers. The DT can integrate data from health care sources like hospitals, general practitioners, and laboratories and from home health monitoring devices such as wearables.

Real-time integrated access to data will allow analysis on dynamically acquired data. This feature enables health care professionals to get a relevant view of factors influencing disease and health conditions of a patient, which can guide actions to optimize health care pathways. Different prediction algorithms can be used to answer medical questions, and real-world personalized predictions can be updated to meet the dynamic circumstances of individuals and their medical history. Testing these prediction models in the real world prospectively can help to define and research hypotheses about potential causal relationships between prediction algorithms and outcomes [16]. Formal testing and evaluating medical decisions in a DT can result in changing the focus of scientific studies to incorporate testing the implementation of a DT-powered intervention such as patient-tailored lifestyle advice in the absence of a data-driven decision framework. Therefore, this DT approach can expedite the process of knowledge translation into medical decision-making in clinical practice while striving for iterative improvement of elements in different domains of the DT.

Example of a Patient Journey From the Earliest Moment of the Life Course

The periconceptional period refers to the 6 months around conception; it is the earliest and one of the most critical periods in life, with long-lasting impact on health and disease later in life and in future generations [17]. The patient journey and data generated during this period can help to illustrate our vision of the dynamic DT and how it would intervene in medical practice. To develop a DT from the earliest moment of the life course, the first set of data would include static (demographics) and dynamic (conditions such as subfertility, hypertension, lifestyle, and vital information from wearables) covariates relating to the parents. This DT supports dynamic medical decision-making (Figure 1). The first example is based on subfertility, for which enrollment in the evidence-based lifestyle coaching program is recommended. This eHealth coaching is not static but dynamic as it gives advice based on the variable input of the individual. For example, if healthy food intake is sufficient, parents will be empowered to maintain this behavior. A second example of how the DT can be used is aiding physicians to prevent short-term adverse outcomes of preeclampsia in pregnant women. The mobile phone app, a smartwatch, and lab measurements combined can yield predictions of risk before and after intervention, empowering medical decision-making and resulting in a better outcome for mother and baby. A third example to highlight the long-term benefits of a DT is hypertension. By mechanisms not fully understood, high blood pressure and preeclampsia during pregnancy yield increased cardiovascular risk for the unborn child and mother in later life years [18-20]. This data-driven association can make more comprehensive follow-up of these individuals valuable as this information can be used to monitor and potentially prevent worse outcomes such as cardiovascular disease. In our example, the DT has enacted medical decision rules driven by pattern recognition in patient data. In general terms, the rules obtained from pattern recognition systems are chosen so that they optimize for healthier outcomes.

Why Now: Innovations That Power the DT

In 2016, a pivotal paper was published that described how data should be managed according to the following 4 principles: findable, accessible, interoperable, and reusable (FAIR) [21]. These principles promote data accessibility to power innovation and are the basis for the following 4 technical innovations that accelerate the progression of DT in health care. The first innovation is data storage, where we can securely store big data in the cloud by designing access rights for each service and algorithm [22]. Data collection innovations such as wearables, which can act as continuous vital sign monitoring systems, feed the new data to the storage keeping the data up to date and relevant. Connectivity innovations such as the Internet of Things then trigger the training and prediction algorithms with the updated information. Finally, computing innovations such as more powerful data processors like tensor processing units can fit bigger and more flexible models with higher performance to yield better representations of disease. These innovations have changed the way we form hypotheses about the physical reality. Increased data and processing power formats require and allow for new dynamic pattern recognition methods that can be described as computer pattern recognition or artificial intelligence [23].

These innovations power the DTs and allow for deep personalized predictions that leverage patient-specific dynamics such as a specific disease pathway relating to biomarker panels based on genomics, metabolomics, or proteomics, which have previously been prioritized or identified. These predictions need to be translated in a concise way to an individual, patient or health care provider. For example, a score, dashboard, or written advice can guide individuals to healthy behavior. In our example, continuously monitored medical and lifestyle data of a couple contemplating pregnancy supplies the DT with data that can be used to give personalized integrated medical advice and brings personalized medicine closer [24].

Themes in Pattern Recognition

There are multiple paradigms for the underlying data analytic platform of a DT relating to pattern recognition approaches. We can use a data-driven or physics-based approach. Physics-based models are based on the understanding of the phenomenon and formulation in a mathematical model, with underlying assumptions that potentially oversimplify the phenomenon [25]. Data-driven (associative) models avoid these assumptions, but they lack interpretability and are sensitive to bias in the model development data. We should choose our (inductive) bias in a way that allows us to optimize interpretability and predictive performance by leveraging domain medical expert-level knowledge. When the physical drivers of disease are not known, these associative data-driven methods can generally be used to diagnose patients, but they cannot identify causation. An alternative approach within the associative methodology is to enrich models with counterfactuals, which leverage counterfactual information to causally weight predictions [26].

In the DT for health care, pattern recognition algorithms are used for predicting continuous outcomes such as blood pressure, classification of diseases, and risk assessment. Medical data generally comprise different types of data and algorithms. For

example, they include longitudinal data with missing predictors and variables, which can be used to enrich the model. Research on how to associatively predict with longitudinal data, population-level data, and data with missing values is ongoing. A statistical modeling approach, which can relate a limited number of predictors to a longitudinal outcome with missing values, uses linear mixed models [27]. Deep-learning methods [28,29], where we flexibly fit either a neural network architecture or many random kernels, are used when a large number of predictors is present. This modeling approach seems to perform well in this domain, but least absolute shrinkage and selection operator regression-based techniques may also yield good performance on this type of data [30].

Medical patient data can comprise higher dimensional data, such as echography, and multiple sensor data, such as electrocardiograms, which are continuously collected in a setting such as the coronary care unit. Imaging data is typically only acquired once and not collected continuously. Although each modality may be information rich, the number of measurements per individual may be limited, causing models to overfit on that individual. To prevent overfitting, we use feature selection and methods that can find a lower representation of these data, such as an auto encoder. Feature selection and engineering can prevent overfitting and summarize high-dimensional data in a feature vector, which can be used to describe disease progression.

As models are implemented in clinical practice, other challenges, such as model drift and the appearance of new class definitions, can arise that cause previously found associations to change. The online learning paradigm can adapt to this change by continuously updating the model with new information and thereby learning to adapt to changes in the environment [31].

Last, because the medical profession generally deals with interventions, another theme of interest is causal machine learning, where the focus is on creating a model that can predict the change in target output if a predictor were changed (eg, change in diet or blood pressure). The challenging part is inferring causality from observational data, as in medicine we would not only like to associatively predict but intervene in patient health journeys to prevent adverse outcomes. The current gold standard is randomized intervention data, but these data are resource-intensive to collect, and the clinical trial setting

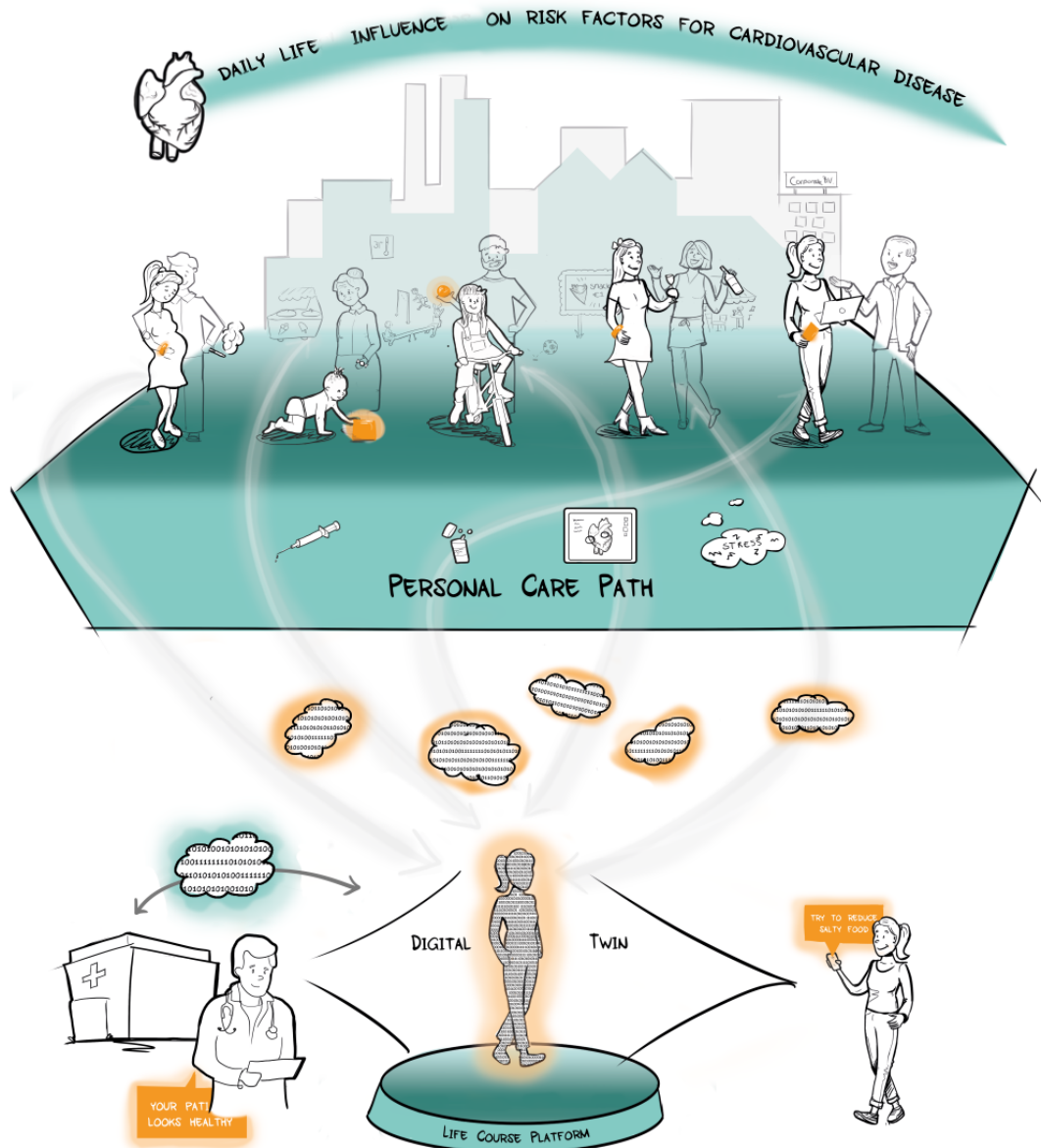
does not reflect real life, so the distribution of these data may not fit the general population [32]. However, there are powerful alternatives for causal questioning that may have a future in drug efficacy and safety evaluation on real-world data [33]. Moreover, there are tools being used for evaluating interventions with methods, which can possibly extend into other medical domains [34].

In summary, applying accurate and understandable models on data that fit well to the target individual will be paramount to the success of the DT. Equally important is deconvoluting the causal factors that influence health outcomes, as this knowledge will power interventions based on personalized causal factors. Using a combination of these methods can form the technical analysis part of the DT framework (Figure 2).

Dynamic Digital Wins

A dynamic DT (Figure 3) is a twin whose relationships—and the data contained within—change over time as the conditions of people's lives change over time (eg, a young healthy individual who needs preventive care to increase the chance of fertility, a patient treated for a life-threatening disease at the coronary care unit). The dynamic aspect also refers to the changing of the model targets according to the relevant question and value trade-off. As an individual experiences different phases of life, the relevant dynamic DTs are triggered to make a prediction using the up-to-date data. To keep the predictions relevant to the patient, the DT needs to be updated in a continuous manner, preferably with the least amount of burden on the individual. The continuous monitoring of vital signs (for patients in a hospital setting and at home) using wearable devices allows for prompt detection of developing pathologies and early intervention, which may not be possible using standard intermittent vital sign measurements [35]. To have a successful implementation of a dynamic DT, a minimal requirement is to start with a FAIR data set. On this data set, algorithms must be trained and tested, and this output needs to be translated into medical or lifestyle advice. The next step is the validation of this advice to see if this decision-support tool improves patient outcomes. For the full DT, there are multiple requirements: (1) continuous collection of, for example, medical and user experience data, (2) integration of new data with existing data in the databases, and (3) improvement of accessibility of systems that access and store new data [36].

Figure 3. A digital twin encompassing different aspects of the life course can be queried to improve medical decision-making to reduce cardiovascular complications. In orange, we show data (and data generators). In green, we display the digital life course platform and real-life course. Together they supply information to each other to support a healthy patient lifestyle by defining a personalized care path.



Data Management

The first requirement in data management is the collection of high-quality data, which includes similar distribution of patients and individuals as the target individual. As the clinical data is collected from different sources, an important challenge is to synergize, homogenize, integrate, and standardize the heterogeneous data to have a standardized data model. In such a data model, we transform the local data source into the common data model via extract, transform, and load procedures [37]. This data model should live in a system that can sustain FAIR principles. Moreover, patients/researchers and health care providers should have an easy and flexible way to add or remove medical and lifestyle data sources [24]. Data security and privacy are important for DT development in the medicine and health care sectors. Open challenges are cleaning and preparing the data [38] and protecting the privacy of data residing on cloud-based data storage platforms [39]. The rising popularity of storing health data on cloud platforms requires deliberate

design in managing data access control and data ownership in combination with schemes such as federated analysis.

Values

Values are the quality of a digital health service as it is experienced by an individual in relation to their needs [40]. Based on this definition, values are created when users such as patients, clinicians, nurses and others use the DT to address their challenges [41,42]. Representative examples of such challenges were identified by the World Health Organization in 2018 [43] and include efficiency, cost, and access to information.

In our early life course patient journey example, the DT can solve information challenges by collecting high-quality data through wearable devices and then providing easy access and visual summaries of the data to the patient and their treating health care professional. Additionally, the DT can provide personalized predicted probabilities of adverse health outcomes

and suggest treatment consisting of lifestyle advice or drug treatment to lower those probabilities. The DT can aid the use challenges of low adherence to treatment by visualizing the effect that treatment may have on future health outcomes so that the patient can better understand and be motivated by the potential benefit of adherence to treatment. The DT addresses efficiency challenges by dynamically updating a patient's predicted probabilities for adverse health outcomes, which enables an early reaction to increased risks. This information may also be used to choose the optimal time between health checks and interventions; in limited care situations, it may be used to prioritize patients with greater needs in triage settings. The cost challenges can be addressed by early risk detection and interventions of the DT, which may prevent the need for costly treatments at later stages of disease progression. To create a valuable DT for users, it may help to include lead users in the development team [24].

Legal and Ethical Considerations

Value Trade-offs

The DT allows us to optimize for healthy outcomes at the cost of other values (eg, fairness). While incorporating more health-related data might increase a prediction's accuracy, it could also raise concerns about surveillance health care and erode the trust between the health care system and individuals [44]. A DT measuring patient data (such as sleep) and the feedback of health guidance could even affect the parameters it is measuring. For example, commercial consumer wearables are not accurate in estimating sleep [45], and the results may worry consumers and even cause sleep problems. An increased number of measurements may identify patients that are considered atypical. Currently, we lack a clear understanding of how to interpret slight deviations from the normal ranges and means and whether these asymptomatic deviations will lead to future diseases. Balancing the ideal of early diagnosis and reducing overdiagnosis and overtreatment can therefore be challenging [46,47]. The attempt to introduce personalized lifestyle advice to curb lifestyle diseases such as obesity, diabetes, and hypertension also risks promoting an overly individualized view of health management.

Additionally, many have expressed concern about worsening existing health inequity [48,49]. People from lower socioeconomic backgrounds might not have the means to access the digital device required for a DT service, or they might not know how to benefit from the information provided to them due to cognitive constraints such as limited digital or health literacy and external constraints. Instead of empowering users by offering them more health-related information, the DT might burden users with a sense of guilt or anxiety and give rise to the idea that users who do not make the advised change could be accused of being responsible for their adverse health outcomes [50-52].

Data Governance and Accountability

The digitalization of health care gives rise to new legal and accountability issues. For instance, who owns the health data gathered by the DT? Is it owned by the patient, the health care provider, or a third party? Is it morally permissible for a DT service provider to sell user data to a third party? To avoid data being used against the patients' best interest, Schwartz et al [24] suggest 3 principles: (1) patients own their data, which empowers patients to protect the privacy of their data against misuse, (2) patients must provide explicit informed consent for the use of their data, and (3) advocacy efforts should enshrine patient data ownership and access into law. The implementation of these principles poses technical and legislative challenges. Additionally, DTs are susceptible to biases present in the data from which they are developed. Data included in the DT should be reviewed and methods to remedy biases should be considered to avoid the perpetuation of historical biases [24]. De Laat [53] and Nissenbaum [54] describe the obstacles of accountability of machine learning algorithms and how this relates to shared development, human and computer errors, and a culture where it is easy to blame the technology. There is also a lack of understanding of how algorithms work [53,54]. Transparency of decision rules and oversight of the decision-making algorithms by governing bodies could solve these outstanding obstacles.

Scientific Domain

In this paper, the DT was defined in terms of requirements stemming from medical, technological, and ethical domains, but to evaluate the effectiveness of such a DT system, we require scientific convergence to test its merit and limitations (Textbox 1). In translational medical science, we aim to generate knowledge in clinical practice so we can intervene and improve health care processes [55]. To achieve this aim, we need to understand the disease process and uncover the relevant causal pathways that influence health outcomes. DTs can be helpful in combining cross-domain knowledge. As a DT incorporates more than one data set, such as physiological measurements, questionnaires, and lifestyle factors, there are many rich features for each patient or individual ready to be used for pattern recognition. Knowledge about individual contributors of disease progression, in turn, could lead to the identification of subtypes of disease with different disease ontologies and treatments, described as deep phenotyping [13]. The DT platform may allow us to test our predictions more efficiently in an umbrella or real-world trial [16,56]. Predicting patients prospectively and monitoring for outcomes may allow us to test the accuracy of the predictions from the DT. In a second step, randomized intervention based on these predictions can unveil causal factors, and this in turn can help us test our scientific hypotheses using real-world evidence and thereby moving science forward by generating medical knowledge and developing novel technical methodologies.

Textbox 1. Benefits of digital twin development in health care.

Health agency and promoting healthy outcomes

- More accurate diagnosis using integrated data
- Improved treatment selection for patients
- Prognosis of patient disease trajectories
- Real-time remote monitoring of health state
- Simulation of treatment and care processes to guide policy

Fair data and validation of research

- Data structures in the digital twin will make data more accessible
- Validation of scientific results provide an implementation framework for decision-making

Algorithm development for medical decision-making, diagnosis, and prevention

- Innovation in algorithms allows us to make more accurate predictions on unseen data
- Causality is inferred from observational data
- Privacy is preserved during data sharing and analysis
- Uncertainty is quantified from different sources and considered during the decision-making processes

Conclusion

In conclusion, translational science and medical care can be improved by following a DT life course approach: high-dimensional data collection and storage, patient trajectory modeling, outcome predictions, testing, model interpretation, and implementation in clinical practice. By committing multiple nonmedical domains to developing a DT, we aim to improve patient care journeys in a systematic and diligent way.

Future Research Opportunities

Our vision of a dynamic DT allows for collaboration of many researchers from different domains, where we can align our research applications in order to develop a dynamic DT. Some subjects related to the DT have not been fully investigated, and new opportunities for research have been defined. A part of the technical domain, which is under active investigation, is how

we can leverage personal and lifestyle factors to reach more accurate predictions on unseen data. This involves but is not limited to using nonrandom and sparse sampling as a predictor matrix to improve our models and infer causality.

We also need more work on how to identify new classes, such as disease types based on new distance metrics, and how to tune our models to meet patient preferences and different privacy settings. A large body of research in the ethical domain is dedicated to value trade-offs in the DT setting and investigates how we can balance values such as fairness, equality, and health. We are also concerned with data storage and federated data analytics, and we performed analysis on harmonized big data sets. Optimal design of the DT application and user interaction is required to meet the requirements of users while optimizing for values arising from different domains. The science domain should come into play to judiciously evaluate the data, models, application, and ability of the technologies to impact patient life course trajectories and decrease the incidence of disease.

Acknowledgments

We would like to thank the convergence program, Digital Twin Flagship Erasmus Medical Center and Technical University Delft for leading the technology program [57]. We acknowledge Marina Bos-de Vos for critically reviewing the paper. We thank Emma Schallers for illustrating [Figure 3](#).

Authors' Contributions

STM, A-HO, AJR-B, K-HK, BB, M Rousian, EB, VV, RH, CA, and RS-T contributed to designing the paper and wrote the first draft of paper. STM, A-HO, AJR-B, and RS-T revised and finalized the draft. All other authors (P-HH, JRvL, SW, PR, DT, M Reinders, and DR) critically reviewed and revised the paper (specifically the parts related to their expertise).

Conflicts of Interest

None declared.

References

1. Grieves MW. Product lifecycle management: the new paradigm for enterprises. *Int J Prod Devel* 2005;2(1/2):71. [doi: [10.1504/ijpd.2005.006669](https://doi.org/10.1504/ijpd.2005.006669)]
2. Roehrs A, da Costa C, Righi R, de Oliveira K. Personal health records: a systematic literature review. *J Med Internet Res* 2017 Jan 06;19(1):e13 [FREE Full text] [doi: [10.2196/jmir.5876](https://doi.org/10.2196/jmir.5876)] [Medline: [28062391](https://pubmed.ncbi.nlm.nih.gov/28062391/)]
3. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Apple Heart Study Investigators. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019 Nov 14;381(20):1909-1917 [FREE Full text] [doi: [10.1056/NEJMoa1901183](https://doi.org/10.1056/NEJMoa1901183)] [Medline: [31722151](https://pubmed.ncbi.nlm.nih.gov/31722151/)]
4. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017 Apr 10;17(1):36 [FREE Full text] [doi: [10.1186/s12911-017-0430-8](https://doi.org/10.1186/s12911-017-0430-8)] [Medline: [28395667](https://pubmed.ncbi.nlm.nih.gov/28395667/)]
5. Usher-Smith J, Emery J, Hamilton W, Griffin SJ, Walter FM. Risk prediction tools for cancer in primary care. *Br J Cancer* 2015 Dec 22;113(12):1645-1650 [FREE Full text] [doi: [10.1038/bjc.2015.409](https://doi.org/10.1038/bjc.2015.409)] [Medline: [26633558](https://pubmed.ncbi.nlm.nih.gov/26633558/)]
6. Dorresteijn J, Visseren F. U-prevent: gepersonaliseerde cardiovasculaire preventie. *Huisarts en Wet* 2020;63(12):75. [doi: [10.1007/s12445-020-0934-9](https://doi.org/10.1007/s12445-020-0934-9)]
7. Björnsson B, Borrebaeck C, Elander N, Gasslander T, Gawel DR, Gustafsson M, Swedish Digital Twin Consortium. Digital twins to personalize medicine. *Genome Med* 2019 Dec 31;12(1):4 [FREE Full text] [doi: [10.1186/s13073-019-0701-3](https://doi.org/10.1186/s13073-019-0701-3)] [Medline: [31892363](https://pubmed.ncbi.nlm.nih.gov/31892363/)]
8. Alber M, Buganza Tepole A, Cannon WR, Dura-Bernal S, Garikipati K, Karniadakis G, et al. Integrating machine learning and multiscale modeling-perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ Digit Med* 2019 Nov 25;2(1):115 [FREE Full text] [doi: [10.1038/s41746-019-0193-y](https://doi.org/10.1038/s41746-019-0193-y)] [Medline: [31799423](https://pubmed.ncbi.nlm.nih.gov/31799423/)]
9. 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):4556-4564 [FREE Full text] [doi: [10.1093/eurheartj/ehaa159](https://doi.org/10.1093/eurheartj/ehaa159)] [Medline: [32128588](https://pubmed.ncbi.nlm.nih.gov/32128588/)]
10. Gkouskou K, Vlastos I, Karkalousos P, Chaniotis D, Sanoudou D, Eliopoulos AG. The "Virtual Digital Twins" concept in precision nutrition. *Adv Nutr* 2020 Nov 16;11(6):1405-1413 [FREE Full text] [doi: [10.1093/advances/nmaa089](https://doi.org/10.1093/advances/nmaa089)] [Medline: [32770212](https://pubmed.ncbi.nlm.nih.gov/32770212/)]
11. Shamanna P, Saboo B, Damodharan S, Mohammed J, Mohamed M, Poon T, et al. Reducing HbA1c in type 2 diabetes using digital twin technology-enabled precision nutrition: a retrospective analysis. *Diabetes Ther* 2020 Nov;11(11):2703-2714 [FREE Full text] [doi: [10.1007/s13300-020-00931-w](https://doi.org/10.1007/s13300-020-00931-w)] [Medline: [32975712](https://pubmed.ncbi.nlm.nih.gov/32975712/)]
12. Chakshu NK, Carson J, Sazonov I, Nithiarasu P. A semi-active human digital twin model for detecting severity of carotid stenoses from head vibration: a coupled computational mechanics and computer vision method. *Int J Numer Method Biomed Eng* 2019 May;35(5):e3180 [FREE Full text] [doi: [10.1002/cnm.3180](https://doi.org/10.1002/cnm.3180)] [Medline: [30648344](https://pubmed.ncbi.nlm.nih.gov/30648344/)]
13. Fagherazzi G. Deep digital phenotyping and digital twins for precision health: time to dig deeper. *J Med Internet Res* 2020 Mar 03;22(3):e16770 [FREE Full text] [doi: [10.2196/16770](https://doi.org/10.2196/16770)] [Medline: [32130138](https://pubmed.ncbi.nlm.nih.gov/32130138/)]
14. WHO Handbook for Guideline Development. Geneva: World Health Organization URL: https://apps.who.int/iris/bitstream/handle/10665/75146/9789241548441_eng.pdf [accessed 2022-08-26]
15. Cosgrove L, Shaughnessy AF, Shaneyfelt T. When is a guideline not a guideline? The devil is in the details. *BMJ Evid Based Med* 2018 Feb;23(1):33-36. [doi: [10.1136/ebmed-2017-110845](https://doi.org/10.1136/ebmed-2017-110845)] [Medline: [29367325](https://pubmed.ncbi.nlm.nih.gov/29367325/)]
16. Yao X, McCoy RG, Friedman PA, Shah ND, Barry BA, Behnken EM, et al. ECG AI-Guided Screening for Low Ejection Fraction (EAGLE): rationale and design of a pragmatic cluster randomized trial. *Am Heart J* 2020 Jan;219:31-36. [doi: [10.1016/j.ahj.2019.10.007](https://doi.org/10.1016/j.ahj.2019.10.007)] [Medline: [31710842](https://pubmed.ncbi.nlm.nih.gov/31710842/)]
17. Steegers-Theunissen R, Twigt J, Pestinger V, Sinclair K. The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum Reprod Update* 2013;19(6):640-655. [doi: [10.1093/humupd/dmt041](https://doi.org/10.1093/humupd/dmt041)] [Medline: [23959022](https://pubmed.ncbi.nlm.nih.gov/23959022/)]
18. Alsnes IV, Vatten LJ, Fraser A, Bjørngaard JH, Rich-Edwards J, Romundstad PR, et al. Hypertension in pregnancy and offspring cardiovascular risk in young adulthood. *Hypertension* 2017 Apr;69(4):591-598. [doi: [10.1161/hypertensionaha.116.08414](https://doi.org/10.1161/hypertensionaha.116.08414)]
19. Maher G, O'Keefe G, Kenny L, Kearney P, Dinan T, Khashan A. Hypertensive disorders of pregnancy and risk of neurodevelopmental disorders in the offspring: a systematic review and meta-analysis protocol. *BMJ Open* 2017 Oct 05;7(10):e018313 [FREE Full text] [doi: [10.1136/bmjopen-2017-018313](https://doi.org/10.1136/bmjopen-2017-018313)] [Medline: [28982838](https://pubmed.ncbi.nlm.nih.gov/28982838/)]
20. Webster K, Fishburn S, Maresh M, Findlay S, Chappell L, Guideline Committee. Diagnosis and management of hypertension in pregnancy: summary of updated NICE guidance. *BMJ* 2019 Sep 09;366:15119. [doi: [10.1136/bmj.15119](https://doi.org/10.1136/bmj.15119)] [Medline: [31501137](https://pubmed.ncbi.nlm.nih.gov/31501137/)]
21. Mons B, Neylon C, Velterop J, Dumontier M, da Silva Santos LOB, Wilkinson MD. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Info Serv Use* 2017 Mar 07;37(1):49-56. [doi: [10.3233/isu-170824](https://doi.org/10.3233/isu-170824)]
22. Itani W, Kayssi A, Chehab A. Privacy as a service: privacy-aware data storage and processing in cloud computing architectures. 2009 Presented at: 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing; 2009; Chengdu p. 711-716. [doi: [10.1109/dasc.2009.139](https://doi.org/10.1109/dasc.2009.139)]

23. Jain A, Duin R, Mao J. Statistical pattern recognition: a review. *IEEE Trans Pattern Anal Mach Intell* 2000;22(1):4-37. [doi: [10.1109/34.824819](https://doi.org/10.1109/34.824819)]
24. Schwartz SM, Wildenhaus K, Bucher A, Byrd B. Digital twins and the emerging science of self: implications for digital health experience design and “small” data. *Front Comput Sci* 2020 Oct 15;2:1-16. [doi: [10.3389/fcomp.2020.00031](https://doi.org/10.3389/fcomp.2020.00031)]
25. Rasheed A, San O, Kvamsdal T. Digital twin: values, challenges and enablers from a modeling perspective. *IEEE Access* 2020;8:21980-22012. [doi: [10.1109/access.2020.2970143](https://doi.org/10.1109/access.2020.2970143)]
26. Richens J, Lee C, Johri S. Author correction: improving the accuracy of medical diagnosis with causal machine learning. *Nat Commun* 2021 Mar 31;12(1):1-2 [FREE Full text] [doi: [10.1038/s41467-021-21494-9](https://doi.org/10.1038/s41467-021-21494-9)] [Medline: [33790263](https://pubmed.ncbi.nlm.nih.gov/33790263/)]
27. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. New York: Springer; 2008.
28. Lee C, Yoon J, Van Der Schaar M. Dynamic-DeepHit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Trans Biomed Eng* 2020 Jan;67(1):122-133. [doi: [10.1109/tbme.2019.2909027](https://doi.org/10.1109/tbme.2019.2909027)]
29. Dempster A, Petitjean F, Webb GI. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Min Knowl Disc* 2020 Jul 13;34(5):1454-1495. [doi: [10.1007/s10618-020-00701-z](https://doi.org/10.1007/s10618-020-00701-z)]
30. Reys JM, Rijnbeek P, Cuthbert A, Ryan PB, Pratt N, Schuemie M. An empirical analysis of dealing with patients who are lost to follow-up when developing prognostic models using a cohort design. *BMC Med Inform Decis Mak* 2021 Feb 06;21(1):43 [FREE Full text] [doi: [10.1186/s12911-021-01408-x](https://doi.org/10.1186/s12911-021-01408-x)] [Medline: [33549087](https://pubmed.ncbi.nlm.nih.gov/33549087/)]
31. Wang J, Zhao P, Libol P, Hoi S. Research collection school of information systems. *J Mach Learn Res* 2014;1 [FREE Full text]
32. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. 2018. URL: https://ftp.cs.ucla.edu/pub/stat_ser/r475.pdf [accessed 2022-08-26]
33. Kyono T, Zhang Y, Van Der Schaar M. CASTLE: regularization via auxiliary causal graph discovery. URL: <http://proceedings.neurips.cc/paper/2020/file/1068bceb19323fe72b2b344ccf85c254-Paper.pdf> [accessed 2022-08-26]
34. Berrevoets J, Jordon J, Gimson A, Van Der Schaar M. OrganITE: optimal transplant donor organ offering using an individual treatment effect. URL: https://www.vanderschaar-lab.com/papers/NeurIPS2020_OrganITE.pdf [accessed 2022-08-26]
35. Leenen J, Leerenveld C, van Dijk J, van Westreenen H, Schoonhoven L, Patijn G. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. *J Med Internet Res* 2020 Jun 17;22(6):e18636 [FREE Full text] [doi: [10.2196/18636](https://doi.org/10.2196/18636)] [Medline: [32469323](https://pubmed.ncbi.nlm.nih.gov/32469323/)]
36. Rivera L, Villegas N, Jiménez M, Tamura G, Angara P, Müller H. Towards continuous monitoring in personalized healthcare through digital twins. *CASCON '19: Proc 29th Ann Int Conf Comput Sci Softw Engin* 2019:329-335.
37. Bogatu A. Cost-effective data wrangling in data lakes. 2020. URL: https://www.research.manchester.ac.uk/portal/files/162160622/FULL_TEXT.PDF [accessed 2022-08-26]
38. Fuller J. Impact of medical device integration on real-time visibility of electronic medical records. URL: <https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=11392&context=dissertations> [accessed 2022-08-26]
39. Barricelli B, Casiraghi E, Fogli D. A survey on digital twin: definitions, characteristics, applications, and design implications. 2019. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8901113> [accessed 2022-08-26]
40. Lepak D, Smith K, Taylor M. Value creation and value capture: a multilevel perspective. *Acad Manag Rev* 2007 Jan;32(1):180-194. [doi: [10.5465/amr.2007.23464011](https://doi.org/10.5465/amr.2007.23464011)]
41. Vargo S, Lusch R. Evolving to a new dominant logic for marketing. *Serv Log Mark Dialog Debate Dir* 2014;68:28. [doi: [10.4324/9781315699035-9](https://doi.org/10.4324/9781315699035-9)]
42. Lusch R, Nambisan S. Service innovation: a service-dominant logic perspective. *MIS Q* 2015 Jan 1;39(1):155-175 [FREE Full text] [doi: [10.25300/misq/2015/39.1.07](https://doi.org/10.25300/misq/2015/39.1.07)]
43. WHO classification of digital health interventions. 2018. URL: <http://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/> [accessed 2022-08-26]
44. Terrasse M, Sisti D. Policing compliance: digital medicine and criminal justice-involved persons. *Am J Bioeth* 2018 Sep;18(9):57-58 [FREE Full text] [doi: [10.1080/15265161.2018.1498937](https://doi.org/10.1080/15265161.2018.1498937)] [Medline: [30235101](https://pubmed.ncbi.nlm.nih.gov/30235101/)]
45. Scott H, Lack L, Lovato N. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. *Sleep Med Rev* 2020 Feb;49:101227. [doi: [10.1016/j.smrv.2019.101227](https://doi.org/10.1016/j.smrv.2019.101227)]
46. Slow Medicine: the Apple Watch and overdiagnosis. URL: <https://www.medpagetoday.com/opinion/slowmedicine/75946> [accessed 2021-07-21]
47. Duffy D. Problems, challenges and promises: perspectives on precision medicine. *Brief Bioinform* 2016 May;17(3):494-504. [doi: [10.1093/bib/bbv060](https://doi.org/10.1093/bib/bbv060)] [Medline: [26249224](https://pubmed.ncbi.nlm.nih.gov/26249224/)]
48. Ibrahim H, Liu X, Zariffa N, Morris A, Denniston A. Health data poverty: an assailable barrier to equitable digital health care. *Lancet Digit Health* 2021 Apr;3(4):e260-e265 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30317-4](https://doi.org/10.1016/S2589-7500(20)30317-4)] [Medline: [33678589](https://pubmed.ncbi.nlm.nih.gov/33678589/)]
49. Cornejo Müller A, Wachtler B, Lampert T. [Digital divide-social inequalities in the utilisation of digital healthcare]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2020 Feb;63(2):185-191 [FREE Full text] [doi: [10.1007/s00103-019-03081-y](https://doi.org/10.1007/s00103-019-03081-y)] [Medline: [31915863](https://pubmed.ncbi.nlm.nih.gov/31915863/)]

50. Marantz P. Blaming the victim: the negative consequence of preventive medicine. *Am J Public Health* 1990 Oct;80(10):1186-1187. [doi: [10.2105/ajph.80.10.1186](https://doi.org/10.2105/ajph.80.10.1186)] [Medline: [2400027](https://pubmed.ncbi.nlm.nih.gov/2400027/)]
51. Owens J, Cribb A. "My Fitbit thinks i can do better!" Do health promoting wearable technologies support personal autonomy? *Philos Technol* 2017 Jun 6;32(1):23-38. [doi: [10.1007/s13347-017-0266-2](https://doi.org/10.1007/s13347-017-0266-2)]
52. Prainsack B. *Personalized Medicine: Empowered Patients in the 21st Century?*. New York: NYU Press; 2017.
53. de Laat P. Algorithmic decision-making based on machine learning from big data: can transparency restore accountability? *Philos Technol* 2018;31(4):525-541 [FREE Full text] [doi: [10.1007/s13347-017-0293-z](https://doi.org/10.1007/s13347-017-0293-z)] [Medline: [30873341](https://pubmed.ncbi.nlm.nih.gov/30873341/)]
54. Nissenbaum H. Accountability in a computerized society. *Sci Eng Ethics* 1996 Mar;2(1):25-42. [doi: [10.1007/bf02639315](https://doi.org/10.1007/bf02639315)]
55. Drucker D. Never waste a good crisis: confronting reproducibility in translational research. *Cell Metab* 2016 Sep 13;24(3):348-360 [FREE Full text] [doi: [10.1016/j.cmet.2016.08.006](https://doi.org/10.1016/j.cmet.2016.08.006)] [Medline: [27626191](https://pubmed.ncbi.nlm.nih.gov/27626191/)]
56. Park J, Siden E, Zoratti M, Dron L, Harari O, Singer J, et al. Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. *Trials* 2019 Sep 18;20(1):572 [FREE Full text] [doi: [10.1186/s13063-019-3664-1](https://doi.org/10.1186/s13063-019-3664-1)] [Medline: [31533793](https://pubmed.ncbi.nlm.nih.gov/31533793/)]
57. Digital Twin. URL: <https://convergence.nl/health-technology/digital-twin/> [accessed 2022-08-31]

Abbreviations

DT: digital twin

FAIR: findable, accessible, interoperable, and reusable

Edited by A Mavragani; submitted 14.12.21; peer-reviewed by H Luo, J Heo; comments to author 18.01.22; revised version received 31.05.22; accepted 09.06.22; published 14.09.22

Please cite as:

Mulder ST, Omidvari AH, Rueten-Budde AJ, Huang PH, Kim KH, Bais B, Rousian M, Hai R, Akgun C, van Lennep JR, Willemsen S, Rijnbeek PR, Tax DMJ, Reinders M, Boersma E, Rizopoulos D, Visch V, Steegers-Theunissen R

Dynamic Digital Twin: Diagnosis, Treatment, Prediction, and Prevention of Disease During the Life Course
J Med Internet Res 2022;24(9):e35675

URL: <https://www.jmir.org/2022/9/e35675>

doi: [10.2196/35675](https://doi.org/10.2196/35675)

PMID:

©Skander Tahar Mulder, Amir-Houshang Omidvari, Anja J Rueten-Budde, Pei-Hua Huang, Ki-Hun Kim, Babette Bais, Melek Rousian, Rihan Hai, Can Akgun, Jeanine Roeters van Lennep, Sten Willemsen, Peter R Rijnbeek, David MJ Tax, Marcel Reinders, Eric Boersma, Dimitris Rizopoulos, Valentijn Visch, Régine Steegers-Theunissen. Originally published in the *Journal of Medical Internet Research* (<https://www.jmir.org>), 14.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.