A new consortium of the EU Innovative Medicines Initiative (IMI) will establish the biggest database of pathology images to accelerate the development of artificial intelligence in medicine.

Nijmegen, 1 February 2021 - To take AI development in pathology to the next level, a European consortium combining leading European research centres, hospitals as well as major pharmaceutical industries, is going to develop a repository for the sharing of pathology data.
The 6-year, €70 million project called BIGPICTURE, will herald a new era in pathology.

Pathology is the cornerstone of the workup of many diseases such as cancer, autoimmune diseases, of the follow up after transplantation and is also critical for the evaluation of the safety of drugs. It's based on the examination of tissue samples (slides) under the microscope. However, despite its pivotal role, it still relies heavily on the qualitative interpretation by a qualified pathologist.

While the microscope symbolizes the profession, the digitalisation of slides in recent years ignited a revolution: not only images can now be shared and accessed from distant locations, they can also be processed by computers. This opens the door for artificial intelligence (AI) applications to assist the pathologist and help study diseases, find better treatments and contribute to the 3Rs (replace, reduce, and refine animal use in research). However, the development of robust AI applications requires large amounts of data, which in the case of pathology means a huge collection of digital slides and the medical data necessary for their interpretation. Sharing these has so far remained challenging due to the data storage capacity required to host a sufficiently large collection and to concerns regarding the confidential character of the medical information.

To allow the fast development of AI in pathology, the BIGPICTURE project aims to create the first European, ethical and GDPR-compliant (General Data Protection Regulation), quality-controlled platform, in which both large-scale data and AI algorithms will coexist. The BIGPICTURE platform will be developed in a sustainable and inclusive way by connecting communities of pathologists, researchers, AI developers, patients, and industry parties.

The project is divided into four main aspects that concern the large-scale collection of data. First, an infrastructure (hardware and software) must be created to store, share and process millions of images that can be gigabytes each. Second, legal and ethical constraints must be put in place to ensure adequate usage of data while fully respecting patient's privacy and data confidentiality. Then, an initial set of 3 million digital slides from humans and laboratory animals will be collected and stored into the repository to provide data for the development of pathology AI tools. Finally, functionalities that aid the use of the database as well as the processing of images for diagnostic and research purposes will be developed.

BIGPICTURE is a public-private partnership funded by IMI, with representation from academic institutions, small- and medium-sized enterprises (SMEs), public organisations and pharmaceutical companies, together with a large network slide contributing partners. The consortium partners involved in the project are:

Academic institutions: Radboud University Medical Center (NL), Linköping University (SE), Leeds Teaching Hospitals NHS Trust (UK), University Medical Centre Utrecht (NL), Uppsala University (SE, ELIXIR node), Haute Ecole Spécialisé de Suisse Occidentale (CH), Eindhoven University of Technology (NL), University of Warwick (UK), Technical University of Munich (DE), Medical University Graz (AT), Institut Pasteur (FR), University of Liege (BE), University of Semmelweis (HU), National Cancer Institute (NL), Region Östergötland (SE), Medical University Vienna (AT), University of Marburg (DE), Helsingin ja Uudenmaan sairaanhoitopiirin kuntayhtymä (FI),

Pharmaceutical companies: Novartis Pharma AG (CH), Janssen Pharmaceutica NV (BE), Bayer AG (DE), Boehringer Ingelheim International GmbH (DE), Novo Nordisk A/S (DK), Pfizer Ltd (GB), Roche (CH), Sanofi Aventis recherche et Développement (FR), Institut de Recherches Internationales Servier (FR), and UCB Biopharma SRL (BE).

Other public & private organisations: CSC – IT Center for Science Finland (FI, ELIXIR node), Biobanks and biomolecular resources research infrastructure (AT), Azienda Ospedaliera Per L'Emergenza Cannizzaro (IT), Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.(DE), Deutsches Institut für Normung E.V. (DE), European Institute for Innovation through Health Data (BE), European Society of Pathology (BE), Digital pathology association (US), GBG Forschungs Gmbh (DE), ttopstart (NL), Sectra AB (SE), Cytomine SCRLFS (BE), Stichting Lygature (NL), Owkin (FR), Deciphex (IE), MedicalPhit (NL), Timelex (BE),

BIGPICTURE starts on 1st February 2021 and will run for 6 years. However, the platform is meant to last, and the consortium will elaborate sustainability plans to maintain and continue to develop the platform beyond this term.

Owkin will participate in the BIGPICTURE consortium through developing, alongside other members, AI models to unveil signatures from WSIs. This will include model training geared towards predicting outcomes such as genetic mutations, treatment

response, recurrence, survival and more. Furthermore, Owkin will engage in efforts to make these models interpretable—or make known the features of an image that trigger the model's prediction. Because of the scale of the datasets involved, and the GDPR regulations on sensitive data access, the consortium will also leverage Owkin Federated Learning technologies to enable the collaborative development of these AI models.

Acknowledgement of support and disclaimer