Journal of Pathology Informatics 14 (2023) 100313

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Short Communication

Digital twin concept: Healthcare, education, research

Maria Peshkova [a,b,*], Valentina Yumasheva [c], Ekaterina Rudenko [d], Natalia Kretova [a,d], Peter Timashev [a,b,e], Tatiana Demura [d]

[a] World-Class Research Center "Digital Biodesign and Personalized Healthcare", Sechenov University, 119991 Moscow, Russia
[b] Institute for Regenerative Medicine, Sechenov University, 119991 Moscow, Russia
[c] Sechenov University, 119991 Moscow, Russia
[d] Institute for Clinical Morphology and Digital Pathology, Sechenov University, 119991 Moscow, Russia
[e] Chemistry Department, Lomonosov Moscow State University, 119991 Moscow, Russia

ARTICLE INFO

Keywords:
Digital twins
Databases
Digital pathology
Biobanking

ABSTRACT

Introducing the concept of digital twins in healthcare, medical education, and research is a complex multistage challenge requiring participation of multidisciplinary teams. In pursuing this goal, we have created a validated database of scans of colorectal tumor slides associated with relevant clinical and histological information. This database is also linked to the blood bank, which opens a wide range of opportunities for further research. Herein, we present our experience within the scope of the digital twins initiative.

Introduction

A digital twin is referred to as a replica of a real object or phenomenon.[1] This term was first introduced in the early 1990s,[2] while the general concept was developed in 2002.[3] It was described as data transmission from real equipment to its virtual twin resulting in a high-precision virtual model that can imitate a real object's behavior. At present, digital twins are widely used in industry, for instance, by NASA, General Electric, Siemens, ANSYS, Dassault, *etc.*[4] There is an emerging use of digital twins in education, especially since the COVID-19 pandemic. For example, this technology was successfully applied in training architects[5] and engineers.[6]

In medicine, though, despite the wide use of digital technologies and artificial intelligence,[7–9] the application of the digital twins' concept is still in its infancy, not only in education but also in healthcare and research. This can be explained by the utmost complexity of biological systems, which should be represented at multiple levels, including molecular, cellular, and tissular ones,[8] to be fully mimicked. While still a challenge, elaboration of digital twins in medicine starts with the first steps, such as consolidation of data between biobanks and creation of databases.

Technique description

To ensure the selection of highly informative samples for the database, 3 pathologists, a biologist, and a laboratory technician perform slide evaluation and scanning. Slides of colorectal cancer (adenocarcinoma) fragments (Fig. 1A), as well as of regional and distant metastases (Fig. 1B) are selected. The slides are scanned using a Leica Aperio AT2 slide scanner with an x40 objective. The digital copies of the slide are then marked by a specialist in QuPath software into several classes: Tumor; Intramural_vascular_invasion; Extramural_vascular_invasion; Perineural_invasion; Tumor_budding; Tumor_infiltrating_lymphocytes; Lymphoid_follicles.

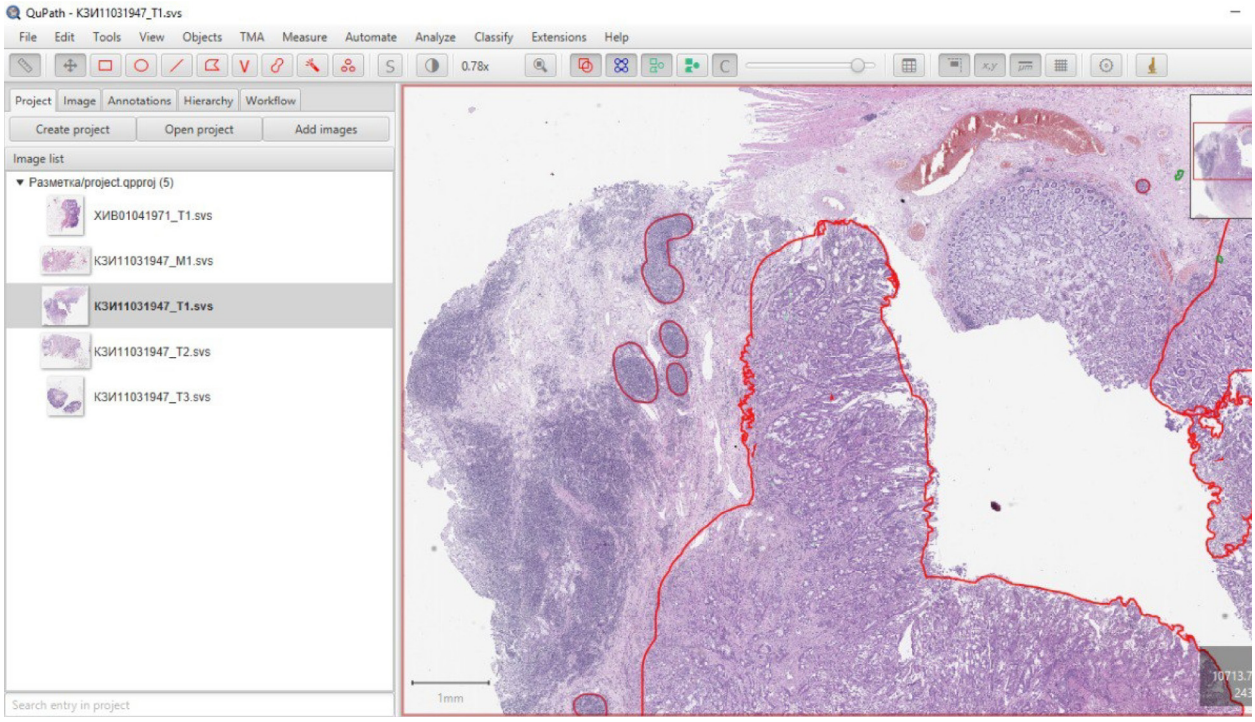
Relevant clinical and histological information is entered into the table based on Microsoft Excel Spreadsheet Software, making it compatible with spreadsheet editors of other software developers on Windows, Linux, and MacOS operating systems. All queries, filtering, and data selection are implemented using standard Microsoft Excel functions.

The table contains information about patient's sex, age, codes of associated histological scans, number of scans, tumor localization and staging, comorbidities, ICD-code, tumor dimensions (mm), circular tumor growth, ulcerations, foci of necrosis, degrees of differentiation and malignancy, invasion (including fat, perineural, blood, and lymphatic vessels invasion), lymphatic nodes with and without metastases, metastases dimensions, resection margins, signet-ring cells, tumor response and budding, number of lymphoid follicles and infiltrating lymphocytes, mutations in APC, TP53, KRAS, NRAS, and BRAF genes, microsatellite instability, chromosomal instability, CpG island methylator phenotype, consensus molecular subgroups, TS, TP, DPD, MSH2, MLH1, PMS2, MSH6, CD133, CD44, CD166, ALDH1, TGFβ, EGFR, CD34, VEGF, CD3, CD4, CD8, CD68, and p53 expression.

* Corresponding author.
E-mail address: peshkova_m_a@staff.sechenov.ru (M. Peshkova).

http://dx.doi.org/10.1016/j.jpi.2023.100313
Received 1 March 2023; Received in revised form 11 April 2023; Accepted 12 April 2023
Available online 16 April 2023

A



B

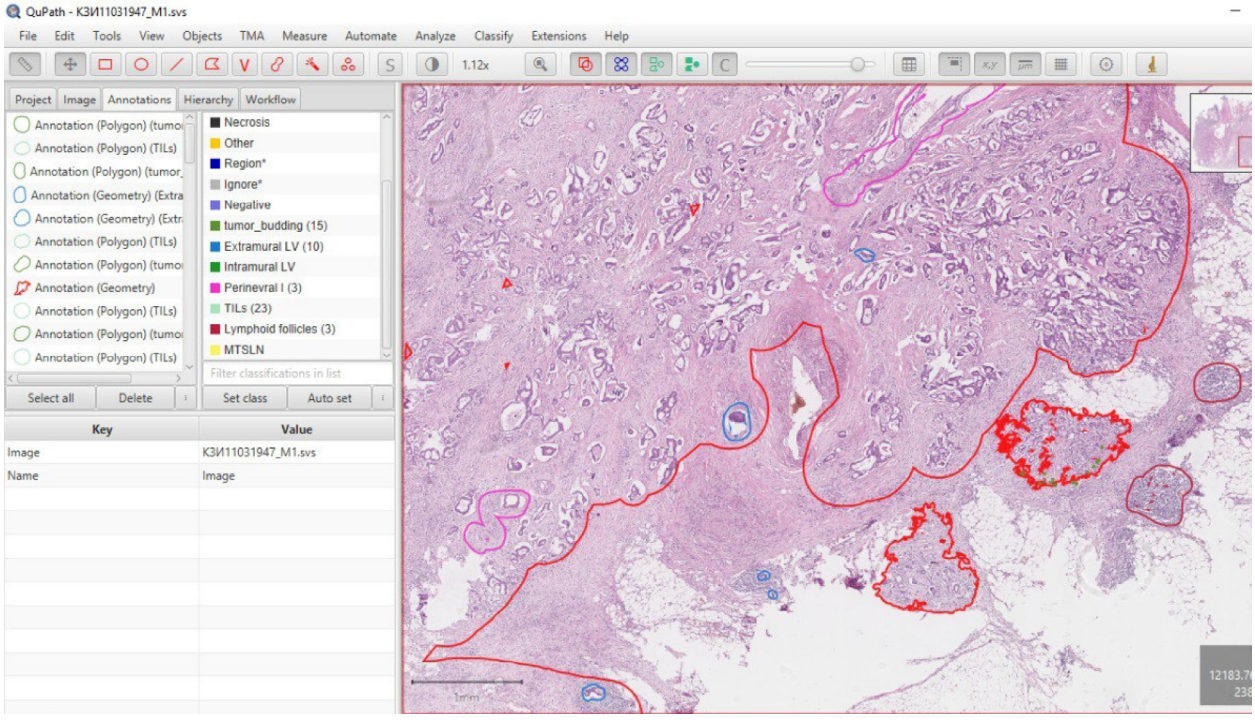


Fig. 1. Markup of a histological slide scan via QuPath software.

Results

As a result, a database of colorectal cancer cases was created. The database consists of 3 main blocks:

1. Scans of histological slides that are marked according to key and prognostically significant morphological patterns using the QuPath software.
2. Anonymized clinical data along with morphological and immunohistochemical tumor characteristics (consists of several sections integrated into a single table and is linked to the data from the blood bank).
3. Reference books and classifiers (contains data for checking values and substituting data into the tables of the previous block).

The database contains cases of colorectal cancer of various localizations (Table 1).

The algorithms developed for creating this database can be applied for the creation of databases for other projects of the digital twins initiative. The parameters chosen to create this database can be optimized considering other pathologic conditions' specifics.

Discussion

Colorectal cancer is the world's third most common cause of cancer death.[10,11] Classical TNM staging corresponding to the primary tumor invasion degree (T), lymph node status (N), and distant metastases (M)[12] is used to characterize a tumor, and further treatment strategy is based on this information. Machine learning methods are increasingly used for histological and immunohistochemical slides analysis,[13] helping to ease the burden on pathologists, decreasing medical error probability and improving overall accuracy and productivity. Still, according to some estimations, more than 50% of cases of colorectal cancer are detected at stages III–IV.[14]

Moreover, due to cancer heterogeneity, there is an unmet need for highly personalized treatment strategies, as well as for the balance between the treatment efficacy and safety. These gaps in diagnostics and treatment urge the need for digital twins development.

Developing complex digital models based on biobanks of tumor tissues and biological fluids of patients representing the disease dynamics could contribute to the improvement of advanced diagnostics and more effective screening of colorectal cancer. Moreover, such digital twins could be used in medical education, helping to develop competencies in personalized healthcare.

International experience broadens the scope for digital twins creation. For example, a digital platform for kidney preimplantation biopsies proposed by Neri and colleagues aims to increase the precision of organ assessment and its predictive value for transplant outcomes.[15] The Archipelago of Ovarian Cancer Research (AOCR), an initiative of Dutch biobanks, promotes fundamental and translational research on ovarian cancer, accumulating blood and tissue samples, clinical and pathological data (including digitalized slides), and encouraging researchers to enrich the biobank with data resulting from their experiments.[16] International cooperation between biobanks could help accumulate large amounts of data faster, thus boosting the creation of digital twins for various pathologies. To facilitate data consolidation, biobanks should follow common standard operating protocols and are encouraged to promote transparency and responsible use of biospecimen.[17]

However, the widespread introduction of digital twins in healthcare, medical education, and research is currently limited by the long and costly development process.[18] Furthermore, some associated ethical issues, including personal data management, need to be addressed.[4,19,20] Finally, the use of digital twins in medical education is sometimes frowned upon due to the fear of supplanting traditional bedside teaching. However, the proposed digital twin concept rather aims to enrich the existing approaches and, in some years, may become an indispensable component of healthcare, education, and research.

Table 1
ICD codes for cases included in the database.

№	ICD
1)	C18.7 Malignant neoplasm: Sigmoid colon
2)	C20 Malignant neoplasm: Rectum
3)	C19 Malignant neoplasm: Rectosigmoid junction
4)	C18.6 Malignant neoplasm: Descending colon
5)	C18.0 Malignant neoplasm: Caecum
6)	C18.2 Malignant neoplasm: Ascending colon
7)	C18.4 Malignant neoplasm: Transverse colon
8)	C18.5 Malignant neoplasm: Splenic flexure
9)	C21.1 Malignant neoplasm: Anal canal
10)	C18.8 Malignant neoplasm: Overlapping lesion of colon
11)	C18.3 Malignant neoplasm: Hepatic flexure

Authors' contribution

M.P. and V.Y. drafted the manuscript with primary editing and revision support from E.R. and N.K.; P.T. and T.D. critically revised the manuscript and coordinated the manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding

This work was financed by the Ministry of Science and Higher Education of the Russian Federation within the framework of state support for the creation and development of World-Class Research Centers "Digital biodesign and personalized healthcare" № 075-15-2022-304.

Conflict of interest

The authors declare no conflict of interest.

References

1. Vikhman V, Romm M. "Digital twins" in education: prospects and reality. Higher Educ Russia 2021;2:22–32. https://doi.org/10.31992/0869-3617-2021-30-2-22-32.
2. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. J Pers Med 2021;8:745. https://doi.org/10.3390/jpm11080745.
3. Grieves MW. Product lifecycle management: the new paradigm for enterprises. Int J Prod Dev 2005;2:71–84. https://doi.org/10.1504/IJPD.2005.006669.
4. Madaliev A, Ivanov VM. Addictive manufacturing and digital twins: from industry to medicine. Health Basis Human Potential Probl Ways Solve Them 2019;14(1):229–234.
5. Sepasgozar, Samad ME. Digital twin and web-based virtual gaming technologies for online education: a case of construction management and engineering. Appl Sci 2020;10 (13):4678. https://doi.org/10.3390/app10134678.
6. Liljaniemi A, Paavilainen H. Using digital twin technology in engineering education – course concept to explore benefits and barriers. Open Eng 2020;10:1377–1385. https://doi.org/10.1515/eng-2020-0040.
7. Williams BJ, Lee J, Oien KA, Treanor D. Digital pathology access and usage in the UK: results from a national survey on behalf of the National Cancer Research Institute's CM-Path initiative. J Clin Pathol 2018;71(5):463–466. https://doi.org/10.1136/jclinpath-2017-204808.
8. He X, Wang L, Wang L, et al. Effectiveness of a cloud-based telepathology system in china: large-sample observational study. J Med Internet Res 2021;23(7), e23799. https://doi.org/10.2196/23799.
9. Evgenievna RE, Alexandrovna DT, Andreevna VK, et al. Analysis of the three-year work of a digital pathomorphological laboratory built from the ground. J Pathol Inform 2022;13, 100111. https://doi.org/10.1016/j.jpi.2022.100111.
10. Langner C, Magro F, Driessen A, et al. European Society of Pathology; European Crohn's and Colitis Foundation. The histopathological approach to inflammatory bowel disease: a practice guide. Virchows Arch 2014;464(5):511–527. https://doi.org/10.1007/s00428-014-1543-4.
11. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136(5): E359–E386. https://doi.org/10.1002/ijc.29210.
12. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin 2017;67(2):93–99. https://doi.org/10.3322/caac.21388.
13. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. Med Image Anal 2016;33:170–175. https://doi.org/10.1016/j.media.2016.06.037.
14. McGarvey N, Gitlin M, Fadli E, Chung KC. Increased healthcare costs by later stage cancer diagnosis. BMC Health Serv Res 2022;22(1):1155. https://doi.org/10.1186/s12913-022-08457-6.

15. Neri F, Eccher A, Rigotti P, et al. Advantages of using a web-based digital platform for kidney preimplantation biopsies. J Pathol Inform 2021;12:41. https://doi.org/10.4103/jpi.jpi_23_21.
16. Zelisse HS, van Gent MDJM, de Ridder S, et al. Archipelago of Ovarian Cancer Research Consortium. Establishment of the Dutch nationwide, interdisciplinary infrastructure and biobank for fundamental and translational ovarian cancer research: archipelago of ovarian cancer research. Gynecol Obstet Investig 2022;87(6):389–397. https://doi.org/10.1159/000528330.
17. Bonizzi G, Zattoni L, Capra M, et al. Standard operating procedures for biobank in oncology. Front Mol Biosci 2022;9, 967310. https://doi.org/10.3389/fmolb.2022.967310.
18. Baxi V, Edwards R, Montalto M, et al. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol 2022;35:23–32. https://doi.org/10.1038/s41379-021-00919-2.
19. Maddahi A, Chen S. Applications of digital twins in the healthcare industry: case review of an IoT-enabled remote technology in dentistry. Virtual Worlds 2022;1:20–41. https://doi.org/10.3390/virtualworlds1010003.
20. Braun M. Represent me: please! Towards an ethics of digital twins in medicine. J Med Ethics. medethics-2020-106134. https://doi.org/10.1136/medethics-2020-106134.