Clin Chem Lab Med 2024; 62(11): 2156–2161 DE GRUYTER

Perspectives

Andrea Padoan* and Mario Plebani

Dynamic mirroring: unveiling the role of digital twins, artificial intelligence and synthetic data for personalized medicine in laboratory medicine

https://doi.org/10.1515/cclm-2024-0517
Received April 24, 2024; accepted April 30, 2024;
published online May 13, 2024

Abstract: In recent years, the integration of technological advancements and digitalization into healthcare has brought about a remarkable transformation in care delivery and patient management. Among these advancements, the concept of digital twins (DTs) has recently gained attention as a tool with substantial transformative potential in different clinical contexts. DTs are virtual representations of a physical entity (e.g., a patient or an organ) or systems (e.g., hospital wards, including laboratories), continuously updated with real-time data to mirror its real-world counterpart. DTs can be utilized to monitor and customize health care by simulating an individual's health status based on information from wearables, medical devices, diagnostic tests, and electronic health records. In addition, DTs can be used to define personalized treatment plans. In this study, we focused on some possible applications of DTs in laboratory medicine when used with AI and synthetic data obtained by generative AI. The first point discussed how biological variation (BV) application could be tailored to individuals, considering population-derived BV data on laboratory parameters and circadian or ultradian variations. Another application could be enhancing the interpretation of tumor markers in advanced cancer therapy and treatments. Furthermore, DTs applications might derive personalized reference intervals, also considering BV data or they can be used to improve test results interpretation. DT's widespread adoption in healthcare is not imminent, but it is not far off. This technology will likely offer innovative and definitive solutions for dynamically evaluating treatments and more precise diagnoses for personalized medicine.

Keywords: digital twins; artificial intelligence; personalized medicine; clinical laboratory

Introduction

The digital twins (DTs) paradigm, characterized by the creation of virtual replicas mirroring physical entities or processes, has recently gained substantial attention in several fields of informatics, technology, and healthcare. Especially in this last context, the advent of integrated digitalization and tailored artificial intelligence (AI) technologies, including advancements like natural language processing, genome sequencing, and imaging, has greatly facilitated the process of creating cornucopias of patient's health data [1]. In addition, generative AI has been proven to generate synthetic healthcare data efficiently, with clear advantages over the experimental collected data, such as the possibility of virtually completely expanding a few datasets from a single patient to a large population [2].

The concept of DT starts from the data gathered from different sources for a single person (physical entity). This requires not only results from diagnostics (including laboratory and pathology test results), but also real-time updates of personal data by wearables, and self-tracking tools capturing physical activities, heart rate, etc. [3]. Thereafter, a specific virtual replica, referred to as a digital twin, is created [4, 5]. Additionally, the creation of DTs could be facilitated by using synthetically generated data [6], particularly in cases where population data are unavailable or limited. The concept of DT in healthcare can be applied to patients, healthcare communities, and entire hospital wards; in addition, DT can be set up for specific medical devices of IVD analytical platforms. Thus, DTs are emerging as transformative tools, offering unprecedented opportunities not only for diagnosis, prognosis, and real-time monitoring of patients but also for assessing the precision and accuracy of medical devices, entire wards, or clinical laboratories.

***Corresponding author: Andrea Padoan**, Department of Medicine, University of Padova, via Giustiniani 2, Padova, Italy; Laboratory Medicine Unit, University-Hospital of Padova, via Giustiniani 2, Padova, Italy; and QI.Lab.Med, Campodarsego, Padova, Italy, Phone: 049 8212792, E-mail: andrea.padoan@unipd.it. https://orcid.org/0000-0003-1284-7885
Mario Plebani, Department of Medicine, University of Padova, Padova, Italy; and QI.Lab.Med, Padova, Italy. https://orcid.org/0000-0002-0270-1711

Currently, DTs in the field of medical disciplines are transitioning beyond their initial stages. It is interesting to note that while DTs are gaining attention in various fields, including healthcare, there is a lack of comprehensive documentation for certain disciplines, such as laboratory medicine [7]. DTs putative uses are rapidly evolving in "-omics" sciences [8], in oncology [9, 10], in radiology [11], cardiology [12, 13], and potentially in all other fields of medicine.

Interestingly, the digital replica that stands behind the digital twins concept is not a static representation, such as a digital shadow or a simulation model. DTs can be considered a dynamic, real-time simulation that mirrors its physical counterparts regarding physiological and operational aspects. Thus, simulation should also include forecasting some clinical measurement trajectories that can be predicted by varying parameters within the AI models, exploring the effect within a range of outcomes [14]. Furthermore, wearable sensors or new electronic monitoring devices have the potential to be fully integrated with DT, contributing to a more holistic approach to health monitoring [15]. Thus, digital twins might offer a window of opportunities to fix or solve the intricacies of healthcare processes, providing actionable insights for improved decision-making and patient outcomes. However, the successful application of this digital revolution requires addressing several technological issues still present, mainly due to the lack of data integration across several platforms and the management of interoperability in medical information generated during the patient's journey within the healthcare system [4, 5].

The potential role of AI and DTs in moving from population-based to patient-based laboratory medicine

The fundamental principle of the scientific method involves employing structured datasets to validate hypotheses reliably. However, it is undebatable that the recent availability of large volumes of data in medical sciences (big data), the use of AI tools, in addition to the increasing availability of genetic and molecular information have suggested the possibility of stratifying data or generating equations and algorithms following unveiled tailored schemes, for personalized treatments [16]. Recent reviews described DTs as associated with using AI for personalized medicine [7]. AI for personalized medicine and digital twins are two related but distinct concepts. AI uses algorithms to perform tasks that typically require human intelligence, such as assisted diagnosis, prognosis, and treatment recommendation. DTs are high-resolution models of individual patients updated in real-time with available data [17]. Despite DT technology being often powered by AI, DTs allow simulation and modeling of patients' results for diagnostic tests and treatments and planning predictive analyses [9]. Hence, DTs are technologies used to dynamically simulate the variation in one or several patients' characteristics (e.g. disease response), based on the real-time effect of changing of one or multiple parameters (Figure 1).

Physiological rhythms and biological variation for personalized medicine

In biological systems, rhythmic variations in the concentrations/activities of molecules can be observed over a wide range of time scales from fractions of a second to years. These variations are classified into three subgroups: ultradian, circadian, and infradian [18]. Ultradian variations represent 'short-term rhythms', with a frequency of fewer than 24 h, but typically with periods in the range of 20 min to 6 h. Circadian rhythms are daily oscillations corresponding to the earth's 24 h rotation cycle, and infradian rhythms are periods that range from days to years [18, 19]. First of all, ultradian and circadian rhythms can greatly influence the diagnosis, screening and monitoring of disease. Diurnal significant variation of albumin and TSH was demonstrated by Andersen and colleagues. For albumin, variations could be attributable to varying hydration and metabolism during the day; for TSH, the pulsatile secretions from the thyroid, the control exerted by hypothalamus-releasing hormone, as well as dopaminergic mechanisms were speculated to explain diurnal variation [20]. Andersen et al. concluded that for both albumin and TSH, the use of floating RI or 2 h intervals RI could be important for the correct clinical classification of patients [20]. Moreover, for a large series of laboratory tests, including hormones, electrolytes and leukocytes it has been demonstrated that light exposure, sleeping conditions and chronophenotype (brevidiurnal, longidiurnal) factors affect their levels during day [18]. In the context of biological variations, AI and DTs applications can be helpful for predicting the influence of circadian and ultradian on laboratory parameters, for ameliorating diagnosis, screening and patient monitoring. In particular, DTs can serve to deliver realistic simulations to physicians, predicting customized, patient-centric perspectives on the physiological fluctuations of clinical values. Furthermore,

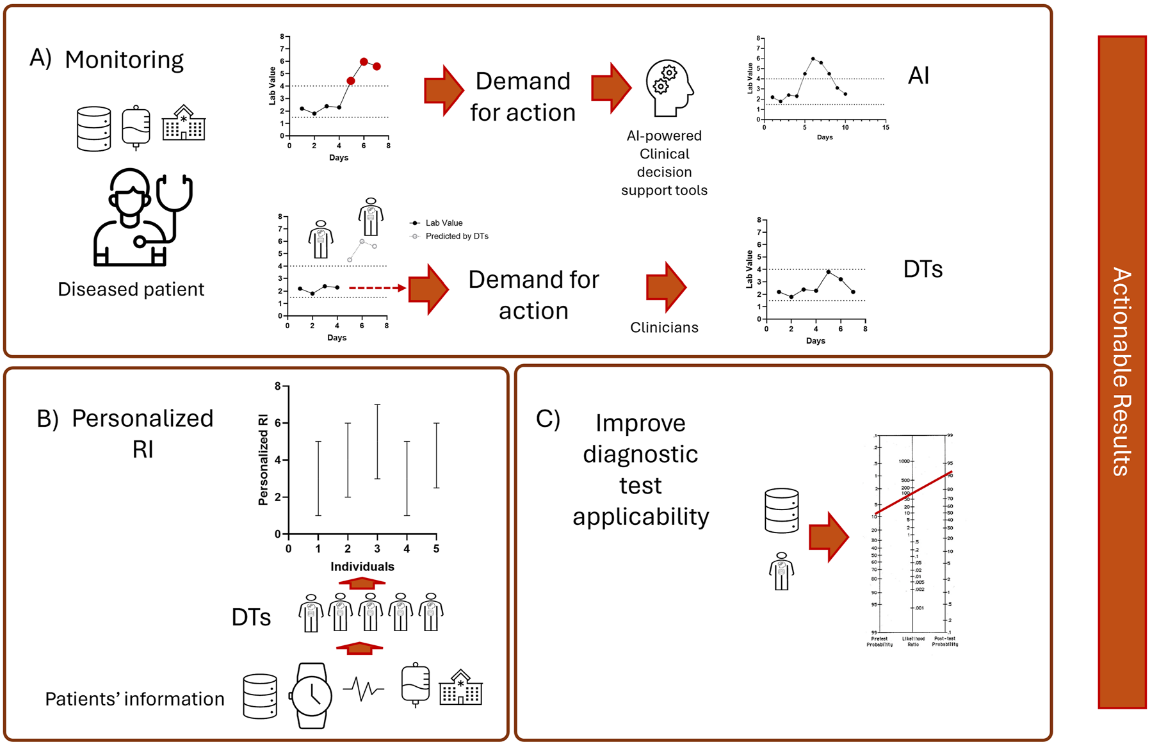

Figure 1: Possible roles of digital twins (DT) in laboratory medicine for creating clinically actionable information. (A) In contrast to AI alone, DTs might facilitate the clinical monitoring of patients and might be used for real-time prediction of worsening of clinical parameters; (B) DTs could be of utility for defining personalized RI, capturing inherent biological diversity among individuals and by using patient's data; (C) by leveraging the pre-test probability, DTs can improve diagnostic test significance.

they will prove beneficial in forecasting the anticipated changes resulting from alterations in patients' conditions, taking into account chronotypes and utilizing insights into the rhythmicity of physiological parameters.

DTs in cancer therapy monitoring: advancements for tailored treatment strategies

In several cancer types, the elevated mortality is due to the recurrence and metastasis of the tumor after surgery. The evaluation of cancer during the therapy is crucial for the effectiveness of the treatment; however, tumor heterogeneity can cause a variation in therapeutic responses, posing challenges and inefficiencies in patients' monitoring [21]. For instance, colorectal cancer (CRC) is known for its significant variability observable not only between patients but also within a single individual [22]. Patients' lifestyle and exogenous factors, together with the remarkable genomic instability of this neoplasia, might lead to an elevated spatial heterogeneity of tumor cells. This latter term refers to clones genetically distinguished between groups of tumor cells with different metastatic and invasiveness that can be detected within the same individuals and that are not often easy to be determined [22]. It has been suggested that to capture the originality of the tumor, in addition to molecular pathology, liquid biopsy techniques could represent a promising diagnostic tool, allowing the evaluation of circulating tumor cells, circulating tumor DNA, or exosomes of microRNA. Pathology and liquid biopsy could serve for determining tumor heterogeneity and evolution, particularly in light of the intra-patient variability observed in the invasiveness of different metastases within the same patient [22, 23]. By using highly accurate patients' data from imaging, molecular diagnostics, clinical biomarkers, clinical parameters, and biopsies, DTs may delineate a virtual copy of the patient's tumor status, metastasis, and treatments. These data could be used by clinicians as well as by AI tools to simulate tumor response to therapies that could be delineated with extreme precision. Hence, DTs might provide reliable predictive representations of tumor growth and differentiation (also of different metastasis), finally offering the opportunity for improvements in the surgical treatment of target metastasis.

Simulation of personalized reference intervals (pRI)

For personalized reference data, individual-specific information is needed, and this can be obtained using repeated measurements in single individuals at specified conditions. For every analyte, each individual exhibits unique fluctuations, known as within-person variation (CV_P), which highlight the inherent biological diversity among individuals [16]. Recently, it has been reported that personalized reference intervals (pRI) might be derived by biological and analytical variation and using previous test results [24]. The idea behind the pRI is that instead of using a reference population to establish reference values, steady state results of an individual (homeostatic set point) could be used, together with data from biological variation (BV), to calculate the width of possible results for that analyte [25]. pRIs have also been suggested as important in evaluating individual metabolomics/proteomic results [26]. The estimation of pRI calls for a complex mathematical approach, including evaluating data quality (analytical accuracy). From this perspective, AI and DTs might potentially forecast (among the multiple predictions of patients' information) pRI using individual and population data, particularly for the parameters with sufficient patient data to compute homeostatic set points and robust reference values(Figure 1). This presents an exceptional prospect for physicians to gain a comprehensive understanding of anticipated fluctuations in patients' parameters, applicable not just in clinical environments but also for monitoring vital function impairments in outpatient settings by general practitioners.

DTs and synthetic data for improving test results interpretation

The correct interpretation of some diagnostic test results is based on the cut-off (threshold). The probabilistic approach at the basis of the dichotomization of patient's results (namely positive or negative) can be either the optimization of sensibility/sensitivity (diagnostic test operating characteristics) or of positive/negative predictive values (PPV/NPV) [27]. However, the last two parameters depend not only on the operating characteristics of the test, but also on disease prevalence (pre-test probability). However, there is a concrete difficulty in correctly estimating the pre-test probability in most clinical contexts, and this might lead to overrating the probability of having the disease in the case of positive results (PPV) or not having the disease for negative results (NPV) [28]. Different methods can be used to estimate the pre-test probability [29]. Direct estimation of disease prevalence from population studies, or studies based on large data cohorts are mostly used. Also, the clinical reasoning process should yield a final probability of having a disease, despite it could be cumbersome [30]. Some specific tools have been speculated to be useful for achieving prevalence estimation, such as clinical prediction rules (that quantify the relative importance of clinical data points) or analysis models [31, 32]. DTs may have the potential to improve the accuracy of disease prevalence estimation by leveraging real-time data integration, finally leading to a more accurate post-test probability of disease [33] (Figure 1). In this context, synthetic data could also be used. Indeed, it has been demonstrated that synthetic patient and population health data for the state of Massachusetts could be useful for generating ML algorithms for predicting cancer risk [6]. Furthermore, the correct interpretation of diagnostic test results, also facilitated by reporting interval-specific probabilistic estimates of disease presence in laboratory reports, is expected to add actionable information to laboratory test results, enhancing interpretation of many tests, including autoimmuno and serological test result [28]. Finally, AI can finally facilitate the interpretation of clinical reports reporting interval-specific probabilistic estimates of disease.

DTs for improving automation

Total laboratory automation (TLA) could be described as systems able to prepare, sort, manage (centrifugation) and transport specimens to the laboratory analyzers to finally provide analytical results. These systems can perform several other tasks, such as measuring HIL indexes, automatically prioritizing emergency samples for reducing TAT and adequately storing samples [34]. The functions performed by TLA systems are similarly a series of rules devised by humans to manage what has been planned in preceding viewpoints. On the contrary, AI applied to TLA could autonomously adapt the flow of operations it carries out to minimize delays and recover from downtimes [35]. However, enforcing AI rules might lead to unforeseen behaviors, possibly resulting in catastrophic outcomes for laboratories. In automation, DTs serve as a "digital replica" of the program's rules and algorithms. Thus, DTs will not only enable precise determination of the behavior of advanced AI-based platforms in various scenarios but also automatically generate tests to ascertain the outcome of each scenario. Finally, every variation in predefined rules will be automatically tested for possible errors, thus avoiding the risk of human errors and unpredicted outcomes.

Final remarks

Currently, there is an incredible interest in AI applications in healthcare. A huge limitation of ML is represented by data availability, and this is primarily true for healthcare information, which includes sensitive data and requires anonymization procedures. Several strategies are being studied to solve this problem; firstly, aggregation procedures could be used to gather data from different studies. Aggregative approaches present the limitation of being preferentially applicable only to datasets collected by specific studies. In contrast, this approach could be successful for some tools (diagnostic algorithms), but it could not be feasible for real-time application when patients are continuously evaluated over time. To overcome this problem, synthetic patient data technologies have already been proven to be an alternative, efficient, and reliable source of information for testing new processes involving healthcare; recently, a tool has been developed for synthetic electronic health record generation [6]. Although population-based ML and AI algorithms offer physicians support in predicting patients' outcomes, they are generated from general data and do not consider the unique inherent complexity of each individual. The emerging technologies of DTs seem to be able to fill this gap. Enhancing the capabilities of AI to individualized data, DTs would present some peculiar features, such as the feasibility of real-time updates. In the field of clinical laboratory, there are many applications that could benefit from DTs. Some examples of potential improvements could be personalized RI, tailored BV rhythmicity and more accurate monitoring of cancer responses through biochemical markers. Further, DTs of analytical platforms (including the automation workflow) and devices, will make advantages for optimizing procedures, including stat tests execution in consolidated labs, and teaching to new personnel. However, AI and DTs' widespread adoption in healthcare in the future can be limited by data silos and lack of data integration, issues that should be solved for upscaling laboratory technological advancements [36]. All in all, while DTs widespread adoption in healthcare is not imminent, it is not far off. Shortly, this technology will likely offer innovative and definitive solutions for dynamically evaluating treatments and more precise diagnosis, thereby achieving personalized medicine.

Research ethics: Not applicable.
Informed consent: Not requested.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

1. Meijer C, Uh H-W, El Bouhaddani S. Digital twins in healthcare: methodological challenges and opportunities. J Personalized Med 2023;13:1522.
2. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. NPJ Digit Med 2023;6:186.
3. Venkatesh KP, Brito G, Kamel Boulos MN. Health digital twins in life science and health care innovation. Annu Rev Pharmacol Toxicol 2024;64:159–70.
4. Venkatesh KP, Raza MM, Kvedar JC. Health digital twins as tools for precision medicine: considerations for computation, implementation, and regulation. NPJ Digit Med 2022;5:150.
5. Katsoulakis E, Wang Q, Wu H, Shahriyari L, Fletcher R, Liu J, et al. Digital twins for health: a scoping review. NPJ Digit Med 2024;7:77.
6. Chen A, Chen DO. Simulation of a machine learning enabled learning health system for risk prediction using synthetic patient data. Sci Rep 2022;12:17917.
7. Machado TM, Berssaneti FT. Literature review of digital twin in healthcare. Heliyon 2023;9:e19390.
8. Nanni U, Ferroni P, Riondino S, Spila A, Valente MG, Del Monte G, et al. Biospecimen digital twins: moving from a "high quality" to a "fit-for-purpose" concept in the era of omics sciences. Cancer Genomics Proteomics 2023;20:211–21.
9. Moztarzadeh O, Jamshidi MB, Sargolzaei S, Jamshidi A, Baghalipour N, Malekzadeh Moghani M, et al. Metaverse and healthcare: machine learning-enabled digital twins of cancer. Bioengineering 2023;10:455.
10. Grieb N, Schmierer L, Kim HU, Strobel S, Schulz C, Meschke T, et al. A digital twin model for evidence-based clinical decision support in multiple myeloma treatment. Front Digit Health 2023;5:1324453.
11. Pesapane F, Rotili A, Penco S, Nicosia L, Cassano E. Digital twins in radiology. J Clin Med 2022;11:6554.
12. Dubs L, Charitatos V, Buoso S, Wegener S, Winklhofer S, Alkadhi H, et al. Assessment of extracranial carotid artery disease using digital twins – a pilot study. Neuroimage Clin 2023;38:103435.
13. Corral-Acero J, Margara F, Marciniak M, Rodero C, Loncaric F, Feng Y, et al. The 'digital twin' to enable the vision of precision cardiology. Eur Heart J 2020;41:4556.
14. Allen A, Siefkas A, Pellegrini E, Burdick H, Barnes G, Calvert J, et al. A digital twins machine learning model for forecasting disease progression in stroke patients. Appl Sci 2021;11:5576.
15. Celik Y, Godfrey A. Bringing it all together: wearable data fusion. NPJ Digit Med 2023;17:149.
16. Coskun A, Lippi G. Personalized laboratory medicine in the digital health era: recent developments and future challenges. Clin Chem Lab Med 2023;62:402–9.

17. Armeni P, Polat I, De Rossi LM, Diaferia L, Meregalli S, Gatti A. Digital twins in healthcare: is it the beginning of a new era of evidence-based medicine? A critical review. J Personalized Med 2022; 12:1255.
18. Coskun A, Zarepour A, Zarrabi A. Physiological rhythms and biological variation of biomolecules: the road to personalized laboratory medicine. Int J Mol Sci 2023;24:6275.
19. Abou EHM, Delvin E, Elnenaei MO, Hoffman B. Diurnal rhythm in clinical chemistry: an underrated source of variation. Crit Rev Clin Lab Sci 2018; 55:516–34.
20. Andersen IB, Brasen CL, Christensen H, Noehr-Jensen L, Nielsen DE, Brandslund I, et al. Standardised resting time prior to blood sampling and diurnal variation associated with risk of patient misclassification: results from selected biochemical components. PLoS One 2015;10: e0140475.
21. Oriuchi N, Endoh H, Kaira K. Monitoring of current cancer therapy by positron emission tomography and possible role of radiomics assessment. Int J Mol Sci 2022;23:9394.
22. Molinari C, Marisi G, Passardi A, Matteucci L, De Maio G, Ulivi P. Heterogeneity in colorectal cancer: a challenge for personalized medicine? Int J Mol Sci 2018;19:3733.
23. Darrigues L, Pierga J-Y, Bernard-Tessier A, Bièche I, Silveira AB, Michel M, et al. Circulating tumor DNA as a dynamic biomarker of response to palbociclib and fulvestrant in metastatic breast cancer patients. Breast Cancer Res 2021;23:31.
24. Coskun A, Sandberg S, Unsal I, Yavuz FG, Cavusoglu C, Serteser M, et al. Personalized reference intervals – statistical approaches and considerations. Clin Chem Lab Med 2021;60:629–63.
25. Coskun A, Sandberg S, Unsal I, Serteser M, Aarsand AK. Personalized reference intervals: from theory to practice. Crit Rev Clin Lab Sci 2022; 59:501–16.
26. Pusparum M, Ertaylan G, Thas O. Individual reference intervals for personalised interpretation of clinical and metabolomics measurements. J Biomed Inf 2022;131:104111.
27. Kost GJ. Diagnostic strategies for endemic coronavirus disease 2019 (COVID-19). Arch Pathol Lab Med 2022;146:16–25.
28. Bossuyt X, Van Hoovels L. Reporting likelihood ratio in autoimmune serology: just a theory or feasible in practice? Clin Chem 2024;70:482–5.
29. Uy EJB. Key concepts in clinical epidemiology: estimating pre-test probability. J Clin Epidemiol 2022;144:198–202.
30. Moreira J, Bisoffi Z, Narváez A, Van den Ende J. Bayesian clinical reasoning: does intuitive estimation of likelihood ratios on an ordinal scale outperform estimation of sensitivities and specificities? J Eval Clin Pract 2008;14:934–40.
31. Novikov I, Olmer L, Keinan-Boker L, Silverman B, Robinson E, Freedman LS. A modified prevalence incidence analysis model method may improve disease prevalence prediction. J Clin Epidemiol 2020;123: 18–26.
32. Akushevich I, Yashkin A, Kravchenko J, Fang F, Arbeev K, Sloan F, et al. A forecasting model of disease prevalence based on the McKendrick–von Foerster equation. Math Biosci 2019;311:31–8.
33. Vallée A. Digital twin for healthcare systems. Front Digit Health 2023;5: 1253050.
34. Yu H-YE, Lanzoni H, Steffen T, Derr W, Cannon K, Contreras J, et al. Improving laboratory processes with total laboratory automation. Lab Med 2019;50:96–102.
35. Ialongo C, Bernardini S. Total laboratory automation has the potential to be the field of application of artificial intelligence: the cyber-physical system and “Automation 4.0”. Clin Chem Lab Med 2019;57:e279–81.
36. Messiou C, Lee R, Salto-Tellez M. Multimodal analysis and the oncology patient: creating a hospital system for integrated diagnostics and discovery. Comput Struct Biotechnol J 2023;21:4536–9.