

# Big Data and Public Health Systems: Issues and Opportunities

David Rojas de la Escalera<sup>1</sup>, Javier Carnicero Giménez de Azcárate<sup>2\*</sup>

<sup>1</sup> Business Development Senior Consultant – eHealth Division, Sistemas Avanzados de Tecnología (SATEC) (Spain)

<sup>2</sup> Health Service of Navarre (Spain)

Received 15 July 2016 | Accepted 3 February 2017 | Published 27 March 2017



## ABSTRACT

Over the last years, the need for changing the current model of European public health systems has been repeatedly addressed, in order to ensure their sustainability. Following this line, IT has always been referred to as one of the key instruments for enhancing the information management processes of healthcare organizations, thus contributing to the improvement and evolution of health systems. On the IT field, Big Data solutions are expected to play a main role, since they are designed for handling huge amounts of information in a fast and efficient way, allowing users to make important decisions quickly. This article reviews the main features of the European public health system model and the corresponding healthcare and management-related information systems, the challenges that these health systems are currently facing, and the possible contributions of Big Data solutions to this field. To that end, the authors share their professional experience on the Spanish public health system, and review the existing literature related to this topic.

## KEYWORDS

Big Data, Health System, Healthcare Organizations, Health Information Systems, Epidemiological Surveillance, Strategic Planning.

DOI: 10.9781/ijimai.2017.03.008

## I. INTRODUCTION

### A. The Health System

According to the World Health Organization (WHO), “a health system consists of all organizations, people and actions whose primary intent is to promote, restore or maintain health. This includes efforts to influence determinants of health as well as more direct health-improving activities. A health system is therefore more than the pyramid of publicly owned facilities that deliver personal health services” [1]. Furthermore, every health system performs the following set of basic functions [2]:

- Delivering health services to individuals and to populations.
- Creation of resources.
- Stewardship.
- Financing the system.

The center of any health system must be the first of these functions, since healthcare constitutes the paramount goal and therefore the reason for the existence of the health system itself. Around it, other functions are organized, essential for ensuring healthcare delivery and public health. Among these, the following ones must be remarked:

- Epidemiological surveillance, which comprises the collection and analysis of large volumes of data directly or indirectly related to people’s health, so as to detect or prevent possible health problems regarding public health.
- Planning and overseeing the management of the health system,

which allows healthcare organizations to set out their strategic goals, allocate the necessary resources, assess the degree of compliance of these goals and apply corrective measures if required.

- Clinical research, focused on generating knowledge and applying it to the development of new diagnostic and therapeutic techniques.
- Education and teaching, in order to train new professionals and keep the practicing ones appropriately updated and competent.

### B. The Health Cluster or Ecosystem

From a structural point of view, a health system is neither an isolated nor homogeneous entity, but it comprehends or relates to entities of diverse nature, both public and private, with interests of their own as well as shared interests. This ensemble is known as health cluster or ecosystem, and among its components the following ones must be pointed out:

- Central or federal government and regional or local authorities.
- Healthcare services, conceived as organizations responsible for the management of a determined healthcare network.
- Hospitals.
- Primary care centers.
- Emergency services.
- Pharmacies.
- Convalescent centers.
- Health professionals acting as external providers to the health system.
- Public health services.
- Insurance companies, mutual societies and other entities which finance healthcare.
- Schools for the education and training of doctors, nurses and other

\* Corresponding author.

E-mail address: javier.carnicero.gimenez@cfnavarra.es

health professionals.

- Research centers.
- Professional associations and colleges.
- Foundations and learned societies.
- Stakeholders, such as patients’ associations.
- Pharmaceutical and other health technologies industries.

**C. Challenges Faced by the Health System**

For decades, the public health systems of European countries, created following the end of World War II, have been frequently mentioned as a reference model to be followed, especially in those aspects regarding coverage, quality of service and contribution to the welfare of society. However, the scene in which these systems arose has suffered a series of major changes, being the most important the following ones [4]:

- The aging of the population, with a continuous increment of chronic and degenerative diseases.
- The financial crisis, which causes important budget cuts in the public funds meant to finance the health systems activities, and makes it more difficult –or even impossible– for the citizens to compensate these cuts with out-of-pocket expenses.
- The creation of new techniques and drugs, more effective but also more expensive, mainly due to the necessity to compensate the research costs caused by their development.
- The increasing demands of the citizens, who require more and better healthcare services in a setting that seeks patient empowerment and promotion of personalized medicine.

As a token of the first two determinants, aging of the population and public budget cuts, the Spanish case is addressed below. Table I shows the progress of these two indicators during the period between 2003 and 2014.

These data reveal that the Spanish population has increased from 42.72 to 46.77 million people during the 2003-2014 period, while the percentage of people older than 64 years has risen from 17.03% to 18.05% over total population, and the dependency ratio, which indicates the ratio between population older than 64 years and population between 15 and 64 years old, has risen from 24.75% to

26.99%. On the other hand, public health expenditure in 2003 meant 5.37% GDP, reaching a peak of 6.77% in 2009 and falling to 6.26% in 2013, experiencing a small recovery in 2014, with 6.34% GDP. Regarding private health expenditure, it was at minimums around 2.14%-2.17% GDP but it has risen year after year since the beginning of the financial crisis in 2008, reaching 2,74% GDP in 2014.

All things considered, the impact of all these determinants is so important that the sustainability of this model of public health system has been questioned in recent years.

**D. The Transformation of Public Health Systems**

Despite the fact that the challenges explained above make clear that a deep transformation of this health system model is needed, and IT is often considered as one of the main facilitators for this change, it is not admissible to think that health systems are going to lose their essential features. Health systems must improve people’s health, from both an individual and a collective point of view, and this final goal will not change in spite of the introduction of new technologies such as Big Data.

The patient must always be the centre of any health system and, in the same way, health information must always be the centre of a health information system, which will be introduced below. The actions of a clinic professional focus on the achievement of specific healthcare goals customized for each one of their patients –improving or keeping their health status–. Besides knowledge, healthcare requires a connected and personalized relationship between the provider and the patient, so that interventions are tailored to the patient’s unique preferences and behaviour as, for instance, drug adherence. Different people will have different reasons for non-adherence [5].

On their behalf, health systems managers must seek the compliance of the general goals defined by their organizations. These goals will be the aggregate of the individual goals related to each one of the professionals in their clinical staff. In addition, these managers will also be responsible for the allocation of the necessary resources and the financing of the whole activity in their organizations.

On the whole, health systems must focus their efforts on the creation of value for both the patient and society. To that end, clear goals must be defined that find an appropriate balance between the patient’s personal interests and the collective interest of society. For instance, in the event

TABLE I. DEMOGRAPHICS AND HEALTH EXPENDITURE IN SPAIN (2003-2014)

Year	Total population	Population 15-64 yrs.	Population older than 64 yrs.		Dependency ratio	Public health expenditure		Private health expenditure	
			People	% over total		M€	% GDP	M€	% GDP
2003	42,717,064	29,396,965	7,276,620	17.03%	24.75%	43,158.4	5.37%	17,354.5	2.16%
2004	43,197,684	29,777,965	7,301,009	16.90%	24.52%	46,992.4	5.46%	18,651.1	2.17%
2005	44,108,530	30,511,110	7,332,267	16.62%	24.03%	51,351.5	5.52%	20,094.2	2.16%
2006	44,708,964	30,849,177	7,484,392	16.74%	24.26%	56,662.2	5.62%	21,520.7	2.14%
2007	45,200,737	31,188,079	7,531,826	16.66%	24.15%	61,612.0	5.70%	23,101.9	2.14%
2008	46,157,822	31,869,008	7,632,925	16.54%	23.95%	68,147.1	6.11%	24,392.9	2.19%
2009	46,745,807	32,145,023	7,782,904	16.65%	24.21%	73,035.6	6.77%	23,863.0	2.21%
2010	47,021,031	32,153,527	7,931,164	16.87%	24.67%	72,852.6	6.74%	24,593.7	2.28%
2011	47,190,493	32,082,758	8,093,557	17.15%	25.23%	71,800.0	6.68%	25,510.2	2.38%
2012	47,265,321	31,980,402	8,222,196	17.40%	25.71%	68,262.9	6.47%	26,594.3	2.55%
2013	47,129,783	31,718,285	8,335,861	17.69%	26.28%	65,718.5	6.26%	26,981.3	2.62%
2014	46,771,341	31,281,943	8,442,427	18.05%	26.99%	65,975.7	6.34%	28,558.1	2.74%

Sources:

- Demographics: Demographic Information System, Spanish National Statistics Institute (INE).
- Health expenditure data: OECD Health Statistics.

of a surgical intervention it is mandatory to measure some indicators such as mortality rate, adverse events, time of recovery, care costs, or time for the patients to return to their jobs at full capacity. Nevertheless, it is necessary to also take into account other indicators, maybe more subjective and thus harder to measure, but equally important because of their impact on the patient, such as post-surgery functionality, pain suffered, or the cost of all these factors from a quality-of-life point of view. If health systems are not focused on their patients' interests and on achieving the corresponding goals, they will hardly be able to change and ensure their sustainability [6].

This article reviews the main features of the European public health system model and the corresponding healthcare and management-related information systems; the problems and challenges that these health systems are currently facing; and the solutions that Big Data tools may potentially offer in that respect. To that end, the authors have based this work on their professional experience on the Spanish public health system, an analysis of the scene that the latter is facing in the upcoming years and decades, and a review of the existing literature on Big Data applied to health.

## II. BIG DATA SOLUTIONS IN A HEALTHCARE ENVIRONMENT

### A. The Health Information System

The individual performance of the different components of the health cluster, as well as their interactions, causes the creation of multiple data flows, which are also greatly varied, since they involve several business processes. This set of data flows gives the health information system as a result.

As mentioned above, just as healthcare is the centre of any health system, patient-related healthcare information must be the centre of the health information system as well, since it may and will also be used for activities other than healthcare, such as epidemiological surveillance, planning, overseeing of the management, clinical research, and education and training, as stated in the "Introduction" section. The fact that these data are stored in healthcare information systems is a consequence of them being generated during the patient's care, but their usefulness goes clearly beyond this limit.

Therefore, the health information system must allow users to register, process, consult and share large amounts of data, ensuring their availability at the appropriate moment and point of the health cluster. On the healthcare side, this cluster reveals itself as a huge generator and at the same time consumer of enormous sets of information, related to personalized healthcare processes that take place on a daily basis and in a massive way. Healthcare is considerably intense regarding data treatment, by constantly creating immense datasets and frequently requiring access to knowledge sources.

For IT to be fully integrated in the health system value chain, it is mandatory to have a health information system which serves as an instrument for knowledge management being useful to all its users. Healthcare professionals cannot perform their duties properly without registering and using patients' information, or without accessing the knowledge sources that allow them to make decisions on a solid basis. Public health departments need to know the population health status in order to detect or prevent potential collective health issues, as well as defining the necessary corrective and preventive measures. To those ends, these professionals must rely on data generated during every patient's individual healthcare, properly aggregated, as well as other data sources.

Managers are not able to plan a strategy, oversee its performance and assess the achieved outcomes without a tool that allows them to process all the necessary information and provides them with accurate data, timely and in due form. These data are required from the very beginning, since the definition of an appropriate strategy must be based

on the knowledge of the population's health status, complemented with projections of its potential progress.

This complexity has been increased in recent years by a major change in healthcare organizations, which have evolved from a clearly paternalist way of interaction with their patients to another one completely different, focused on seeking their empowerment. In addition, patients are not content anymore with the information provided to them by their doctors, but search the Internet for additional data about their diseases, engage on social networks, make their own decisions, and register information on their health records. This new role is indeed required if, for instance, health authorities seek to promote one of the most important lines of action in the field of chronic patient management, self-care encouragement, which has a beneficial impact on both the patient and the health system. However, this requires also a more varied interaction between them, combining traditional simple events, like setting up appointments with a general practitioner, with more complex actions, like monitoring health data measure and stored by wearable devices.

Despite this, it is clearly positive that both society and the medical community have evolved from a discussion about giving patients clearance to access their own healthcare information, to a totally different one about seeking the best way for the patients to register data in their health records, either in an active and conscious way or in an passive and automated one via specific devices. In any case, it must be always taken into account that the management of healthcare information is not a process unrelated to healthcare, but an inseparable component of healthcare itself, hence its management and supervision are the healthcare professionals' responsibility, even though the patients take a more active role. Furthermore, every professional must accept this new reality and provide the patient with the necessary training, so that this initiative ends up being successful [7].

Apart from this, the temptation of exploiting the information stored on the different social networks turns out to be very powerful. It is true that, to the health system, it is a possibility worth exploring, but several conditioning factors must also be considered. The first one is data protection as a consequence of people's right to privacy, something that, from the very first moment, seems to collide clearly with the business model of social networks themselves, designed to share large amounts of information in a quickly, heterogeneous and, up to some point, uncontrolled way.

Precisely these features represent another important conditioning factor, since social networks are nothing but huge repositories that store unstructured, poorly classified or simply uncategorized data, not to mention the more than likely irrelevance of most of them regarding healthcare and, moreover, their doubtful veracity, a feature essential to this field. Anyway, given their market penetration, with millions of users around the world, it seems advisable to assess the possibility of using social networks as an information source for health systems, as long as a model can be defined that solves or at least mitigates all the inconveniences mentioned above.

### B. Potential Contributions of Big Data to Health Systems

The field of Big Data analytics is rapidly expanding, up to the point that it has begun to play a main role in the evolution of healthcare practices and research, by providing tools to register, manage, and analyse huge amounts of both structured and unstructured data produced by current healthcare information systems [8].

Health-related Big Data streams can be classified into three categories [5]:

- Traditional healthcare data are generated within the health system and stored in datasets such as health records, medical imaging tests, lab reports or pathology results, among others. Analysing this information allows to achieve a better understanding of disease

outcomes and their risk factors, and also to reduce health system costs, thus making them more efficient.

- “Omics” data deal with large-scale datasets in the biological and molecular fields, such as genomics, microbiomics or proteomics, for instance. The study of this information leads to deeper knowledge about how diseases behave, in order to accelerate the individualization of medical treatments.
- Data from social media allow to figure out how individuals or groups use the Internet, social media, apps, sensor devices, wearable devices or any other tools, to better inform and enhance their health.

In addition, the inclusion of geographical and environmental information may further increase the ability to interpret gathered data and extract new knowledge [11][12].

Combinations of several types of data must also be taken into account. The concept of personalized medicine, partially introduced above, seeks to combine the patient’s health record and genomic data in order to support the clinical decision-making process, making it predictive, personalized, preventive and participatory, an idea known as “P4 Medicine” [9].

At the micro level, personalized medicine aims to customize the diagnosis of a disease and the subsequent therapy by taking into account the individual patient’s characteristics, instead of relying on decisions taken according to general guidelines, defined as a result of population-based studies and clinical trials. This will require the integration of clinical information, mainly patient records, and biological data such as genome or protein sequences. These data are generated from different and heterogeneous sources, and have very diverse formats [9].

In fact, healthcare data no longer needs to be restricted to traditional datasets such as electronic health records. For instance, mobile or wearable devices monitoring physiological signals can provide timely access to multiple data points that are increasingly interconnected. Traditionally, the data generated by this sort of devices have not been stored for more than a brief period of time, being discarded afterwards and therefore preventing any extensive investigation to benefit from the exploitation of these data. However, attempts to use this kind of datasets have been increasing lately, in order to improve patient care and management [8][10].

Nevertheless, there is a difference between collecting data, having access to data, and knowing how it should be used to improve healthcare. Now that the technology for handling massive amounts of data is available, the next step is developing tools for information sharing and knowledge management, which are seriously limited by the lack of system interoperability [9][10].

For instance, with full interoperability the ability to collect data in a timely manner from several different sources leads to an increase in registries. Disease registries are still in an early stage, but they might be valuable tools when it comes to supporting patient-centred self-management of chronic illness and defining customized treatment plans. Besides, the integration of computer analysis with appropriate care will help doctors to improve diagnostic accuracy. In a similar way, the integration of medical images with other types of electronic health record data and genomic data can also improve the accuracy of a diagnosis and reduce the time required for it [8][10].

A major emphasis of personalized medicine is to match the right drug with the right dosage to the right patient at the right time. Moreover, gene sequencing and the use of the subsequent genetic data in diagnosis and treatment will be essential to the future of personalized medicine, with actions such as the prescription of drugs based on genomic profiles of individual patients, known as pharmacogenomics. However, analytics of high-throughput sequencing techniques in genomics is a problem inherent to Big Data itself, since the human

genome consists of 30,000-35,000 genes. Some ongoing projects aim to integrate clinical data from the genomic level to the physiological level of a human being. These initiatives will surely help when it comes to deliver personalized healthcare [8][9][10].

At the macro level, faster access to data allows any hospital to define and apply quality improvement policies based on the constant monitoring of outcomes, so as to ensure that the strategic goals of the organization are achieved. Hospitals have also used electronic health records, datasets originally intended to document individual healthcare processes, to identify system-related inefficiencies and quality issues. Faster access to data has also been hugely useful for the identification and management of disease outbreaks, allowing public health initiatives to be targeted to specific areas, as a result of population analysis [10].

The mining of electronic health record data made possible for researchers to identify possible sources of adverse events. Healthcare professionals used this information to improve organizational practices and reduce error rates. Moreover, many clinical information systems such as electronic health records and computerized physician-order entry systems capture a large amount of metadata about their use, which can be used for auditing purposes, thus allowing the organization to detect user-device interaction problems, shrinking safety margins and other technology-related safety issues and concerns, before any adverse event takes place [10].

The potential impact of Big Data is not easy to estimate, let alone on such an early stage. A report sponsored by the McKinsey Global Institute states that the proper use of Big Data within the United States healthcare sector might allow improvements with an estimated value of more than \$300 billion every year, two-thirds of which would be achieved by reducing the healthcare expenditure of the whole country [13].

However, healthcare IT history has made clear that technology-based panaceas do not exist. The potential of IT for transforming health systems seems to be widely accepted, as a consequence of its contribution to the improvement of healthcare processes, but IT has also caused new issues and risks, such as user-computer interaction problems or technology-induced errors. As a consequence, it seems clear that much IT outcomes-based research is still needed, in order not only to prove its value, but also to quantify it [10][14].

### *C. Requirements for the Use of Big Data within the Health Information System*

While the ability to manage massive amounts of data provides a huge opportunity to develop methods and applications for advanced analysis, the real value of Big Data will only be achieved if the information extracted from these data is useful to improve clinical decision-making processes and patient outcomes, as well as lower healthcare costs [9]. To that end, several basic requirements must be met, though they are very similar to the requirements of the health information system itself.

First of all, it is essential to ensure the quality of the information. This involves the development of thorough protocols which define the criteria required for data input, validation, harmonization if necessary, registry, processing and transmission to other components of the information system. In fact, several of the main requirements of data mining are the technical correctness of data, the accuracy and statistical performance, and the update or reassessment of the analysis [15].

In the health field, the information managed is so complex and heterogeneous that it is necessary to employ data carefully structured and, as long as it is possible, categorized. This is useful for data identification and error control purposes. Furthermore, healthcare information is a perfect example of three major features, commonly known as the three Vs, widely accepted as defining characteristics of Big Data: volume, variety, and velocity. In addition to these a fourth V, the veracity of healthcare data, is obviously critical for its meaningful use [8].



All possible information sources and data flows within the health information system must be perfectly identified as well. Since the information system must store all data required for the performance of the different corporate functions of the health system, it is clear that all of its components must be interoperable, as stated above, so that any data can be accessed from any point of the health system that needs them. Hence another cardinal requirement is the interoperability of systems, subsystems and components, defined as their capability of exchanging information without altering the meaning of the exchanged data, regardless of their source and their use within each system.

For instance, a medical consultation generates information used for the patient's healthcare, the management of the employed resources and the billing of the service, but it can also be used in the medium and long-term for outcome assessment, strategic planning, research, education, epidemiological surveillance, or even as evidence in legal proceedings. Moreover, the aggregate of every data generated during that consultation and the ones generated during millions of similar healthcare events will be useful to create knowledge, on which clinical decision-making support systems will be based.

Therefore the cycle comprehends the transition from data to information, from information to knowledge, and from knowledge to practice. All of this needs the interoperability of clinical information systems, logistic and economic-financial systems, business intelligence systems, and universities and R&D centres systems, among others. As a consequence, every system must be capable of filtering the information received in order to extract the data it needs, so as to not compromise their processing, thus avoiding the risk of producing adulterated results.

Finally, from a technological point of view, it is mandatory to have a high-performance IT infrastructure on which to rely for the generation, storing, processing and exchange of large data volumes, in a quick and efficient way. Luckily, hardware, software and communications solutions have experienced a huge progress in recent years, so technological viability is hardly an obstacle nowadays.

#### D. Some Additional Issues and Barriers

The implementation of Big Data solutions and tools in the health field requires addressing not only the organizational and technological issues detailed above, but also several legal and ethical questions.

From a legal point of view, the first cause of conflict may be data propriety. As explained in previous sections, every data properly processed and analysed can be turned into knowledge, and the latter can be easily made profitable. The first companies working this angle are tech giants such as Google, which provides personalized advertisements based on navigation and search history, and Facebook, which admitted to focus part of its efforts on sociological research based on its users' data, and has even tried to take possession of these information in a completely unilateral way [7].

Given that the generation, registry and processing of all this information requires a powerful hardware and software infrastructure as a base, and therefore a large investment by these companies, their intention to make it profitable may be considered legitimate to a certain point, especially if they are not charging users for the service provided. However, limitations regarding the use of the stored data must be clearly established, something that seems to be far from being solved with the current legal framework, which is quite confusing. For instance, in the case of Spain, this framework combines European Union, national, regional and sectorial (both health and e-government) regulations [7]. Moreover, most of this legislation is outdated to a large extent, since it was passed in a time when IT progress was far from the current one [16].

Once at this point, a revision of these regulations, taking into account the current potential of Big Data solutions, as well as the foreseeable one on the short and mid-term, seems to be more than appropriate.

Of course, this revision must be addressed with the goal of balancing the individual interests of patients (right to privacy) and professionals (legal certainty in the performance of their healthcare and management duties), as well as the general interests of society (research, education or improvement of healthcare services, among others). To that end, protocols must be defined that combine both a priori measures, such as data anonymization, and a posteriori measures, such as thorough audits regarding the access and use of data. Having the human factor in mind, one of the most crucial a priori measures will always be raising the awareness of patients, professionals and organizations.

From an ethical point of view, quite a few similarities to the legal field can be observed. The fact that IT is going to play an increasingly important role in health systems seems to be widely accepted, since its potential as a key instrument for the transformation of the current model is appreciated. Nevertheless, there is also a great concern about the lack of transparency in the management of the large amounts of data guarded by healthcare organizations. For this reason, the promotion of more and better control measures is backed by bioethics experts, starting with the development of a specific legal framework that can be turned into clear and visible actions, thus transmitting a sense of security and contributing to promote the trust in healthcare data mining [15].

### III. BIG DATA IN SERGAS: A CASE STUDY

Within the Spanish National Health System, healthcare is accountable to the Autonomous Communities, which represent the regional level of the government and each has a health service. In the case of the community of Galicia, this would be the Galician Health Service (Servizo Galego de Saúde, SERGAS).

SERGAS relies on a Business Intelligence solution for the exploitation of structured datasets, these being provided by a regional database in which information supplied by the different hospitals and primary care centers of this health service is aggregated. In addition, a management system for information related to human resources and pharmaceutical expenditure provides structured data as well.

In order to complement this BI system, SERGAS has implemented Big Data technologies so as to exploit unstructured data stored in the patients' electronic health records. This innovation makes SERGAS the first Spanish health service to use Big Data in a systematic way. On a total budget of 982.278 euros, several projects have been developed regarding the following lines of action:

- Rare diseases management:
  - Detection of suspicious cases.
  - Creation of a Rare Diseases Registry.
- Chronic diseases management:
  - Detection of Diabetes Mellitus type 2 patients, chronic obstructive pulmonary disease patients and patients with pluripathology, yet uncategorized as such in their health records.
  - Calculation of prevalence and incidence indicators, as well as risk factors.
- Clinical research: decision-making support regarding the selection of the most appropriate kind of vascular endoprostheses (stents).
- Nosocomial infections management:
  - Research and categorization of detected cases.
  - Automated alerts.
- Surveillance of several syndromes:
  - Case identification.
  - Detection of food toxi-infection and acute respiratory symptoms outbreaks.
- Exploitation of lab test results (currently in progress).

As a whole, these systems are handling information belonging to 2.900.000 patients, provided by 63 different data sources. Up to the year 2016, 59.000.000 normalized events have been compiled, 12.000.000 documents (of 50 different kinds) have been semantically processed, and 500.000 cases have been detected.

Regarding information security, SERGAS applies a set of corporate criteria, with standard measures such as the definition of user profiles and access authorization levels, the anonymization of aggregate data and the performance of audits to verify regulation compliance. Besides, there are several committees that define the guidelines for the management of ethics and governance, always within the current legal framework.

#### IV. CONCLUSIONS

As it happens with IT in general, the successful implementation of Big Data solutions in a healthcare environment will depend on their capability to generate an added value that benefits patients, professionals and organizations. No one seems to doubt the need to improve public health systems by evolving their current model, or the potentially valuable contributions of Big Data in this respect, but the great complexity that characterises the implementation of this kind of tools seems to be proven too, according to the requirements and, in some cases, obstacles of different nature that must be dealt with.

Once technological viability is apparently achieved, it is time for healthcare organizations and authorities to face the challenge of studying the possibilities of Big Data and seeking the best way of applying it to the solution of their issues, problems and needs. In order to achieve this, they must not start wondering what information they have now and what they can achieve with it, but what information they need and how they can get it. The most frequent problem will not be the availability of the necessary data, but the screening of the relevant information and how to assess it. In summary, the most important thing is not having the data, since this is already happening, but being able to ask the right questions at the right moment, process them to provide only the necessary and relevant information, and show the latter to healthcare professionals in a way that they can assimilate it in a quick, correct and easy manner, in order to make the right decisions at the right time.

On the healthcare side, Big Data must become the foundation of clinical decision-making support systems, and also an instrument for data aggregation concerning public health departments, as well as research and education. On the management side, managers will be able to have a more accurate and timely knowledge of the real status of their organizations, and adopt a prospective planning instead of a retrospective one. In addition, they will be capable of detecting deviations from objectives earlier and applying the appropriate corrective and, preferably, preventive measures.

In conclusion, the implementation of Big Data must be one of the main instruments for the change of the current health system model, turning it into another one with improved effectiveness and efficiency, calculated taking into account both healthcare and economic outcomes of health services, thus being meaningful to patients and also to society, and taking advantage of the patients' potential as active participants in their own care.

#### ACKNOWLEDGEMENTS

The authors wish to thank Ms. Pilar Carnicero and Mr. Guillermo Vázquez, for their contributions in elaborating and improving this article.

#### REFERENCES

[1] World Health Organization (WHO). *Everybody's business: Strengthening health systems to improve health outcomes*. WHO's framework for action,

2007.

[2] World Health Organization (WHO). *The Tallinn Charter: Health Systems for Health and Wealth*. WHO European Ministerial Conference on Health Systems. Tallinn (Estonia), June 25th-27th, 2008.

[3] Rojas D., Carnicero J. *A model of information system for healthcare: Global vision and integrated data flows*. In: Berhardt L.V., editor. *Advances in Medicine and Biology*, Vol. 82. Nova Publishers, NY, Enero 2015.

[4] Carnicero J., Rojas D. *La explotación de datos de salud: Retos, oportunidades y límites*. In: Carnicero J., Rojas D. (editors). *La explotación de datos de salud: Retos, oportunidades y límites*. Sociedad Española de Informática de la Salud (SEIS), June 2016.

[5] Hansen M.M., Miron-Shatz T., Lau A.Y.S., Paton C. *Big Data in Science and Healthcare: A Review of Recent Literature and Perspectives*. *Contribution of the IMIA Social Media Working Group*. Yearb Med Inform 2014; 9(1): 21–6.

[6] Porter M.E., Lee T.H. *The Strategy That Will Fix Health Care*. Harvard Business Review. Oct 2013. 10:50–70.

[7] Martínez R., Rojas D. *Gestión de la seguridad de la información en atención primaria y uso responsable de Internet y de las redes sociales*. In: Carnicero J., Fernández A., Rojas D. (editors). X SEIS Report, "Manual de Salud Electrónica para directivos de servicios y sistemas de salud – Volumen II: Aplicaciones de las TIC a la atención primaria de salud". Economic Commission for Latin American and the Caribbean (ECLAC) and Sociedad Española de Informática de la Salud (SEIS), August 2014.

[8] Belle A., Thiagarajan R., Soroushmehr S.M.R., Navidi F., Beard D.A., Najarian K. *Big Data Analytics in Healthcare*. Hindawi Publishing Corporation, 2015, 1–16.

[9] Panahiazar M., Taslimitehrani V., Jadhav A., Pathak J. *Empowering Personalized Medicine with Big Data and Semantic Web Technology: Promises, Challenges, and Use Cases*. Proc IEEE Int Conf Big Data. 2014 Oct; 2014:790-5.

[10] Kuziemsky C.E., Monkman H., Petersen C., Weber J., Borycki E.M., Adams S., Collins S. *Big Data in Healthcare - Defining the Digital Persona through User Contexts from the Micro to the Macro*. *Contribution of the IMIA Organizational and Social Issues WG*. Yearb Med Inform. 2014 Aug 15; 9:82-9.

[11] Luo J., Wu M., Gopukumar D., Zhao Y. *Big Data Application in Biomedical Research and Health Care: A Literature Review*. Biomed Inform Insights. 2016; 8: 1–10.

[12] Alper J. (rapporteur). *Big Data and Analytics for Infectious Disease Research, Operations, and Policy: Proceedings of a Workshop*. The National Academies Press, December 2016.

[13] Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A. *Big Data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, June 2011.

[14] Rojas D., Carnicero J. *La utilización de las TIC como un factor de riesgo para la seguridad de los pacientes*. En: Carnicero J., Rojas D., Martínez R. (editors). XI SEIS Report, "Las TIC y la seguridad de los pacientes: Primum non nocere". Sociedad Española de Informática de la Salud (SEIS), November 2016.

[15] León P. *Bioética y explotación de grandes conjuntos de datos*. In: Carnicero J., Rojas D. (editors). *La explotación de datos de salud: Retos, oportunidades y límites*. Sociedad Española de Informática de la Salud (SEIS), June 2016.

[16] Andérez A. *Disposiciones legales aplicables*. In: Carnicero J., Rojas D. (editors). *La explotación de datos de salud: Retos, oportunidades y límites*. Sociedad Española de Informática de la Salud (SEIS), June 2016.



David Rojas de la Escalera

Mr. Rojas has a degree in telecommunications engineering (with a specialization in telematics) from the Universidad de Cantabria. Currently, he is a Business Development Senior Consultant in the eHealth Division of the company SATEC. He has been an external advisor to the Spanish Ministry of Health, Social Policy and Equality for the eHealth Governance Initiative of the European Union. He is also a member of the Spanish Society of Health Informatics, and acts as a referee for the editorial board of the journal *Gestión y Evaluación de Costes Sanitarios* (Management and Evaluation of Healthcare Costs) published by Fundación Signo.



Javier Carnicero Giménez de Azcárate

Javier Carnicero Giménez de Azcárate, MD, PhD. Dr. Carnicero received a medical degree from the Universidad de Zaragoza and then went on to obtain a doctorate from the Universidad de Valladolid. Currently, he works in a management position for the Health Service of Navarre. He has been Director of the Spanish Healthcare System Observatory (Quality Agency - Spanish Ministry of Health, Social Policy and Equality). He is also on the executive board of the Spanish Society of Health Informatics.