



**HAL**  
open science

## **”Artificial intelligence”: Which services, which applications, which results and which development today in clinical research? Which impact on the quality of care? Which recommendations?**

Vincent Diebolt, Isaac Azancot, François-Henri Boissel, Isabelle Adenot, Christine Balagué, Philippe Barthelemy, Nacer Boubenna, Hélène Coulonjou, Xosé Fernandez, Enguerrand Habran, et al.

### **► To cite this version:**

Vincent Diebolt, Isaac Azancot, François-Henri Boissel, Isabelle Adenot, Christine Balagué, et al.. ”Artificial intelligence”: Which services, which applications, which results and which development today in clinical research? Which impact on the quality of care? Which recommendations?. *Thérapie*, 2019, 74 (1), pp.155 - 164. 10.1016/j.therap.2018.12.003 . hal-02053243

**HAL Id: hal-02053243**

**<https://hal.science/hal-02053243>**

Submitted on 22 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## THERAPIE

### GIENS WORKSHOPS 2018 / CLINICAL RESEARCH

Epub ahead of print puis numéro 1 (janv fév) 2019

**“Artificial intelligence”: Which services, which applications, which results and which development today in clinical research? Which impact on the quality of care? Which recommendations ?\***

Artificial intelligence

*Vincent Diebolt<sup>a,\*</sup>, Isaac Azanco<sup>b</sup>, François-Henri Boisset<sup>c</sup>, the participants of Giens XXXIV Round Table « hot topic N°2 », Isabelle Adenot<sup>d</sup>, Christine Balague<sup>e</sup>, Philippe Barthélémy<sup>f</sup>, Nacer Boubenna<sup>g</sup>, Hélène Coulonjou<sup>h</sup>, Xosé Fernandez<sup>i</sup>, Enguerrand Habran<sup>j</sup>, Françoise Lethiec<sup>k</sup>, Juliette Longin<sup>l</sup>, Anne Metzinger<sup>m</sup>, Yvon Merlière<sup>n</sup>, Emmanuel Pham<sup>o</sup>, Pierre Philip<sup>p</sup>, Thomas Roche<sup>q</sup>, William Saurin<sup>r</sup>, Anny Tirel<sup>s</sup>, Emmanuelle Voisin<sup>t</sup>, with the contribution of Thierry Marchal<sup>u</sup>*

<sup>a</sup> F-CRIN, 31059 Toulouse, France

<sup>b</sup> AP-HP, hôpital Lariboisière, 75475 Paris, France

<sup>c</sup> Novadiscovery, 69009 Lyon, France

<sup>d</sup> Haute autorité de santé, 93218 Saint Denis la Plaine, France

<sup>e</sup> Institut mines-telecom business school, 91011 Evry, France

<sup>f</sup> Astrazeneca, 92400 Courbevoie, France

<sup>g</sup> INSERM transfert, 75013 Paris, France

<sup>h</sup> AP-HP, DRCI, 75475 Paris, France

<sup>i</sup> Institut Curie, 75005 Paris, France

<sup>j</sup> *Fonds FHF recherche et innovation, 75993 Paris, France*

<sup>k</sup> *Janssen Cilag, 92787 Issy les Moulineaux, France*

<sup>l</sup> *Merck santé, 69008 Lyon, France*

<sup>m</sup> *Hospices civils de Lyon, HCL, DRCI, 69229 Lyon, France*

<sup>n</sup> *CNAMTS, 75986 Paris, France*

<sup>o</sup> *IPSEN, 91940 Les Ulis, France*

<sup>p</sup> *CHU de Bordeaux ; université de Bordeaux, 33076 Bordeaux, France*

<sup>q</sup> *Delsol avocats, 69001 Lyon, France*

<sup>r</sup> *Dassault systèmes – 3DS, 78946 Vélizy Villacoublay, France*

<sup>s</sup> *MSD, 92400 Courbevoie, France*

<sup>t</sup> *Voisin consulting life science, 92100 Boulogne, France*

<sup>u</sup> *ANSYS, université catholique de Louvain, 1348 Louvain la Neuve, Belgique*

Text received 5 november 2018; accepted 13 november 2018

\*Corresponding author. F-CRIN, pavillon Leriche, hôpital Purpan, place du Dr Blayac, 31059 Toulouse cedex 9, France.

*E-mail adress: vincent.diebolt@inserm.fr (V. Diebolt)*

*\* Articles, analyses and proposals from Giens workshops are those of the authors and do not prejudice the proposition of their parent organization.*

## Summary

Artificial Intelligence (AI), beyond the concrete applications that have already become part of our daily lives, makes it possible to process numerous and heterogeneous data and knowledge, and to understand potentially complex and abstract rules in a manner human intelligence can but without human intervention. AI combines two properties, self-learning by the successive and repetitive processing of data as well as the capacity to adapt, that is to say the possibility for a scripted program to deal with multiple situations likely to vary over time.

Roundtable experts confirmed the potential contribution and theoretical benefit of AI in clinical research and in improving the efficiency of patient care. Experts also measured, as is the case for any new process that people need to get accustomed to, its impact on practices and mindset. To maximize the benefits of AI, four critical points have been identified. The careful consideration of these four points conditions the technical integration and the appropriation by all actors of the life science spectrum: researchers, regulators, drug developers, care establishments, medical practitioners and, above all, patients and the civil society.

- 1st critical point: produce tangible demonstrations of the contributions of AI in clinical research by quantifying its benefits
- 2nd critical point: build trust to foster dissemination and acceptability of AI in healthcare thanks to an adapted regulatory framework
- 3rd critical point: ensure the availability of technical skills, which implies an investment in training, the attractiveness of the health sector relative to tech-heavy sectors and the development of ergonomic data collection tools for all health operators
- 4th critical point: organize a system of governance for a distributed and secure model at the national level to aggregate the information and services existing at the local level

Thirty-seven concrete recommendations have been formulated which should pave the way for a widespread adoption of AI in clinical research. In this context, the French “Health data hub” initiative constitutes an ideal opportunity.

## **KEYWORDS**

AI; Data; Knowledge; Clinical research; Clinical trials; Real-life studies; Assessment; Training; Interdisciplinary; Interoperability; Governance

## **Abbreviations**

AI: artificial intelligence

AMM: marketing authorization

ANR: *Agence nationale de la recherche*

ANSM: *Agence nationale de sécurité du médicament et des produits de santé*

CFDA: China food and drug administration

CNAMTS: *Caisse nationale d'Assurance maladie des travailleurs salariés*

CNIL: French *commission nationale de l'informatique et des libertés*

CPTS: territorial professional communities of health

DGOS: *Direction générale de l'offre de soins*

DMP: shared medical record

EMA: European medicines agency

EHESP: *Ecole des hautes études en santé publique*

ESSYM: *Ecole supérieure des systèmes de management des organisations*

FDA: Food and drug administration

GAFAM: Google, Apple, Facebook, Amazon

GPRD: general data protection regulation

HAS: *Haute autorité de santé*

i~HD: European institute for innovation through health data

IFPs: interoperability platforms

IMI: innovative medicines

IVF: in vitro fertilization

PMDA: Japan pharmaceutical and medical device agency

PTA: territorial support platforms

RIPH: research involving human intermediation

RM: reference methodology

Not a single day goes by where artificial intelligence (AI) is not mentioned in a publication or a news website. The subject, which lacks rigor in its definition, triggers debates and controversies, feeds fantasies, hopes and fears about its promises and the risks it entails. The augmented man (“superman”) or the obsolete man?

Numerous fields of activity and every moment of our daily life are likely to be impacted by the latest revolution in progress: transport; security; education; banking and insurance services; and of course health both in its dimension of care but also in research, development and experimentation.

At its most basic, AI can be a substitute for human intervention to make routine decisions, such as turning on a vehicle's headlights when the night begins to fall. In the field of health, the insulin pump measures the level of glucose in the blood and possibly the level of activity of the patient: from these data it automatically makes the decision to inject insulin. The defibrillator collects information about the movement of the heart and, if its algorithm detects fibrillation, sends a defibrillation pulse.

But AI's potential in health goes far beyond that and, if its most eager proponents are to be believed, would be an accelerator of medical progress and a source of substantial savings. While the round table's primary focus was to investigate these lofty claims, the starting point has been to reflect on the terms and define the field of investigation.

AI is first and foremost an analytical technology that has the ability to process numerous and heterogeneous data and knowledge, and to understand potentially complex and abstract rules in a manner human intelligence can but without human intervention.

The notion of data is tied to that of knowledge, a necessary level of abstraction to give data meaning. We also find in this definition two properties specific to AI:

- Self-learning, a system that by successive and repetitive processing of data has the capacity to increment its mode of treatment and to continuously improve its relevance.
- A certain adaptability, that is to say the possibility for the same scripted program to deal with multiple situations that may vary over time.

In a first field of application, AI differs from the “IoT” (internet of things): “vertical” or “simple” AI equips an increasing number of implanted or connected devices in healthcare. As far as therapeutic research is concerned, the entire process of developing new drugs has been investigated starting from the earlier stages (screening for molecules) until the introduction of the finished drug in the system of routine care, which implies recognition by all regulatory actors.

In a second field of application, for the purpose of improving the efficiency and coordination of care, “transverse” or “complex” AI is associated with conversational capabilities which are intended to facilitate the synthesis and understanding of multiple information from both connected and implanted objects, electronic medical records and various incidental services at the disposal of patient and practitioners.

With regards to the “care coordination” aspect, the prerequisite for having a unified access to information from which the cross-cutting AI agents can develop is to ensure the convergence of health networks. This involves governance changes to migrate from non-interoperable centralized architectures to a distributed model based on Interoperability or Intermediation Platforms and likely secured by blockchain technology if this proves to be relevant and meets the constraints of this type of system.

All roundtable experts were unanimous and confirmed the potential contribution and theoretical benefit of AI in clinical research and in improving the efficiency of patient care. Experts also measured, as is the case for any new process that people need to get accustomed to, its impact on practices and mindset. To maximize the benefits of AI, four critical points have been identified. The careful consideration of these four points conditions the technical integration and the appropriation by all actors of the life science spectrum: researchers, regulators, drug developers, care establishments, medical practitioners and, above all, patients and the civil society.



## **1st critical point: artificial intelligence and clinical research, from promises to demonstration: what evidence?**

The failure rate in clinical research is very high. Less than 10% (5,1% in cancerology) of drug candidates reach the market [1]. In phase III alone and a sample of 640 drug candidates, the failure rate is 54% according to a study published in 2016 [2]. While modeling and simulation is common in other sectors, e.g. automobile and aerospace industries, it is not yet prevalent in therapeutic research and development.

For medical devices, the failure rate is likely lower because “patient-device” interactions are easier to observe and therefore to resolve. On the other hand, a not insignificant rate of recall of products is observed: the products pass the clinical tests but when applied to a larger population, several problem cases emerge. Since it is impossible to test for the planet's 7.4 billion inhabitants, other approaches are to be found, such as digital twins, replicas and numerical simulation models of an object, process or a system made from data analysis and machine learning [3].

Owing to the lack of dissemination of scientific evidence in the field, the potential of AI in clinical research is still largely theoretical. Nevertheless, despite the confidential nature of drug R&D and the protracted nature of the drug development cycle, tangible use cases start to accumulate [4-6].

AI impacts clinical research in four different ways:

1. Impact on failure / success rates.
2. Impact on the reduction of time to market (“time to patient access”) [7].
3. Impact on the reduction of development costs in absolute terms.
4. Significant reduction in the number of post-marketing product recalls.

For real-life studies and the course of care, evaluation of AI’s impact would focus on three criteria:

1. Predictive and pathological predictive signals
2. Procedures for intervention

### 3. Impact over time on the patient's condition and quality of life

In this field, we have a first set of references which, with the help of figures, confirms the comparative advantage - compared to current medical practices - of using artificial intelligence tools. In a study of embryonic quality in the in vitro fertilization (IVF) practice setting, compared to human evaluation of embryo quality, an artificial neural network algorithm was more accurate and allowed better identification of embryos of good quality, thus improving the score of an index (Kappa index) that quantifies the level of concordance. For artificial intelligence, it reached 88% while the agreement between the three embryologists was only 50 to 54% [8].

Still in the gynecological field, in the context of the preservation of fertility, the use of a “deep learning” algorithm (test on 3 million images) allowed to divide by 30 the count and the analysis of ovarian follicles according to a comparative study on animal model (murine model) [9]. In dermatology, a team from the pigmented lesion and melanoma program at Stanford cancer institute (Stanford university) has developed a learning algorithm that performs as well in detecting and diagnosing skin cancers as a group of 21 experienced dermatologists [10].

The diffusion, credibility and acceptability of AI, in particular by the regulatory agencies issuing marketing authorizations, depends on the demonstration capacity of its inputs and the perception of benefits by regulatory agencies, health professionals and patients.

#### **To overcome the lack of tangible demonstrations, three recommendations are made:**

1. Create a national warehouse that centralizes use cases and collaborates at European level with the i ~ HD (European institute for innovation through health data) set up in 2016. This warehouse would involve contributions from national health system databases / cohorts, but also the contributions of drug developers who should agree to make available some of their development plans and their benefits / failures. Open innovation would collectively save valuable time in clinical development.

2. Design, finance, and build a battery of prototypical projects in the form of comparative clinical trials (double-blind: with / without AI).

3. Build a methodology for assessing the impact of AI, known and shared, and incorporating a new taxonomy.

**2nd critical point: create confidence to ensure the dissemination and acceptability of ai in health research, in a suitable regulatory context.**

Confidence in these new approaches will be built through information, the definition of a number legal and technical rules and safeguards, and due consideration for ethical issues.

**Inform to enlighten and reassure public opinion**

If “fear arises sooner than anything else” (Notebooks / Leonardo da Vinci), we know those ignited by artificial intelligence:

- The AI is a black box where we do not know what is happening.
- Doubts about the reliability and robustness of the results (“rubbish in / rubbish out”).
- Dangers if AI algorithms fall into malicious hands.
- A disembodied and dominating form of intelligence.

From an ethical point of view, it can be noted that AI in clinical research could make it possible to answer complex scientific questions through the use of existing health data by avoiding the need for additional experiments, and thus exposure to risks inherent in any clinical trial of a large number of patients.

According to a study conducted by the “Odoxa” study institute in May 2016 with a sample of nearly 1,000 people, in the face of AI, the French are very divided, 49% perceiving it as an opportunity and 50% as a threat to fear [11]. Another survey conducted by the same institute (results published on 16/11/2017) for the “health care data institute” puts 83% of French people who would agree to share their anonymous health data if it meets at least one of the following objectives: improvement of medical diagnoses and treatments; faster progression of medical research [12].

Two lessons are to be learned:

- First that, disembodied, the AI sometimes remains for our fellow citizens:
  - o either basic digital applications such as pre / post-hospital online questionnaires, online pre-admissions or hospital kiosks,
  - o either to the fictional image of a self-aware entity replacing the human mind (such as HAL 9000 in Stanley Kubrick's "2001, A Space Odyssey" or to allude to a more pop culture reference, "Terminator 3: rise of the machines")
- The second lesson is that survey respondents would give the go-ahead for tangible results on the effectiveness of care and scientific breakthroughs without going so far as to accept the appropriation of their data by a dehumanized AI who would manage their destiny.

### **Recommendation**

- Launch an education and awareness campaign in care centres, with and patient associations, on the interest and the benefits of using AI in healthcare, as well as on the legal and technical safeguards existing, taking into account ethical issues. It should be based on concrete examples and by adapting communication tools to each target.

### **Towards recognition of AI in marketing authorization and approval procedures by regulatory authorities**

In its annual Appropriations Bill voted on July 20, 2017, the US Congress has mandated the Food and drug administration (FDA) to expand the use of in silico clinical trials and the development of digital twins. Internally, the FDA has set up a dedicated "Modeling & Simulation" working group of nearly 200 experts for the development, validation and exploitation of these approaches.

Europe is also mobilizing and, at the Commission level, preparing its “digital” directive. digital health / big data is one of the four priorities of future editions of the “innovative medicine innovative / IMI / think big” call for projects. The European medicines agency has also set up a working group on the subject.

Lastly, the European patent office has for the first time, in the annual update of its guidelines applicable from 1 November 2018, to describe on the basis of which criteria it is appropriate to to appreciate the patentability of AI-based devices and machine learning [13].

For national regulatory and evaluation authorities - in France the *Agence nationale de sécurité du médicament et des produits de santé* (ANSM) and the *Haute autorité de santé* (HAS) the recognition of in silico during the evaluations and authorizations of health products requires an adaptation of their practices and procedures.

## **Recommendations**

Need for regulatory agencies of therapeutic products and health practices to adapt their method of evaluation:

- Harmonization of good practices between the main European medicines agency/Food and drug administration/Japan pharmaceutical and medical device agency/China food and drug administration (EMA / FDA / PMDA / CFDA) agencies for the integration of “digital evidence” approaches into authorization or certification procedures (marketing authorization [AMM] for medicines and “CE mark” for medical devices).

- Establishment of an evaluation body for AI algorithms in health, including embedded software (centralization of skills) if possible European, including aspects of certification / validation of systems.

- Integration of specific expertise.

- An evaluation process that is no longer "photographic" but evolutionary (“versioning”). Thus, in files submitted with a dimension of IA, the initial algorithm would provide entry of

tracking keys “post-registration” with clinical evaluations in parallel (and not ex post) not to constrain the development of evolving AI technologies

### **An adapted legal framework that requires some adjustments and clarifications**

An adapted legal framework

The legal framework for the use of health data for research purposes seems to be well-defined with regards to the notion of ownership attached to data, and more specifically to databases, thanks to the existence of harmonized European legislation. This legislation recognizes the rights of database producers and the capacity they have at their disposal to grant exploitation rights on these bases to the benefit of various users.

Other harmonized legislation at European level is applicable to AI in order to guarantee the fundamental rights of European citizens:

- First and foremost the protection of privacy and therefore personal data for each individual. The general data protection regulation (GDPR), which entered into force on 25 May 2018, pursues this objective beyond the borders of the European Union. Henceforth, any non-national operator implementing a data processing relating to persons who are on the territory of the Union and in particular from data making it possible to carry out a follow-up of the behavior of these persons which takes place within the territory of the Union, must abide by the RGPD.

- The protection of health is guaranteed by the supervision of health software that can be qualified as a medical device. New regulations, reinforcing the obligations of manufacturers and other players in the medical device sector, come into effect in 2020.

The next revision of the bioethics law should make it possible to take stock of existing legislation and developments that might be desired in order to regulate certain aspects of AI that are not yet covered by the legislation in force. This seems to be the wish of the Council of State in a report

made public on July 11, 2018 concerning the revision of the bioethics law [14]. This review will be based on groundwork carried out by organizations such as the Avicenna alliance (association for predictive medicine), a public-private partnership set up in 2016 and supported by the European commission.

Two weaknesses of the legal framework have been pointed out:

- There is indeed a gap between the European RGPD and the French *Commission nationale de l'informatique et des libertés* (CNIL) reference methodology (RM) 004, which regulates in France the processing of personal data for purposes of study, evaluation or research not involving human intermediation in the process (studies on data reuse). MR 004 imposes not only general information on the persons concerned (posting in the premises, mention in the welcome booklet, etc.) but also “*individual patient information included in the searches. It must be performed for each project in which the patient participates or for which the patient's data will be processed*”. A constraint that is not imposed by the European GDPR.

- The law dated 5 March 2012 [15] on research involving human intermediation (known as the Jardé law) is based on the notion of risk. It creates different categories of “research involving human intermediation” (RIPH) according to a gradual level of risk exposure, all of them, even those non-interventional or observational “that involve no risk or constraint in which all acts are performed and products used in the usual way” [16] to be submitted to the favorable prior opinion of an institutional review board.

This may seem anachronistic insofar as these searches are not, in fact, practiced on the human being but made from data which is legally covered by the protection of personal data. The Institutional Review Board also, lacking specific skills, may be ill-equipped in the face of the protocols that they must yet assess.

## **Recommendations**



- Alignment of the patient information obligations of the French legal framework regulation with the European GDPR for the processing of personal data for purposes of study, evaluation or research not involving human intermediation (studies on reuse of data).

- Clarification of the regulation of “RIPH” (research involving the human person) to be made on clinical trials including the exploitation of data. Studies of data (collection and processing) should be framed only by data protection legislation.

Use of data and collection of consent of those from whom they originated

Since the “Huriet / Sérusclat” law of December 20, 1988 [17] which constitutes the basis for the current regulation of clinical research in France, any research is subjected to the obligation of data collection prior to any research, with the explicit and informed consent of any participant whether he is healthy or not.

The intrusion of artificial intelligence techniques into clinical research, with the massive use and aggregation of data of various kinds and origin for multiple purposes and at times that can be very far apart, upsets the bond of trust between a patient and her doctor stated in 1937 by Georges Duhamel. How, given this spatial and temporal distance, can we ensure a sufficiently qualified level of information to the sources of these data?

Another question relates to the the fact that the data are aggregated and anonymized. Does this allows to override the principle of explicit consent, to the extent the intended use does not harm the person and it serves the common good, i.e. improving medicine?

All of these constraints lead to the formulation of a number of recommendations to reconcile deployment and use while respecting patients as the sources of data.

## **Recommendations**

- Rather than explicit consent at each stage of data use by those from whom it originates, favour the principle of “no objection” or implicit consent to be collected at the initial stage of data collection, with, possibly, the establishment of a recourse authority to which anyone could turn to should they consider that their data were used against their will or for an unethical purpose.

- Develop patient awareness tools and promote the systematic consultation of local user committees in health facilities.

- Promote the sharing of files and the so-called “broken glass” practice allowing healthcare professionals involved in the follow-up of the management of the patient, outside the attending physician, to access a patient's data: this under the constraint of increased traceability and the duty of information given to the patient at his request on the who and why has her medical information been accessed, without any additional legal constraint.

### **The algorithm a black box to illuminate by the contribution of knowledge**

In AI, raw patient data has little intrinsic value for the prediction of the therapeutic benefit of a drug candidate by modeling and simulation. It is its integration with knowledge publicly available in scientific articles on the disease of interest - a more robust modeling material, and its mechanistic representation with mathematical equations, that data can produce actionable insights.

This requires to solve two technical difficulties:

- The auditability and interpretability of the algorithm, which in the context of AI, can reach a level of abstraction that can be difficult for humans to understand, making it difficult to decipher simulation results and disentangle causation from correlation.

- Making scientific knowledge easier to exploit - the volume of articles is staggering (30 million articles to date on PubMed), and it grows at a fast clip with a doubling every 10 years. Furthermore, the fact that most of these documents are in pdf format complicates the technical exploitation of this unstructured information. The resulting complexity makes it impossible for a single human brain to assimilate.

## Recommendations

- Use hybrid modeling approaches combining data and knowledge (compilation of existing literature)
- Encourage the emergence of open innovation platforms for knowledge sharing, such as GitHealth (<https://githealth.io/>)
- Impose responsible algorithms “by design”, that is to say integrating in their programs elements allowing the traceability and the explicability of the results as well as controls ensuring the absence of potential negative effects of the algorithms (bias , discrimination, value judgments).

### 3rd critical point: “Know-how and ability” - Availability of technical skills and ergonomic tools for data collection

Many preconceived ideas circulate about AI: today we are close to the saturation of available human resources and the shortage of skills and experts; there would be a delay in the adaptation of the educational system, in the establishment of quality vocational training for these new jobs. More specifically in healthcare, the sector lacks attractiveness, would face stark competition from the US tech giants (Google, Apple, Facebook, Amazon [GAFAM]), who already collaborate with “big-pharmas”, as well as the four Chinese dragons: Baidu, Alibaba, Tencent and Xiaomi. More dynamic, offering better pay, more attractive, they drain the labor market of its available resources, depriving other sectors - including healthcare, thus slowing the pace of adoption of AI in the rest of the economy.

In fact, the situation does not corroborate this very dark picture. The situation is more complicated and more contrasted.

The AI is a field in itself which, despite its recent and still emerging nature, has branched into technical specialties, be it for the collection, management and exploitation of data, for the

development and computer programming of tools and processing of data. These new data science jobs are supposed to operate at the interface with doctors, the gatekeepers of medical science but also in direct contact with patients who provide the raw material, their data, to fuel these tools.

Regarding training of these data scientists, France holds its ground against other developed nations, far from the backwardness sometimes described and that one could fear. In particular, there are training courses for AI experts in leading engineering schools, such as at the *Institut Mines-Télécom* (Telecom Paris Tech, Telecom Sud Paris, IMT Atlantique, ...), or within the Dataia institute (data sciences, artificial intelligence and society) and its members (*École polytechnique*, *Centrale Supélec*, *ENS Cachan*, *Ensaé Paris Tech*, university Paris Sud, etc.). The constitution of the “Data” department of the Institut Curie was made without major recruitment difficulties and with a limited turnover.

On the other hand, there is a lack of multidisciplinary training ensuring the necessary dialogue between “doctors / engineers / computer scientists” expertise in healthcare institutions (whereas multidisciplinary approaches to data science and the human and social sciences are at the heart of some training programs - see Dataia and Institut PRAIRIE), the situation being aggravated by the inadequacy of the apparatus for data collection and interfacing in hospitals, either because they have an administrative and financial vocation or because they did not take into account user-specific ergonomics. But some engineering schools, such as Centrale Lyon for example, have developed a master's degree in health engineering to bridge these gaps.

With regards to health facilities, the structuring of medical records is notoriously insufficient, making it difficult to extract and use meaningful scientific and medical information which is contained in medical records, most often archived in pdf format.

The conversion of medical records in computer-readable format remains low in hospitals, with a still significant use of hard copies, hampering the exploitation of patient data by AI algorithms.

These difficulties reflect insufficient appropriation of IT tools by healthcare professionals, which originates from insufficient ergonomics and complex settings requiring the use of specialized services. They are the result of the often old and rigid architectures of integrated management software packages that are found in most university hospital centers, based on “client-server” platforms that have most often not completed their migration to web-based architectures. The latter provide for lighter, more ergonomic, more cost-efficient and above all more adaptable environments

for healthcare professionals, whose cultural needs must be better respected to ensure quality medicine, for which computer science should not impose any functional regression.

With regards to family practice, the decreasing density of practitioners requires an increased optimization of consultation time. AI could also allow doctors to gain valuable time by allowing her to have an analysis of his patient and focus on the most at-risk patients. In this space, too, the multitude of software publishers can hinder coordination efforts between family practices and hospitals.

### **Five recommendations are formulated**

- Encourage the establishment of bridges between engineering schools and medical curricula and foster the emergence “industry - hospital / engineer – clinician” pairs for entrepreneurial projects such as Sim & Cure, Therenva and Predisurge, three French start-ups working in the field of cardio and endovascular surgery simulation.

- Promote and develop inter-sectoral, equitable and arm's length partnerships for sharing skills (eg GAFAM / health industries / health institutions / research organizations).

- Ensure a sufficient number of specialized training courses, such as “Data Science”, including a “Health” dimension in these offerings to ensure an adjustment of demand and supply.

- In hospitals: reverse the top-down paradigm of equipment in data collection tools to enable health actors to better define the needs, functionalities and settings of the files used, in order to facilitate statistical analysis for clinical and research purposes.

- Evaluate (in particular during certification processes) and fund health information systems based in particular on:

- o 1. The quality of information structuring (necessary for quality statistical exploitation, especially in clinical research).

- o 2. The effectiveness of the switch to digital files and the possibility of gradually eliminating paper archives, both to reduce costs but also - and above all - to improve quality by

making the information immediately available, usable and accessible, shareable to the actors in the coordination of care in family practice or in the patient's home.

#### **4th critical point: getting ready to act: organization and governance**

##### **Organizational transformations to realize**

The organizational impact of AI in clinical research has not been modeled. For healthcare companies, hospitals and research and development organizations, which will be the main users of artificial intelligence, this impact is far from being neutral. It has not yet been thought, theorized and anticipated.

##### **The integration of AI in the R&D departments of pharmaceutical companies**

“Big-pharmas” are multinational entities which have grown in size and complexity over the past 20 years, even if their plasticity has been put to the test by successive waves of mergers & acquisitions and the transition from the "chemical" paradigm to the “biological” one. While AI represents a new mutation of their organizations, it is also a remarkable opportunity for growth and positive change, along the lines of what has been observed in other industries (technology, automotive, aeronautics, energy). At the crossroads of AI, three branching paths of integration and adaptation are offered to large pharmaceutical companies (biotech and medtech companies’ situation differ to the extent that they can build their organization from the ground up around these new approaches):

- Creation of new functions identified in addition to existing departments.
- Integration of these new functions into existing workflows.

- Outsourcing the function through the use of partners, technology-driven small and medium size companies specializing in AI, knowing that even in this case, this implies, internally, for laboratories, the creation of new functions and expertise to establish and maintain a link between these external providers and internal capabilities.

Whichever solution is chosen, AI initially results in an investment in expertise and therefore an additional cost that can be offset by reformatting the existing R&D operations, which consist for the most part of PhDs, doctors, pharmacists, engineers, biologists and veterinarians. In all cases, existing employees must be prepared to the profound changes in their organization in order to seize the opportunities that may open up to them in terms of professional development.

The use of *in silico* clinical trials will also have consequences on pre-clinical phases, those on animal models, some of which may be at least partially substituted by these new approaches.

### **The recommendations are intended to ensure**

- Realization of studies on the impact of AI in the management of large pharmaceutical organizations (*École supérieure de systèmes de management des organisations* [ESSYM], *École des hautes études en santé publique* [EHESP]...).

- The full functional integration of AI in the R&D processes and the preparation of organizational adaptations.

- Support for existing staff to adapt to this new normal.

## **Health establishments (“Établissements de Santé” or ETS), collectors and users of data:**

Places of care, training, research, health institutions, public hospitals, cancer research centers, private clinics, all acquire a 4th vocation, as (1) a source of collection and (2) a contributor to research on AI (see the 1st AI and health symposium organized by the Nancy university hospital on November 21, 2018, or some research projects based on AI funded by *Agence nationale de recherche* [ANR] or *direction générale de l’offre de soins* [DGOS]) and as (3) end-users of AI algorithms.

Ultimately, the goal of the approach is to provide the patient and those who care for him, doctors and care teams, tools and indicators that they could not have had without the contribution of AI. What we summarize through the formula: “The augmented practitioner”.

Health facilities, data collectors: If they become essential pillars of the system, this implies that they have the capacity to meet the challenge in terms of collecting, hosting, indexing and exporting data in an interoperable format, each of these dimensions constituting a singular imperative.

Owkin- and Bpifrance-supported “Substra” collaborative projects involving eight other partners, including four healthcare institutions, seem to be leading the way. While hospitals may be reluctant to share their data, the “open source” platform that will be operational in 2019 uses a technique of relocated servers in each data warehouse between which deep learning algorithms travel. The data is neither moved nor shared.

### **At the end of the chain, for the routine use of AI by medical teams embedded in health establishments (ETS), it is recommended**

- To integrate AI and necessary equipment into institutional health establishments strategies.
- To optimize the management of chronic diseases, automated remote control via connected patients and the possibility of administering certain routine or emergency care by intelligent equipment, through the upstream monitoring of risk factors and risk of relapse.



- Integrate epigenetic, behavioral and environmental factors to define, with the help of AI, a new taxonomy (holistic approach to patient management).

- To delineate between management tools - that can produce medical information and that are inherently heterogeneous owing to the multitude of software disseminated in France in those ETS - with the communication, consultation and file sharing tools which must be user-friendly and universal, based on web platforms and / or accessible through tablets or smartphones.

- To encourage the implementation in ETS of modern web-based architectures to facilitate the aggregation of components according to a national framework of interoperability, allowing the structuring of information and massive digitization, the integration of AI tools, including intelligent conversational agents and text analytics agents.

## **Towards convergence of health networks to improve coordination of care, the role of AI:**

While health technology agencies (HTAs) represent the major source of medical information from a quantitative point of view, care coordination cannot be envisaged without the convergence of different health actors at the local, regional and national levels: public and private medical institutions and medico-social, liberal medicine, pharmacies, medical analysis laboratories and imaging centers in particular.

However, compounding the aforementioned difficulties (ergonomic software to structure the information, digitalize it, make it interoperable according to a framework of interoperability defined at the national level), is the absence of convergence between platforms defined at local, regional and national level:

- GHT (territorial hospital groups, purely public structures excluding other local actors in their organizational model).
- Regional platforms set up by the ARS through various projects such as TSN (digital care territory).
- National platforms including the communicating cancer dossier or the DMP (shared medical record, whose revival is expected to be announced soon at national level).

Urbanization and the logic of convergence of these different platforms appear difficult to understand for users, a reflection of misalignment issues in the governance of these different projects. It is impossible to envisage that all the information produced in all the regions is hosted within the DMP, nor that all the information produced in the different localities of a region be hosted within a regional platform.

## **The technical solution exists and could be deployed throughout the ecosystem**

- Create interoperability platforms (IFPs) through which information is passed from one point (information producer) to another (primary information host) securely and traced.

- Transmit to the different interconnected hosts, from these IFPs, URLs rather than physical documents allowing access to documents (depending on prespecified access rights) within the approved primary hosting structures. This will allow the interconnection of different health hosting "clouds" in a distributed logic, governed within a common framework of interoperability. This architecture can use blockchain technologies to ensure the traceability of data usage.

Finally, if interconnected “clouds” manage to host health information in “big-data” mode, the nature of the human brain, the time constraints, the multiplication and the legally binding nature of institutional recommendations, will make it impossible to use devices proposed to health actors and patients in the absence of powerful agents integrating AI.

### **The recommended areas for improvement are fourfold**

Enable the convergence of health platforms, by setting up a strong and unique identification counter for healthcare professionals and patients, from which a simple navigation without re-identification will be possible between the different local, regional and national platforms, depending on the need for the user to access the information, and her access rights, integrating traceability and communication of who has requested access to them.

- Integrate, without re-identification within shared medical file access platforms, the services offered to the patient and / or the caregiver and information from the connected devices that will have been made available to patients, in particular for home monitoring or as part of research protocols.

- Migrate current centralized hosting architectures to a distributed, interoperable and secure architecture using interoperability platforms that can be secured by blockchain technologies.

- Deploy and validate intelligent conversational agents and AI tools, including machine learning and text analytics, to facilitate synthesis and understanding of medical records and to integrate institutional recommendations.

*AI therefore appears as the inevitable outcome of processes, provided that improvements are made to these architectures, implying a change in the governance of healthcare systems.*

### **National governance of the system**

Three factors encourage thinking at the national level the deployment of AI in the healthcare pathway:

- Letting go would lead to the disorderly juxtaposition of “silo” initiatives, overlapping or even duplicating which will be source of financial mismanagement and incommunicability between these tools.

- The shared medical record (DMP) will enter into its widespread deployment phase with, for the French *Caisse nationale d'Assurance maladie des travailleurs salariés* (CNAMTS), a hosting goal of 10 million medical records at the end of 2018 and 40 million in 2022. If the DMP does not can claim to integrate all health information produced at the national level, it aims to provide healthcare professionals and patients an immediately accessible medical summary that will ultimately rely on AI tools. It will also be necessary for specialized needs to navigate in a seamless way, without identification, from the DMP to the regional or local platforms containing the necessary, more exhaustive information.

- The coming years will see a proliferation of AI-based solutions that require evidence of efficiency and effectiveness that can only be achieved by soliciting an information system with foolproof resilience.

## **Given these factual elements, four recommendations emerged**

- Because the coordination of care takes place at the local level, rely on the territorial professional communities of health (CPTS) and territorial support platforms (PTA) established by the Law of modernization of our healthcare system voted in 26 January 2016
- Define a distributed and secure model to converge the information and services that benefit the patient
- Which implies on a technical level:
  - o For regional platforms and the DMP to have access without re-identification to the medical information hosted by the approved hosts of the locality (choice to be made)
  - o Importance of relying on operational demonstrators (to qualify)
- Extend the purposes of the DMP to public health to allow collective exploitation for research purposes (legislative avenue to be identified)

## **Conclusion**

“Artificial intelligence will not replace doctors. But the doctors who will use AI will replace those who will not do it”. This is the catchy title of a recent article by Xavier Comtesse, mathematician and creator of several start-ups, and Daniel Walch, director of groupement hospitalier de l’Ouest lémanique (GHOL) [18]. Because the march of innovation is irresistible and does not wait, if France misses the train of AI, it will be relegated definitively, tomorrow, to supporting roles. In clinical research, the damage would be incalculable for both patients deprived of early access to innovation and for actors of the system demoted to mere providers of services and no longer initiators and promoters of scientific advances.

AI in clinical research is a chance for the healthcare system if all its actors, patients, professionals, drug developers and especially regulatory agencies rise up to the challenge and build an ecosystem conducive to its proper use.

This new ecosystem will first impact mindsets, information systems architectures, organizations, governance, which implies an effort in pedagogy, awareness and acculturation, mastery of technology and the appropriation of adapted tools. However, if France will probably struggle to win against the US and China in terms of devices intended for the general public shipping with AI embedded into them, especially if it engages alone while the action must be taken collectively at the European level, it has the skills and know-how to play a leading role in research, professional niches and especially software and algorithms. Healthcare represents a major opportunity to develop this emerging sector, provided that the necessary changes in the architecture of healthcare networks and governance are made.

With the “health data hub”, one of the main axes of the French Artificial Intelligence strategy, whose construction originated in the “prefiguration” mission report delivered to the minister of health on October 12, 2018, the effort is launched, and is the perfect vehicle for proposals and recommendations from the group of experts gathered around this round table.

### **Disclosure of interest**

Authors have no competing interest to declare.

### **References**

[1] Thomas DW, Burnes J, Audette J, Carroll A, Dow-Hygelund C, Hay M. Clinical development success rates 2006-2045. June 2016. <https://www.bio.org> [Accessed 16 november 2018 (28 pp.)]

- [2] Hwang TJ, Carpenter D, Lauffenburger JC, Wang B, Franklin JM, Kesselheim AS. Failure of investigational drugs in late-stage clinical development and publication of trial results. *JAMA Intern Med.* 2016 Dec 1;176(12):1826-1833. doi: 10.1001/jamainternmed.2016.6008.
- [3] Dhainaut JF, Huot L, Bouchara Pomar, Dubray C, Augé P, Barthélémy P, et al. Utilisation des objets connectés en recherche clinique. *Thérapie* 2018 Feb;73(1):41-51. doi: 10.1016/j.therap.2017.11.002. Epub 2017 Dec 11.
- [4] Pharnext. Pharnext annonce les résultats positifs de l'essai pivot de Phase 3 de PXT3003 pour le traitement de la maladie de Charcot. 2016. <https://www.pharnext.com/fr/actualites/communiqués-de-presse> [Accessed 16 november 2018]
- [5] Simm J, Klambauer G, Arany A, Steijaert M, Wegner JK, Gustin E, et al. Repurposing high-throughput image assays enables biological activity prediction for drug discovery. *Cell Chem Biol* 2018 May 17;25(5):611-618.e3. doi: 10.1016/j.chembiol.2018.01.015. Epub 2018 Mar 1.
- [6] Kogan E, Twyman K, Heap J, Milentijevic, D, Lin JH, Alberts MJ. Use of machine learning to determine stroke severity of patients diagnosed with stroke in integrated claims-medical records dataset. *Circulation* 2018;136:A15849. [https://www.ahajournals.org/doi/abs/10.1161/circ.136.suppl\\_1.15849](https://www.ahajournals.org/doi/abs/10.1161/circ.136.suppl_1.15849) [Accessed 16 november 2018]
- [7] Avicenna Alliance. Association for predictive medicine. Conference report. International Avicenna Alliance conference report. 4th September 2018. European Parliament (Brussels). <https://www.avicenna-alliance.com/conference-2018/conference-report/> [Accessed 16 november 2018 (16 pp.)]
- [8] Rocha JC, Da Silva DLB, Dos Santos JGC, Benham Whyte L, Hickman C, Lavery S, et al. Using artificial intelligence to improve the evaluation of human blastocyst morphology. ESHRS congress, Genova, Switzerland. Juillet 2017. [www.assis.unesp.br/Home/PosGraduacao/Biociencias/artigo-15.pdf](http://www.assis.unesp.br/Home/PosGraduacao/Biociencias/artigo-15.pdf) [Accessed 16 november 2018 (6 pp.)]
- [9] Sonigo C, Jankowski S, Yoo O, Trassard O, Bousquet N, Grynberg M, et al. High-throughput ovarian follicle counting by an innovative deep learning approach. *Sci Rep* 2018 Sep 10;8(1):13499. doi: 10.1038/s41598-018-31883-8.
- [10] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blaud HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115.

- [11] Odoxa. Le rendez-vous de l'innovation : l'intelligence artificielle emballe les « gagnants » du système mais fait peur aux autres. Mai 2016. <http://www.odoxa.fr/sondage/rdv-de-linnovation-lintelligence-artificielle-emballe-gagnants-systeme-peur-aux-autres/> [Accessed 16 november 2018]
- [12] Health care data institute. Données de santé : les français sont prêts à les partager ! 16 novembre 2017. <http://www.odoxa.fr/wp-content/uploads/2017/11/CP-HDI-Odoxa-16nov17.pdf> [Accessed 16 november 2018]
- [13] European patent office. Guidelines for examination in the european patent office. Novembre 2018. <https://www.epo.org/law-practice/legal-texts/guidelines.html> [Accessed 16 november 2018]
- [14] Conseil d'État. Révision de la loi de bioéthique : quelles options pour demain ? 11 Juillet 2018. <http://www.conseil-etat.fr/Decisions-Avis-Publications/Etudes-Publications/Rapports-Etudes/Revision-de-la-loi-de-bioethique-quelles-options-pour-demain> [Accessed 16 november 2018]
- [15] Legifrance. Loi n° 2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine (1). Version consolidée. 16 novembre 2018 [https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000025441587\\_](https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000025441587_) [Accessed 16 november 2018]
- [16] Legifrance. Ordonnance n° 2016-800 du 16 juin 2016 relative aux recherches impliquant la personne humaine. JORF n°0140 du 17 juin 2016 <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000032719520&categorieLien=id> [Accessed 16 november 2018]
- [17] Legifrance. Loi n° 88-1138 relative à la protection des personnes qui se prêtent à des recherches biomédicales. 20 décembre 1988. Consolidée le 16 novembre 2018. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000508831> [Accessed 16 november 2018]
- [18] Walch D, Comtesse X. L'intelligence artificielle ne remplacera pas les médecins. Mais les médecins qui utiliseront l'IA remplaceront ceux qui ne le feront pas. Rev Med Suisse 2018;14:1728. <https://www.revmed.ch/RMS/2018/RMS-N-620/L-intelligence-artificielle-ne-remplacera-pas-les-medecins.-Mais-les-medecins-qui-utiliseront-l-IA-remplaceront-ceux-qui-ne-le-feront-pas> [Accessed 16 november 2018]