Guide to fast-track AI-enhanced clinical studies for viral pandemics

Lessons learned from COVID-19

The DRAGON consortium

Guide to fast-track AI-enhanced clinical studies for viral pandemics

Introduction

Topic 1 Using clinical data
- How is data collected (when a pandemic hits)?

Topic 2 People involved
- Patient perspective
- Clinician perspective
- Researcher perspective

Topic 3 Regulations and standardizations
- Ethics
- Regulations & Standardization

Lessons learned

Introduction

The SARS-CoV-2 (COVID-19) virus was first detected in Wuhan, China in December 2019. By March 11th, 2020, the World Health Organization (WHO) had declared COVID-19 a global pandemic and the virus has since claimed over 6.5 million lives (1). COVID-19 differs from previous pandemics in the sense that it is the first pandemic to take place in the interconnected modern world. This had several implications for its development and management. First, global travel meant that the virus could spread faster and wider than any other virus in history (2). Furthermore, in the age of the internet and social media, in which information also travels at a rapid rate, false and misleading information among the general public undermined measures taken by public health authorities to control the unfolding pandemic.

On the other hand, this global interconnectedness also offered tremendous opportunities to tackle COVID-19. For the first time in history, the large volumes of clinical data generated and stored in healthcare systems around the world could potentially be extracted and analyzed through artificial intelligence (AI) and machine learning (ML) methods. These methods could be used to answer critical questions, such as: which treatment protocol could offer the highest survival probability for an individual patient, how healthcare systems' capacity could be efficiently leveraged for testing and treatment, and how clinical trials could be conducted to ensure patient safety and treatment effectiveness when in-person visits are limited?

Since future viral pandemics may take place in a possibly more interconnected world, it is crucial that the global community is able to respond in a coordinated manner. A responsive healthcare system is required that can perform accurate and rapid diagnoses and predict which treatment is likely to benefit the individual patient. Such a system requires a coherent data infrastructure and global coordination in order to deploy clinical resources in the most efficient manner possible and avoid overloading the healthcare system. It is also crucial to have participation from an informed public.

In current clinical practice, anamneses, examinations, diagnoses, and treatment are performed based on the information the patient shares with the clinician and their own expert knowledge. The exchange of information is primarily between patient and clinician in a clinical setting, as illustrated in Figure 1. In order to improve current care, the clinician may share information/data collected from their patients with researchers who are increasingly using AI/ML techniques on these large volumes of data to generate models that can aid in diagnoses or predict an individual patient's response to a certain drug or treatment. It is important to note that while the researcher has the technical expertise to develop and interpret the model, there needs to be an informational exchange between the clinician and the researcher in which the clinician can communicate to the researcher the challenges they face in the clinic so that the researcher can develop an AI model and train it, using the data provided by the clinician, which can be interpreted by the researcher and communicated to the clinician.

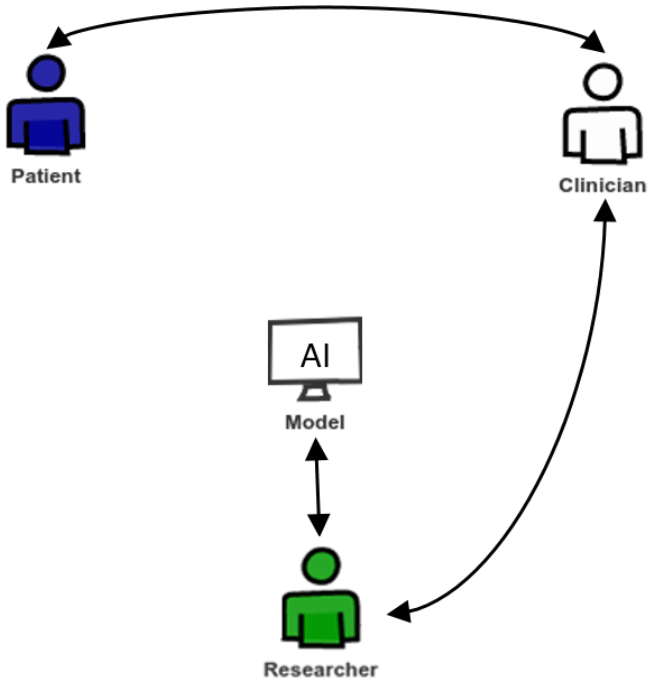

Figure 1. Current situation

The situation described above is not optimal as it places a considerable burden on the clinician to give data to the researcher or the AI model. Given the fact clinicians experience a huge workload already, optimizing the process of collecting high-quality data, using it to train and validate an AI model, and deploying it in the clinic would benefit patients, clinicians, research institutes, and the wider healthcare system.

The DRAGON project is an international collaboration between universities, research institutes, patient-led organizations, companies, and non-profit groups to develop an AI/ML-based clinical decision support system to diagnose COVID-19 and predict its outcomes for individual patients. In order to learn from multiple datasets stored in various hospitals and institutes throughout Europe, the project used a federated learning system in which AI/ML algorithms could be trained by sending them to different institutes and improving continuously as new data became available. Patient privacy was preserved by ensuring that patient data did not need to leave the institute. The successful deployment of such a federated learning system depends on:

- Collected and curation of the most relevant clinical data
- Successful engagement of multiple stakeholders, not only clinicians, researchers, patients, and model developers but also members of the general public
- Regulations and standardization to facilitate cross-country collaboration.

The aim of this guide is to outline the main challenges faced by the DRAGON consortium in these three areas and to distill the lessons learned into possible strategies that can be used in future (viral) pandemics.

Topic 1 - Using clinical data

Clinical research during a pandemic relies on rapid data collection and sharing of findings to build an evidence base for informing and implementing a fast response. Different data collection methods are currently used in clinical research: questionnaire surveys, patient self-reported data, proxy/informant data, hospital/ambulatory medical records, and biological samples. When a pandemic hits, the first critical step is to determine which variables are most important, keeping in mind that they might evolve during the course of the study. Healthcare workers then need to learn and implement new research protocols, which can be challenging as they already juggle new clinical practice guidelines and increased workload. The major pitfalls in data collection at the start of a pandemic are, therefore:

1. The absence of a research manual containing a precise description of the variables to collect and of the procedure to follow;
2. The difficulty in adequately training data collectors;
3. Time constraints to gather, systematically and reproducibly, high-quality, complete data.

The COVID-19 pandemic helped develop alternative and innovative data collection methods from remote electronic reporting to improve efficiency, speed, and accuracy. This includes different types of data; within DRAGON the focus was mostly on clinical data and images (CT scans).

Clinical data

During public health emergencies, clinical data are key elements to assess the impact of new treatments, track trends in the quality of care, and identify the needs of the most vulnerable. In a later phase of the pandemic, clinical data help researchers and policymakers understand longstanding issues and shortcomings of national healthcare systems. Clinical data can be used for local studies, but the COVID-19 pandemic taught us that broader studies were more informative, emphasizing how data access and re-use are critical for ensuring the efficiency and effectiveness of the global scientific research effort to address pandemics. Of note, initiatives are underway in Europe and other regions to develop Open Science Clouds, an infrastructure providing data, analytical tools, and information about computation, security, and data storage.

Imaging data

Medical imaging processing can help in the diagnosis, prognosis, treatment, and early detection of diseases. This technology has been of major importance in the fight against COVID-19. Several artificial intelligence, machine learning, and deep

learning techniques have been deployed in medical image processing in the context of COVID-19 disease, leading to various applications to predict the outcome and survival of infected patients. Given the power of medical imaging to quickly and non-invasively evaluate anatomy and physiological processes, medical imaging will most likely play a crucial role in any future public health emergencies, including pandemics.

Biological specimens

Biological specimens, such as blood, saliva, and feces, can help twofold. On one hand, it can provide information related to the variety of the virus. On the other hand, it can provide information about the individual patient's genetics. Knowing such information can unravel a lot about a disease and will help develop better treatments and avoid further spread.

Combining clinical data, imaging data, and biological data the right way could help develop personalized approaches, also known as precision medicine. To do so, it is important to collect all data in a standardized way, both within and between clinics. When data is not standardized from the beginning, mapping exercises (both syntactic and semantic) must be done to combine the disparate data effectively. This takes much longer and requires more effort than if the same structure was used initially.

DRAGON experiences

When the COVID-19 pandemic hit, hospital workers did what they could and treated patients as written in their regular protocols. When the world started to know more about COVID it became clear which variables were most important. It may be clear that due to the lack of knowledge and the overburdened staff, data was not collected in a very standardized way, and in some instances, it was not collected at all. Nevertheless, the partners involved in clinical trials conducted as part of DRAGON were all very motivated to get the best out of it. Hours were spent extracting the right variables and CT scans from the hospital databases. Extra materials such as saliva and breath aerosols were collected as part of the clinical trials, in multiple centers across Europe. Besides data collected within hospitals also patients who had COVID but recovered at home were included. This resulted in yet another way of collecting data and yet another way of having to align it with the rest of the database.

Typically in research, data is collected for a single trial and a single purpose. This lack of global data harmonization and data standards adoption impedes realizing innovative and efficient research from multiple sources of disparate data. A data standard is (1) a set of defined data elements, their characteristics, and the relationship among them, (2) rules for creating, managing, and using the data elements, and (3) a design that enables consistent collection and representation of these elements. Data standards enable combining data meaningfully when they have been collected in many places and many ways. This enables different organizations to share data within clinics and between clinics.

The Clinical Data Interchange Standards Consortium (CDISC), is one of the partners within DRAGON and defines standards to deal with medical research data linked with life sciences, to "enable information interoperability to improve medical research and related areas of life sciences". As mentioned, most of the data were collected on the go and needed to be mapped according to CDISC standards in hindsight.

Suppose the opportunity arises where a health system can set up the mapping and refresh the clinical data. In that case, once the mapping has been completed to the CDISC standards structure, these tables can be refreshed to provide information for continuous data and ongoing surveillance to support both public health use cases and clinical trials. Having a process whereby the data is refreshed will prepare for fast-tracking clinical trials in future pandemics. The data would be able to inform science from the beginning and enable reuse beyond one clinical trial thus enabling a responsive healthcare system.

Topic 2 - Engaging key stakeholders

Once the relevant clinical data has been identified and collected in a standardized format, it needs to be used to develop accurate and explainable AI models to answer the relevant clinical questions that arise during a viral pandemic. AI models can be improved by going beyond data collected during standard clinical visits and gathering data directly from patients, such as patient-reported symptoms. An example of this is the Zoe COVID Symptom Study in the UK which used a publicly available app to let individuals track their symptoms and was instrumental in detecting key predictors of COVID-19 like loss of taste and smell (anosmia) (3). The development and release of these digital tools for the wider public can enable citizens to contribute valuable data, participate in its analysis and interpretation, and even inform research priorities. Therefore, to get the best out of clinical data, it is critical to engage key stakeholders both within and outside the clinic, namely clinicians, researchers, and citizens throughout the model development and deployment process.

Development of AI tools

In contrast to current clinical practice described previously in Figure 1, an optimized process for the development stage of an AI model would involve interactions between clinicians, researchers, patients, and systems to capture and process the relevant data (Figure 2). In the proposed scenario, when the clinician and the researcher agree on the new data-driven decision support system, the researcher will start developing the model based on an ongoing exchange with the clinician about what data is required (clinical data, images, lab work, biological specimens, et cetera) and also how this data will be collected (automatically derived from existing electronic journals, manual input).

Simultaneously, relevant data will also be collected from patients or even members of the general public who are infected but not admitted to the hospital. This will ensure the data collection will be richer as it also includes participants who are not

seen by a clinician but, in the case of DRAGON, did have COVID. For this reason, we include patients during the developmental phase as well, along with a publicly available tool, such as a mobile app or a website, that they can use to capture the data of interest. It is important to make the tool as user-friendly as possible, which can only be done when the patients get to test the tool and give their feedback to the researcher and tool developers.

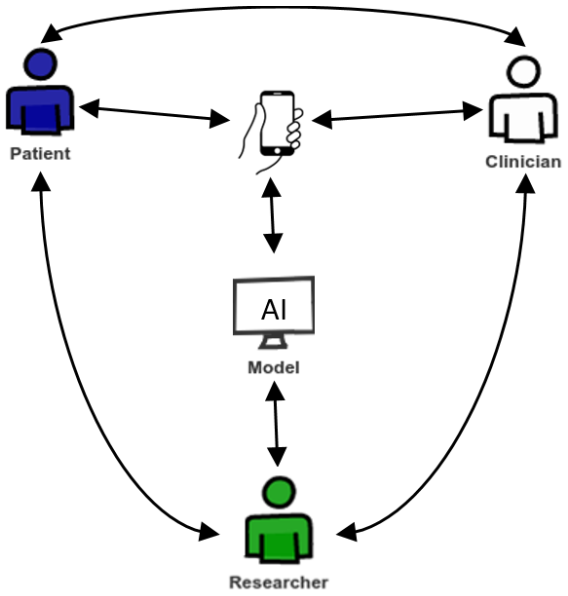

Figure 2. Development phase

The feedback from both the patient and the clinician and also from the researcher will be used to fine-tune the tool. This tool need not necessarily include the AI model but is used to collect data in a standardized way. This will save time for the curation process later on and makes sure less information will get lost. The researcher will be able to extract the data that was gathered through the device and develop the AI model with it.

Clinician-researcher perspective

Successful development of data-driven decision support tools relies on collaboration between clinicians and researchers. Failing to integrate the perspectives of clinicians as end-users is one of the most significant barriers to AI model implementation (4). It is crucial to involve clinicians in model design, development, and testing in an iterative manner. One method is a user-centered design process in which key end-users are consulted throughout the development and testing of a model or prototype (5). The main advantage of this approach is that it provides developers with a deeper understanding of the organizational and behavioral factors that influence the acceptance of new tools, besides the technical requirements. This results in tools that are more effective, safer, easier to integrate into existing organizational structures, and more acceptable to users.

A second challenge, aside from understanding and integrating clinician workflows, is gaining clinicians' trust in AI. In the research community, there has been a relatively greater emphasis on improving AI model performance and less on how

to ensure that clinicians trust models enough to use them as decision support. Prior findings suggest that the barriers to successful deployment vary according to clinician attitudes regarding their perceived level of control and autonomy (6). When clinicians perceive AI models as unnecessary tools or even threats to their clinical autonomy, implementation efforts must focus more on creating awareness about AI and its potential value in decision-making. As clinicians become more informed about AI and specific tools, the issue of trust in the models and the scientific evidence base comes up. Researchers can build trust among clinicians throughout the model development process by involving them in the development stage, appointing clinical champions, providing the necessary training, and measuring performance (7).

Citizen/patient perspective

The ability of citizens and patients to participate in data collection during a viral pandemic is greater than ever in the time of the Internet, social media, and smartphones. Real-world data can be collected at a rapid rate and used to enrich medical data that is collected in clinics. In order for citizens to contribute to these efforts, it is critical that they have access to trusted evidence-based information. Citizens and patients must be informed about topics such as AI and data sharing, as well as be aware of their potential role in research and the value of their involvement. Specific actions taken within the DRAGON project to engage patients and citizens are described in the subsection 'DRAGON experiences'.

To summarize, the development phase is an iterative process, involving patients, clinicians, researchers, and tool developers. Once the tool is considered user-friendly, facilitates data collection, shares results based on its user's needs, and generated enough trust among both the clinician and the patient in order to make use of it, the development phase is considered completed and the tool is ready for deployment.

Deployment of AI tools

During the deployment phase (Figure 3), the AI model and the user tool are completely integrated. It is user-friendly, so both the patient and the clinician know how to put data in it and are able to interpret the results that they get out of it. Ideally, any lack of trust in the AI model has been overcome during the development stage, and the clinician is able to use the results they are provided in their decision-making. This implies that the researcher's role is complete and their presence in the clinical process is no longer necessary.

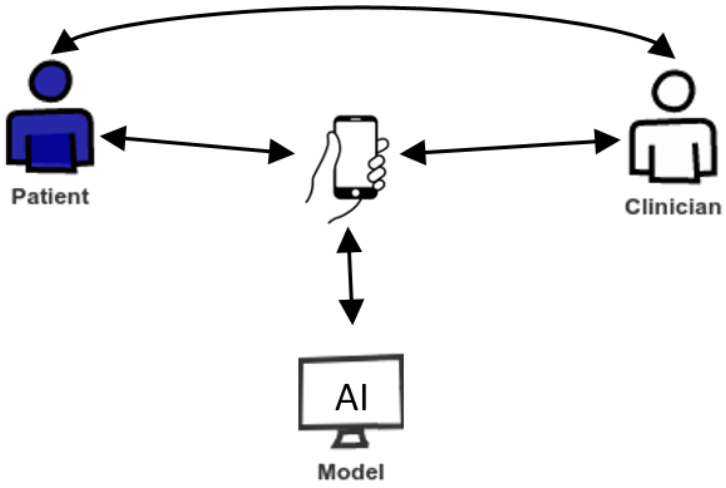

Figure 3. Deployment phase

DRAGON experiences

A clinical use case within DRAGON was the accurate diagnosis of a new patient based on symptom presentation and existing clinical data. In this situation, an AI model is being trained to recognize instances of viral infection by using large volumes of multifactorial data (e.g. CT images, lab results) to compare patients known to be infected with COVID-19 to those who are not.

Empowerment of patients and citizens has been a central theme within the DRAGON project. A citizen and patient stakeholder advisory group (CPAG) was set up, drawing from the patient stakeholder groups organized by the European Lung Foundation (ELF), as well as individuals who have been infected and recovered, individuals who lost family or friends to COVID-19, and healthy citizens. Members of the CPAG took part in Project Integration Team meetings, anonymous online surveys, and interviews as well as the project's general assembly, providing their perspectives as patient stakeholders and guiding the project's research priorities.

These perspectives were used in the development of a set of tools developed by Comunicare for patients (MyCareAvatar) and clinicians (MyPatientCheck). MyCareAvatar is aimed at informing patients and citizens about COVID-19 and contains the following features: (1) an AI-based Self-Assessment Risk Score Questionnaire based on currently available nonograms and predictive models, as well as existing self-assessment questionnaires to help decrease the false negative detection rate; (2) Health Literacy e-learning modules with information on the diagnosis, symptoms, potential, complications, treatment options, and the current WHO guidelines in terms of prevention and containment; (3) Monitoring tools to keep track of relevant measures such as validated electronic Patient Reported Outcome (ePRO) questionnaires (anamnesis, symptoms, quality of life), wearable tech vitals, medication adherence, and enabled early detection and/or efficient follow-up of contaminated citizens.

The clinician tool MyPatientCheck was developed to triage patients more efficiently based on their age and clinical variables that were found to be relevant during

initial analysis of COVID-19 patients (8). Figures 4 and 5 show the input and output screens for this app respectively. The aim is to integrate AI models for diagnosis, severity, and prognosis into this app to create a data-driven decision support system that helps clinicians to determine the most suitable course of action for the individual patient.

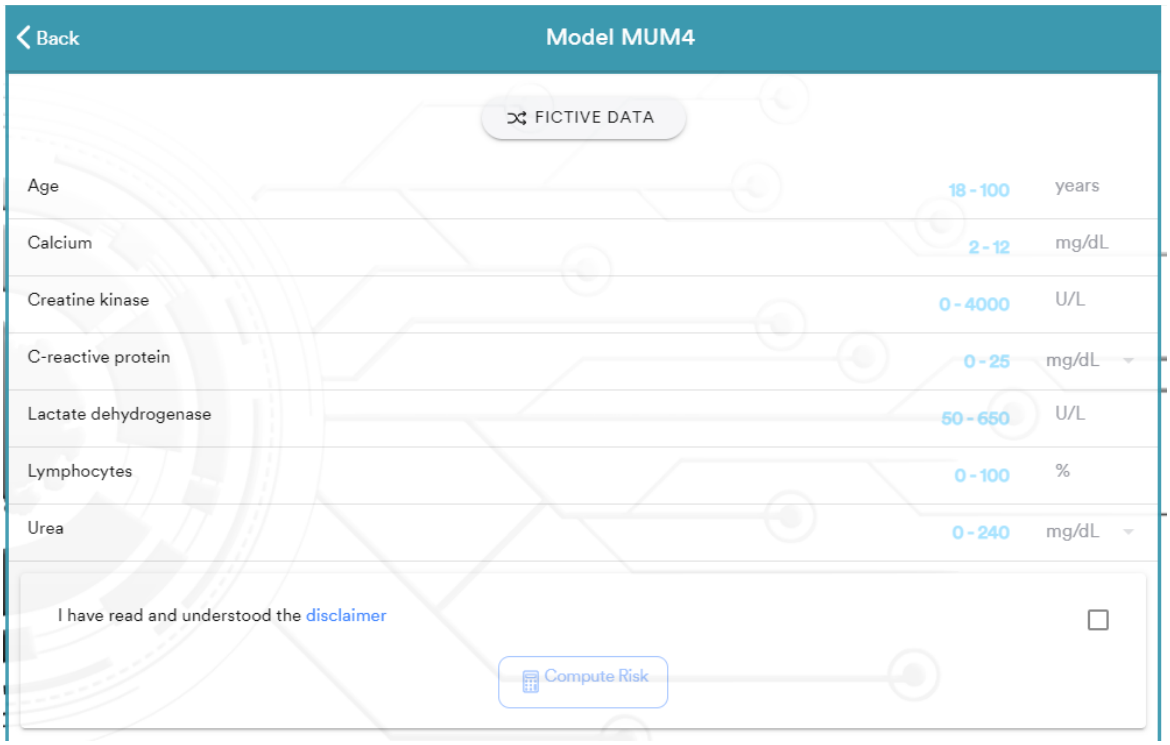

Figure 4: Input screen for one of the models in MyPatientCheck

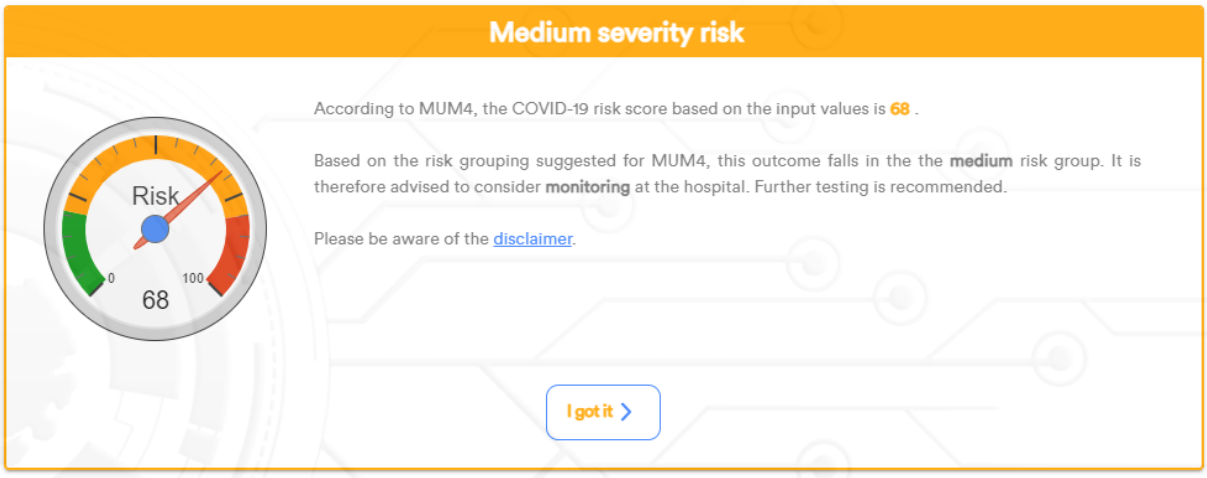

Figure 5: Prognostic calculation for a fictitious patient aged 51 years with calcium 6.2 mg/dL, creatine kinase 113 U/L, C-reactive protein 8.28 mg/dL, lactate dehydrogenase 265 U/L, lymphocytes 22.14%, urea 36.86 mg/dL.

The ultimate aim within the DRAGON project is to integrate MyCareAvatar and MyPatientCheck into a single clinical decision support system with interfaces for both patients and clinicians so that they can make clinical decisions together based

on the latest evidence from the AI model as well as knowledge of the patient's individual situation and symptoms.

Topic 3 - Regulations and standardizations

Given that future viral pandemics will be spread across multiple countries and continents, the processes of data standardization and stakeholder engagement highlighted in the previous topics will likely take place across country borders. Therefore, measures must be taken to protect patient privacy, harmonize various data types, and ensure that the necessary (inter)national ethical approvals can be obtained in a timely manner.

Preserving patient privacy

Developing an accurate AI model requires training data that is voluminous and diverse. This necessitates multicentric collaboration, where many institutes and countries contribute their data to build a rich dataset for model training. The European General Data Protection Regulation (GDPR) and the United States Health Insurance Portability and Accountability Act (HIPAA) provide strict rules on how personally identifiable health data from individuals should be stored and shared. Thus, sharing clinical data between researchers while preserving patients' privacy is a considerable challenge.

In the DRAGON project, data was gathered from multiple sources (e.g. clinical data, CT images) and integrated into a federated learning system. In contrast to static systems, in which models are developed using data collected at one point in time, a federated learning system is a dynamic system that trains AI/ML models from different data sources without the data being gathered in a centralized repository. The untrained model is sent to different centers to be trained on local data and the results are sent back to the central server so that the algorithm can be trained iteratively (9). This decentralized training method ensures that sensitive patient information is not sent outside the center, thus preserving privacy and enabling the model to learn from larger volumes of data.

Regulations and standardizations

Harmonization and standardization of data between clinics during the early stages of a pandemic are essential to allow data sharing and merging. The COVID-19 pandemic saw a massive investment in collaborative research projects, highlighting the importance of developing guidelines for data collection. Since DRAGON is a European project, several points needed to be considered, in particular the compliance with the European Union general data protection regulation (GDPR) and the ethics approvals, the selection and implementation of standards for data collection in the funding framework, and the training and capacity building for data owners.

Recently, the European Commission has adopted a recommendation on a European electronic health record (EHR) exchange format to unlock the flow of health data

across borders. Citizens will be able to securely access and exchange their health data across borders in the EU. This measure might facilitate patient recruitment to clinical trials and streamline clinical trial conduct. With future global consortia, it would be necessary to keep in mind the complexity of various regulations within and outside Europe. A way to overcome this would be to develop global standards.

Ethical approval across countries

One of the biggest and most time-consuming challenges faced within the DRAGON project is the repetitive need for ethical approval to conduct clinical research. Once a new clinical trial has been developed, it needs to be checked by an ethical committee. After approval, the study can be enrolled in that particular center, but usually not in other participating centers. With multicentric clinical trials, most centers prefer to have their ethical committee check the study again before they are allowed to participate in it with their center, while the protocol is the same for all. This slows down the recruitment of participants considerably and is especially harmful during a pandemic when rapid testing and results are necessary in order to take swift action in treating and containing the disease.

It should of course be kept in mind that different countries or continents have different rules and regulations. This could explain the difficulty partially, as each of them needs to make sure it does not collide with their rules and regulations. Nevertheless, this could also be the solution. When looking at Europe the GDPR has been put in place since May 2018. If there'd be an ethical committee operating at a European level, applying the rules of the GDPR, this could solve the issue.

Lessons learned

In summary, the DRAGON project brought to light a number of issues that would need to be addressed in order to ensure that the world is better prepared for a future viral pandemic, namely:

- Avoid harmonization and standardization of data in hindsight. Instead, standardize ways of data collection during regular care, so this is common practice when a pandemic hits. Common standards include CDISC, which was helpful within DRAGON.
- Involve all stakeholders (patients, clinicians, and researchers) while developing tools to collect data that are integrated with AI/ML-models
    - Stay aware of the question from the patient/clinician to make sure the tool will be used to its full advantage
    - Implement their feedback to make the tool user-friendly
    - Explain what the AI/ML model does, work together to develop the output interface in order to gain trust and increase the likelihood of implementation of the tool
    - Deploy the relevant tools among the general population as well rather than limiting them to hospitalized patients, as there is much that can be learned about an unfolding pandemic from citizens who are not yet in the clinical trajectory.
- Develop a tool that can both collect data but also show precision medicine-related results to clinicians and/or participants.
- Be aware that viral pandemics necessitate collaboration across countries and continents. Therefore, ensure that models can learn from data that is distributed across various centers by employing a federated learning system.
- Multicenter clinical trials across Europe each need their own ethical approval, based on their own ethics board. An EU-level ethics board would unify clinical research regulations and speed up the process of starting multi-center studies.

References

1. Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus Pandemic (COVID-19) [Internet]. 2020. Available from: https://ourworldindata.org/coronavirus
2. Barouki R, Kogevinas M, Audouze K, Belesova K, Bergman A, Birnbaum L, et al. The COVID-19 pandemic and global environmental change: Emerging research needs. Environ Int [Internet]. 2021 Jan [cited 2022 Dec 23];146:106272. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0160412020322273
3. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med [Internet]. 2020 Jul [cited 2022 Dec 13];26(7):1037–40. Available from: http://www.nature.com/articles/s41591-020-0916-2
4. Tulk Jesso S, Kelliher A, Sanghavi H, Martin T, Henrickson Parker S. Inclusion of Clinicians in the Development and Evaluation of Clinical Artificial Intelligence Tools: A Systematic Literature Review. Front Psychol [Internet]. 2022 Apr 7 [cited 2022 Nov 29];13:830345. Available from: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.830345/full
5. Abras C, Maloney-Krichmar D, Preece J. User-centered design. Bainbridge W Encycl Hum-Comput Interact Thousand Oaks Sage Publ. 2004;37(4):445–56.
6. Liberati EG, Ruggiero F, Galuppo L, Gorli M, González-Lorenzo M, Maraldi M, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. Implement Sci IS. 2017 15;12(1):113.
7. Browne JT, Bakker S, Yu B, Lloyd P, Ben Allouch S. Trust in Clinical AI: Expanding the Unit of Analysis. In: HHAI2022: Augmenting Human Intellect. IOS Press; 2022. p. 96–113.
8. Wu G, Yang P, Xie Y, Woodruff HC, Rao X, Guiot J, et al. Development of a Clinical Decision Support System for Severity Risk Prediction and Triage of COVID-19 Patients at Hospital Admission: an International Multicenter Study. Eur Respir J [Internet]. 2020 Jul 2 [cited 2022 Jan 25];2001104. Available from: http://erj.ersjournals.com/lookup/doi/10.1183/13993003.01104-2020
9. Dayan I, Roth HR, Zhong A, Harouni A, Gentili A, Abidin AZ, et al. Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med [Internet]. 2021 Oct [cited 2022 Jan 31];27(10):1735–43. Available from: https://www.nature.com/articles/s41591-021-01506-3