1st Symposium "Big Data and Artificial Intelligence: Opportunities for Clinical Decision Making and Health Care Assessment"

Faculty of Medicine and Health Sciences, UB (Barcelona). 6 November 2019

Welcome and opening presentations

After the welcoming ceremony the opening presentations were given by Dr Oriol Pujol Vila, Vice-president for Digital Transformation at the University of Barcelona (UB), Dr Ana Ripoll, Professor in Architecture and Computer Technology, former rector of the Autonomous University of Barcelona (UAB), and President of Bioinformatics Association Barcelona (BIB), Dr Jaime Grego, President of Laboratorios LETI, S.L.U., and Dr Robert Fabregat, Director General for Research and Innovation in Health at the Ministry of Health of the Generalitat de Catalunya.

Dr Pujol talked about the plan to transform the University of Barcelona into a data-focused university and the adaptation of the institution to the digital ecosystem, where data are the focus of everything. For Dr Pujol it is necessary to develop this kind of debating forum to take the right policies and regulations not just into research but into medical practice too. It is necessary for us to know the possibilities of data, to draw conclusions and to know how far we can go, so that the challenges, the opportunities and the development of the legal and ethical framework may help us to progress.

Dr Ripoll focused on what is implied by the arrival of the big data revolution after the Internet tsunami. One of the main challenges is the volume and complexity of data, which are growing exponentially and have to be managed, analysed and interpreted to generate new knowledge. She pointed out the opportunities in the paradigm shift in the diagnosis and treatment of genetically based illnesses, and in gene-based personalized medicine. Spain in general and Catalonia in particular now have the conditions for the development of an ambitious system: genome sequencing centres, storage and management infrastructure, hospitals with basic genomics, technological companies, and so on. For Dr Ripoll there are ethical, social, legal and healthcare challenges, which are holding back the potential of precision medicine for improving the quality of life. She stressed as the objective of the symposium the need to tackle these approaches in order to reinforce, facilitate and not hold back the use of these technologies.

Dr Grego highlighted how digitization is leading organizations to be flexible and able to do more with greater efficiency and quality. These aspects are also applicable to the world of healthcare, whereby working together we can contribute more to the patient, to the health system, and work more efficiently. The most important thing about digitization is that it has to allow us to improve control of the security and the effectiveness of the decisions and the products, and to allocate more resources to innovation, always keeping the patient and their health at the heart of decision-making.

Dr Fabregat focused on families and clinical decisions and he recalled that data are not enough if there is no infrastructure to manage them. He pointed out that, although there are juridical, legal and ethical difficulties, these challenges must not represent a brake, but an opportunity

for decision-making and transparency. We have immense amounts of data, but we have to decide how to use them, guaranteeing the best security for all.

Opening presentation

BIOINFORMATICS BARCELONA (BIB): A STRATEGIC ECOSYSTEM IN HEALTH AS A VISION OF SOCIETIES IN THE 21ST CENTURY. *Dr Ana Ripoll.* *Professor in Architecture and Computer Technology and former rector of the Autonomous University of Barcelona (UAB). President of the Bioinformatics Association Barcelona (BIB).*

Responsibility for health is a global responsibility; it depends on all of us, and it is one of the 17 sustainable development targets. With respect to big data the challenge is to manage the volume, but also the complexity, of data.

Bioinformatics Barcelona is an ecosystem of more than 50 institutions, including universities, research centres, scientific infrastructures, health care research centres, patients' associations and businesses. Why Bioinformatics Barcelona? Because it's not enough to have the technology, it is also a question of common values and of having a common language, understanding the same concepts and speaking the same language among scientists, technologists, ethics committees and health professionals.

The principles that should be applied when designing algorithms are: Justice; Autonomy, the attribution of responsibility, greater intelligence, non-substitution, trust – through transparency; Education and equality; Training citizens, legislators, politicians, legal operators, and so on; Not doing harm and minimizing the negative aspects, always preserving people's privacy.

Five years ago genome sequencing was restricted to the field of research. Its use is now growing in clinical practice, and in the next five years we expect to have genomic data of more than 60 million patients generated within health systems. Governments and institutions have the overall responsibility to speed up the introduction of personalized medicine and to facilitate the timely materialization of the benefits of genomics for individuals, families and health systems.

- Bioinformatics Barcelona: http://www.bioinformaticsbarcelona.eu/

Session 1. Scientific and technical aspects in healthcare practice

Where we are, where we are going and challenges to be overcome from the point of view of patients, professionals and the healthcare system.

Introduction

SCIENTIFIC AND TECHNICAL ASPECTS IN HEALTHCARE PRACTICE. *Dr Xavier Pastor.* *Head of Medical Informatics at the Hospital Clínic in Barcelona and lecturer at the UB.*

Why talk about big data and Artificial Intelligence (AI) in medicine? Hospitals and primary health centres (CAPs) are already working without paper and patients' data are digitized. Why not use advanced data analytics tools for research purposes? Demand is growing and there are great

promises and expectations. But in healthcare there are barriers to the use of data analytics tools. The first barrier is ignorance, due to a shortage of academic training in the biomedical community, which is in need of the necessary investment to ease this situation. The second is fear. The fear of losing privacy, from the user's point of view; the fear of losing the data gathered from the researcher's perspective, or losing what it has been very difficult to achieve from the perspective of organizations. To counter these fears we must invest in transparency, consensus and acknowledgement of the value of data. The third barrier is scepticism about the results and their real returns, which is countered by investing in quality, methodology and assessment.

ARTIFICAL INTELLIGENCE: A REVOLUTION IN HEALTH? ***Dr Josep Lluís Arcos.*** *Head of the Department of Machine Learning at the Artificial Intelligence Research Institute (IIIA-CSIC).*

Artificial Intelligence (AI) is defined as a series of "programs that solve tasks that would require intelligence if a human being solved them." This behaviour is not just inferring a diagnosis on the basis of knowledge and rules, which was the first approach taken with respect to AI. It is a rather simplistic approach, since intelligence requires perception, representation, reasoning, learning and autonomy.

Health data are particularly complex. They can be found in various sources, from medical records to laboratory results. They have variable formats, they are not always digitized and the definitions are subjective depending on the source. The value of data is not the data themselves, but the knowledge that can be inferred from them. Therefore, we must avoid what is known as the "Garbage In, Garbage Out" syndrome: if we enter poor-quality, non-standardized or incomplete data, the result of AI will be of poor quality, because the value of the knowledge obtained depends on the trustworthiness of the data used.

The opportunities for big data in health involve progressing from descriptive and diagnostic analyses towards predictive analyses, with the capability of making decisions in health. The volume of data necessary in this evolution is growing exponentially, including the challenges posed by their use. At the same time, opportunities are increasing as the data can be used to go from questions beginning with "What" and "Why" to "What will happen" and "How to prevent".

- Frost & Sullivan (2016), Growth Opportunities for Healthcare Big Data - An Analysis of Global Case Studies - Research and Markets https://www.researchandmarkets.com/research/gp4bdn/growth

SISCAT PLAN FOR HEALTH INFORMATION SYSTEMS, DEPARTMENT OF HEALTH. STAGE 1 THE ELECTRONIC MEDICAL RECORD: CENTRAL DATA REPOSITORY. ***Sr Pol Pérez Sust.*** *TIC General Coordinator, Department of Health, Generalitat de Catalunya.*

The Plan for Health Information Systems of the Integrated Public Use Healthcare System of Catalonia (SISCAT) is an information management project that hopes to position the information systems of Catalonia among the best in the world.

One of the strategic objectives between 2020 and 2023 is the Electronic Medical Record (HES), conceived as a central repository of healthcare data. This commitment makes it possible to consolidate a citizen-focused digital information model, to facilitate decision-making in

healthcare and management throughout the healthcare chain. The important thing is not just to have an algorithm, but to know if it is the algorithm that we as a society want.

Also important is the creation of the Artificial Healthcare Intelligence Centre of Catalonia (CIASCAT), introduced in the CatSalut system. The objective of CIASCAT is to coordinate and guarantee security, the quality of the data, transparency and efficiency, so that it can be aligned with the SISCAT centres' healthcare policies.

ARTIFICIAL INTELLIGENCE FOR SUPPORTING CLINICAL DECISION-MAKING PROCESSES. ***Dr Carolina García Vidal.*** *Senior specialist, Infectious Disease Service of the Hospital Clínic-IDIBAPS in Barcelona.*

The presentation focused on analysing cases in order to improve healthcare practice through research projects. The infections that immune suppressed patients suffer from are extremely difficult to treat, a situation that forces us to use more broader spectrum antibiotics, thus generating more resistance and greater toxicity, more cost for the health system and a variety of problems associated with multi-drug resistant infections.

To personalize antibiotic treatments and minimize the errors in their administration solutions were sought based on the personal data of the patients that they had and how to use them. The majority of the data was destructured and hard to process, whereby an important part of the process has been converting them into legible data available to the scientific community. The result of this processing is more than one billion items of patients' data in ten years at the Hospital Clínic in Barcelona. The data have also been corrected and revised to achieve 99% precision in them.

With respect to diagnosis and choice of treatment, the results tell us that doctors are wrong 55% of the time, whereas with the system we have developed error has been reduced to less than 5%. Among the achievements derived from this improvement in diagnosis are the reduction of the use of broad-spectrum antibiotics by 99%, saving €200,000, a reduction of between 30 and 50% in the generation of resistance, and 20% less toxicity.

- Garcia-Vidal, Carolina et al. Artificial intelligence to support clinical decision-making processes EBioMedicine, Volume 46, 27-29
  https://www.ebiomedicine.com/article/S2352-3964(19)30454-2/fulltext

Session 2. Ethical, social and legal aspects

Criteria to be applied by ethics committees in the assessment of study proposals, the state of the question and how to go forward. How to boost the use of new technologies in health. How to deal, in a practical approach, with the questions that arise on an ethical, legal and social level.

TRANSFORMING HEALTHCARE AND INCREASING VALUE, DATA ANALYTICS AND AI. ***Dr César Velasco.*** *Director of the Agency for Health Quality and Assessment of Catalonia (AQuAS).*

In order to talk about data we have to know where we are going, and we must do it as responsibly and ethically as possible. We face the challenges of a dynamic society with new expectations, where we are faced with new challenges such as chronicity and demand in society far outstripping the supply of services at the present moment. Also, new technologies are being developed and applied in a context in which resources are limited. At the same time, we are witnessing a paradigm shift that affects other sectors and which generates new demands.

The goal is more personalized and preventive medicine, but the current system is based fundamentally on purchasing products and drugs, and not on services. The objective of the Public Data Analysis for Health Research and Innovation Program (PADRIS), which already exists, is the analysis of data for decision-making, and it is being promoted and improved to achieve a system focused on the person, where the real needs of society are attended to.

- *PADRIS Program* http://aquas.gencat.cat/ca/ambits/analitica-dades/padris/

ETHICAL, SOCIAL AND LEGAL ASPECTS OF BIG DATA AND ARTIFICIAL INTELLIGENCE: CHALLENGES FACING RESEARCH ETHICS COMMITTEES. ***Dr Fernando García López.*** *Doctor at the National Epidemiology Centre of the Carlos III Health Institute, member of the Bioethics and Law Observatory, and lecturer on the master's degree course in Bioethics and Law at the University of Barcelona (UB). Member of the Carlos III Health Institute's Animal Research and Welfare Ethics Committee (CEIyBA).*

Research ethics committees are key players in the future of the regulation of research that uses big data and Artificial Intelligence (AI). Dr García pointed out the possibility that, in future, consent in research could be dispensed with when, as pointed out in the International Ethical Guidelines for Biomedical Research Involving Human Subjects of the Council for International Organizations of Medical Sciences (CIOMS), it is not feasible to do the research without this dispensation or modification; when the research has important social value, and when the research only entails minimal risks for participants.

The development of AI requires a framework of ethical principles and democratic requirements. Various independent and inter-disciplinary advisory bodies of the European Commission have drafted reports in this respect. The European Group on Ethics in Science and New Technologies, as an independent advisory group of the European Commission, establishes the following principles for AI, robotics and autonomous systems: human dignity, autonomy, responsibility; justice, fairness and solidarity; democracy, the rule of law and accountability; security, and physical and mental integrity; protection of data and privacy, and sustainability. Moreover, the Independent High-level Expert Group on Artificial Intelligence, also created by the European Commission, lays out the ethical guidelines for trustworthy AI through the principles of respect for human autonomy, prevention of harm, fairness and explicability.

In the age of big data and AI, it is the duty of research ethics committees to assess in every research project the social value and the value of justice, fairness and solidarity; the prevention of harm, which includes data protection, and the avoidance of discrimination of vulnerable parts of the population; transparency and explicability with the consequent responsibility for

automated decision making; accountability. A democratic debate also becomes necessary, as well as the involvement of society in the conception, design and execution of healthcare research, tasks in which research ethics committees must be involved; and finally, establishing the relevant security measures to protect personal data.

Dr García formulated the following questions: Are research ethics committees ready to carry out these functions? Is the make-up of research ethics committees good for this? How should members of research ethics committees be trained? Research ethics committees must not be used as mechanisms to cover interests contrary to justice, fairness or respect for people, and they must be answerable to society in order to defend human rights and scientific integrity.

- European Group on Ethics in Science and New Technologies. "Statement on Artificial Intelligence, Robotics and 'Autonomous' Technologies" https://ec.europa.eu/info/publications/ege-statements_en
- Independent High-level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy Artificial Intelligence https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

ASSESSMENT OF THE ETHICAL, LEGAL AND SOCIAL ASPECTS OF HEALTH RESEARCH PROJECTS. ***Dr Itziar de Lecuona.*** *Associate Professor in the Department of Medicine and Deputy Director of the Bioethics and Law Observatory – UNESCO chair in Bioethics of the University of Barcelona (UB). Member of the Research Ethics Committee of the Hospital Clínic in Barcelona and of the UB Bioethics Committee.*

Big data and artificial intelligence (AI) are not the future, they are the present. The first obstacle for understanding the paradigm shift this poses is the speed at which progress in these fields is taking place, and the second the convergence of technologies, which adds greater complexity for analysing the technical, ethical, legal and social implications. Big data and AI are a scientific, political and economic commitment of the European Union to create a competitive single digital market. We are witnessing an excessive production of laws that should be focused more on down-to-earth principles and rules rather than procedures. Research ethics committees are facing this challenge.

The data-driven digital society is characterized by re-identification, due to the huge amount of information (datasets, personal ones among them) stored in databases. Big data refers to the correlation of datasets using mathematical algorithms to improve decision-making, predicting behaviours through the determination of patterns of behaviour. Investigation has shown that just with someone's gender, date of birth and postcode it is possible to re-identify a high percentage of people. We must stop using the concept of anonymization because it is obsolete. Now, as well as traditional codification there is a new concept, pseudonymization: "Processing personal data in such a way that they can no longer be attributed to an interested party without using additional information that must be dissociated and subject to technical and organizational measures."

Dr De Lecuona makes the following proposals so that research ethics committees can suitably assess research projects based on big data and AI: incorporating new types of members, in particular, data scientists, in order not to make technical problems ethical; investing in specific

training, especially in digital literacy. Collaboration between research ethics committees, legal services and data protection delegates, innovation units and information and communication technologies. Delimiting agreements with third-party service providers, adapted to the field of healthcare to protect personal data (including the tech giants). Avoiding secondary uses of personal information and covert discrimination. In short, personal data markets disguised as research must be avoided.

It is necessary to explain to potential participants, who provide the raw materials, how these data are going to be collected and used. To encourage transparency and accountability because informed consent is not the only safety valve. To protect people today it is necessary to look after their personal data. Finally Dr De Lecuona reminded us that we are witnessing a new model of research in which research is done with databases created for commercial purposes that are later ceded to the pharmaceutical and biotech industry. While in analogue mode participants were sought to do research with them, in digital mode direct-to-consumer genetic testing opens the door to future research uses.

- De Lecuona, I. (2018). Evaluación de los aspectos metodológicos, éticos, legales y sociales de proyectos de investigación en salud con datos masivos (big data). Gaceta Sanitaria, 32(6), pp. 576-578. DOI: 10.1016/j.gaceta.2018.02.007 http://www.gacetasanitaria.org/es-evaluacion-los-aspectos-metodologicos-eticos-articulo-S0213911118300864

THE ROLE OF THE DATA PROTECTION DELEGATE IN RESEARCH ETHICS COMMITTEES IN THE ASSESSMENT OF BIG DATA AND ARTIFICIAL INTELLIGENCE PROJECTS. ***Sra Miriam Méndez.*** *Lawyer and Data Protection Delegate of the legal services of the Hospital Clínic in Barcelona. Member of the Research Ethics Committee of the Hospital Clínic in Barcelona.*

The function of the Data Protection Delegate (DPD) is to advise on matters of data protection so that the research projects assessed by the research ethics committee of the relevant research centre comply with the applicable law, such as the General Data Protection Regulation (GDPR). The pseudonymization of the data gathered for research projects must be checked, and also those that are based on the reuse of data already stored in databases. Thus, the DPD must confirm that the project submitted for assessment complies with the main data protection obligations, like for example the assessment of its impact on the rights of the potential participants; international transfers and the systems made available for the use of personal data. The DPD must ensure that there is a legitimate basis for the use of the data, that it is possible to apply the principle of data traceability, and to ensure that the project does not allow for the re-identification of the people taking part. On this point Sra Méndez reminded us of the possibility of drawing up profiles of people in research and the procedures for the exercise of the rights of access, rectification, cancellation and opposition, as well as the right to data portability, and to be forgotten, among others.

It is essential for Research Ethics Committees to have informed consent models adapted to the needs of research today and which make it possible to protect the rights of the participants, as well as protocols for researchers and evaluators.

The DPD takes part in the review of processes of information and informed consent, including the models of consensus agreed on by bodies such as the Spanish Agency for Medicines and Medical Devices. The DPD can contribute to the creation of tools that facilitate the application of the law of data protection, but collaboration between the different research ethics committees authorized in our system is essential to establish common criteria, and between the different services of the institution where the research is being done.

- Resources of the Spanish Data Protection Agency: Innovation and technology: https://www.aepd.es/areas/innovacion/index.html Project BIGDATIUS www.bigdatius.com

Session 3. The Application of Big Data and Artificial Intelligence in Clinical Research and Health Assessment

What aspects are critical for big data and artificial intelligence studies to be considered in decision-making in the assessment of the effectiveness, the safety and the efficiency of medicines and other healthcare technologies? What areas for action should be given priority?

HOW TO DEAL WITH THE INEVITABLE? ***Dr Laura Sampietro-Colom.*** *Deputy Director of Innovation, and Head of the Health Technology Assessment Unit at the Hospital Clínic in Barcelona.*

New sources of data and new capture and analysis technologies are inevitable, but the decisions are based on evidence, not on data. Real-World Data (RWD), observational and administrative data, are necessary but not sufficient to generate Real-World Evidence (RWE), evidence obtained from studies that are not clinical trials. RWE is key in the cases where it is not possible to conduct a randomized clinical trial (rare diseases, the need for long-term monitoring, when there is uncertainty about the dose of the drug or the external validity of the clinical trial, etc.).

The life cycle of healthcare technologies and RWD/RWE is defined through the level of adoption in time. An idea is considered to be an 'emerging' technology around 10 years before its market launch; when early adopters incorporate it, it is considered 'new' technology, and when it is adopted by the majority it is considered to be 'established' technology. As time goes by it is either 'optimized' or it becomes 'obsolete'.

Current challenges in the generation and use of RWD to achieve RWE lie in the quality of the data, the skills of professionals and their training, the transferability of results, the differences in infrastructure and its cost, and access to data due to the lack of incentives or restrictions. There are different initiatives for overcoming these challenges and one of the key factors is that the actors must be involved from the start of the technology's life cycle, including especially people from the field of data, but who are not from the health sector.

- Oortwijn, W., Sampietro-Colom, L., & Trowman, R. (2019). How to Deal with the Inevitable: Generating Real-World Data and Using Real-World Evidence for HTA Purposes – From Theory to Action. International Journal of Technology Assessment in Health Care, 35(4), 346-350. https://doi.org/10.1017/S0266462319000400

BIG DATA IN CLINICAL RESEARCH AND ASSESSMENT: a view from IQVIA. *Sr Carles Illa.* *Director of Healthcare in IQVIA.*

Carles Illa, in his presentation, claimed that we could take advantage of the potential of information-based systems by increasing the diagnosis and treatment database. The value of using this information in daily practice is more standardized results-oriented healthcare. Precision medicine increases diagnostic and therapeutic efficiency and specific strategies are developed to learn what works and what does not.

Among the examples of data transformation and use carried out by IQVIA, Sr Illa mentioned several examples of big data use such as that developed for Genomics England, a case in which access to and the processing of the data of more than 100,000 genomes was necessary for the purpose of studying rare diseases and neoplasms. It has also been used to obtain safer and more efficient clinical trials, through monitored risk assessment, oriented towards the centralization of information to improve its register, minimizing errors and increasing the reporting of adverse effects in clinical trials. The challenge of processing natural language is another approach in which big data can be used in health. Using natural language processing algorithms it was possible to convert the text of the medical records of 100,000 patients and more than 30 million documents into data with structured value.

Lastly, the difficulty of managing this type of data for their use in the field of healthcare was discussed. This requires the use of technological solutions capable of combining the potential of the data for decision-making, with the necessary privacy due to its sensitivity, applying the existing laws and regulations.

Closing speech

Dr Joan Bigorra. *Senior advisor on Innovation of the Hospital Clínic in Barcelona and Director of Strategy and Innovation of the Barcelona Institute for Global Health.*

To close the symposium, Dr Bigorra shared some thoughts on digital transformation in the field of healthcare. This transformation is not a fashion, it is an opportunity and an imperative, but in order to develop its entire potential it is necessary for everyone to agree and to comply with methodological, ethical and legal standards that are sufficient and efficient. For this, innovation has to be open and must include all the actors, from academe, health centres, researchers, bioethicists, jurists, the pharmaceutical industry, that of medical technologies and digital tools, patients, etc. And the approach must be focused on the patients, based on “for the patient and with the patient”. Another challenge for this digital transformation is to generate leaderships at the critical moment of the management of this change in which we find ourselves. For this, it is necessary for organizations and committed people to generate meeting places where agreements and actions for change are developed.

Dr Antoni Trilla, Dean of the Faculty of Medicine and Health Sciences of the University of Barcelona, where the symposium was held, and as the representative of one of the organizing institutions, closed the event, congratulating all the speakers, moderators, and especially the people who made the symposium possible, who filled the hall to capacity.