

A Work Project, presented as part of the requirements for the Award of a Master  
Degree in Management from NOVA – School of Business and Economics

## **INTELLIGENT DOCTOR PATIENT MATCHING**

HOW JOSÉ DE MELLÓ SAÚDE EXPERIMENTS TOWARDS DATA-DRIVEN AND  
PATIENT-CENTRIC DECISION MAKING

**MAILIN HÜLSMANN**

**3657**

A Project carried out on the Master in Management Program, under the  
supervision of: Leid Zejnilovic PhD

**MAY, 23<sup>rd</sup> 2018**

## **Abstract**

While data-driven decision-making is generally accepted as a fundamental capability of a competitive firm, many firms are facing difficulties in developing this capability. This case demonstrates how a private healthcare organization, José de Mello Saúde, engages in collaboration with a global university-led program for such capability building, in a pilot project of intelligent doctor-patient matching. The case walks the reader through the entire data science pipeline, from project scoping to data curation, modelling, prototype testing, until implementation. It enables discussions on how to overcome managerial challenges and build the needed capabilities to successfully integrate advanced analytics into the organization's operations.

**Keywords:** José de Mello Saúde/ JMS, Healthcare, doctor-patient relationships, data-driven decision-making

## **Intelligent doctor-patient matching**

How José de Mello Saúde experiments towards data-driven and patient-centric  
decision making

*“Building and sustaining a relationship of trust is not something we invented two years ago, it has been and will be key in managing healthcare.”*

Rui Salinas, JMS

*The healthcare system of the future will look very different, with a crucial change being the move to ‘patient-centric’ healthcare, which will allow greater emphasis to be placed on prevention and access, using digital tools to improve productivity and boost efficiency.”*

World Economic Forum

Miguel Grade, who was recently promoted to lead the Marketing Intelligence team at José de Mello Saúde, was in an animated conversation with Rui Salinas, who had previously managed his team. They intensively discussed the results of the pilot project “Intelligent doctor-patient matching”, part of the company’s efforts towards a new data-driven and patient-centric approach to healthcare delivery. In February 2017, Rui Diniz, Vice President of the group, proposed this project for a collaboration with Nova School of Business and Economics (Nova SBE, n.d.) and Data Science for Social Good Europe (DSSG Europe, 2017). After spending over 20 weeks with a team of skilled and experienced data scientists at Nova School of Business and Economics to develop use cases and the analytics prototype, the organization is facing a challenge: what steps to take next to turn the experimental data analytics efforts into a scalable, valuable long-term project with commitment from all stakeholders of the organization, from the executive committee to medical professionals. They were curious to understand the broader impact of the project on the organization’s outcomes. However, they also knew the complexities of the healthcare environment, where historically change and innovation did not always happen easily. Traditional management structures had been embedded for more than 70 years and the path of developing something new, unknown and different than the norm, was long and required substantial investments to be implemented and aligned with the organization and culture.

## **José de Mello Saúde and healthcare in Portugal**

In 1998, the private healthcare organization José de Mello Saúde was established as a partnership of José de Mello SA<sup>1</sup> and ANF (National Association of Pharmacies), owning 70% and 30% respectively. Until the turn of the millennium, private healthcare accounted for a small part of the overall healthcare landscape in Portugal, since most healthcare providers were held publicly. When three other major players, Luz Saúde, Lusiadas and Trofa Saúde, joined the private healthcare market in Portugal in the early 2000s, José de Mello Saúde opened the Hospital CUF Descobertas in Lisbon in response to growing competition and with the goal to become the pioneer in private healthcare in Portugal. With an investment of around 35 million euros, the opening of this hospital was the largest private investment in healthcare in Portugal at that time. Within 15 years, the private healthcare market evolved from 25% to close to 40% of the total Portuguese healthcare market, playing an increasingly important role in the country [Exhibit 1]. Growing at a 9% CAGR, José de Mello Saúde evolved to be a significant part of the Portuguese national health services and established itself as one of the largest European healthcare networks and the largest private operator of healthcare in Portugal [Exhibit 2]. The organization held a 23% market share of the total private healthcare market, which accounted for approximately 1.6 billion euros in 2014 (José de Mello Saúde, 2018). As of 2017, the organization operated 16 privately owned units and two units under private public partnership (PPP) encompassing a workforce of 8,278 employees and serving 1.5 million patients a year. The private CUF network was composed of seven hospitals, eight outpatient clinics and one research institute [Exhibit 3]. In fiscal year 2017, JMS achieved revenues of 637 million euros (including private and PPP), an increase of more than 12% compared to 2016 with 586 million euros. Net profit was 23.3 million euros in 2017 (José de Mello Saúde, 2018) [Exhibit 4].

---

<sup>1</sup> José Mello SA is an important economic group operating in several strategic sectors in Portugal, including infrastructure and mobility, chemical, energy and health.

## **Innovation**

While José de Mello Saúde expanded at a fast pace, innovation was one of the fundamental corporate values [Exhibit 5] and key drivers for competitiveness and growth of the organization as the market leader and reference for private health care in Portugal. JMS established and maintained strong collaborative relationships with the academic and scientific community as indispensable means for generating and sharing knowledge to strengthen a culture of scientific excellence. In 2015, the “Tagus Academic Network for Knowledge” (Tagus TANK) was created between JMS and Universidade Nova de Lisboa to enable joint participation in projects, clinical research, training and exploration of new ideas and innovative solutions to constantly improve clinical care (José de Mello Saúde, 2016). To further accelerate the internal and external use of innovation, JMS joined the ‘Grow’ Program in February 2017, which was operated across the entire organization of the José de Mello Group (Grow José de Mello, 2017). The objective of this open innovation initiative was to strengthen the link between JMS and the start-up ecosystem and support and accelerate the development of innovative health care projects. The ‘Grow’ program allowed start-ups to have access to the knowledge, experience and unique infrastructures of the hospital group and its partner companies to develop joint pilot projects and to test and adapt products and services in a real market environment.

## **Clinical excellence**

Focus on safety and quality continuously guided the company’s way of providing health care. Marketing Manager, Rui Salinas, underlined the importance of quality of the business: “It is not about whether an apple in the supermarket is rotten or not, it is literally about life and death of a human”. In this context, Salinas and his team had an increasing focus on measuring results and patient outcomes. They used a system to measure risk adjusted mortality rate (RAMR), which was adjusted to predict the risk of death of patients considering their International

Classification of Diseases (ICD) profile. “We saved 18 lives that, without this quality measure, would not had survived.”, he emphasized. Moreover, all health units were equipped with most advanced equipment and excellent comfort conditions to guarantee the most demanding standards of safety and quality for patients. With a strong focus on clinical and information technology to continuously improve and optimize complex clinical procedures, JMS was on constant search for innovative tools and means to provide top-level clinical care [Exhibit 6]. For example, the CUF Robotic Surgery Unit used new robotic solutions like the Da Vinci Xi, which was at the top line of surgery robots. The Da Vinci Xi, with an investment of nearly two million euros, was a major step towards improved surgery efficiency enabled by advanced medical technology (José de Mello Saúde, 2016). Management at JMS felt very confident in making considerable investments into clinical systems. Salinas underlined: “When it is about new medical machinery, devices and diagnostic methods, we are all in for it, because we live on the cutting edge of science and therefore we are in need of the latest technologies” (Grade, M., Salinas, R., personal communication, April 10<sup>th</sup>, 2018).

### **Changes of expectations**

Rising expectations and increased digital capabilities posed new challenges and opportunities in the way healthcare was provided at JMS. With technological advancements, customers expected a level of convenience and personalization in healthcare similar to that offered by leading online retail companies or banks. Patients were better informed than ever, and they increasingly demanded access to information about their care. They wanted an active role in their health, manage appointments and access their personal health information when and where it suited them best. To keep up with these changing needs and take advantage of new digital capabilities, management at JMS saw the potential of streamlining processes such as setting up appointments with doctors. “By automating routine tasks, doctors can become more effective with their time and provide better diagnostic and treatment options for patients”, Salinas

reflected. In 2010, automatic check-in desks were implemented in all hospital units aiming to not only facilitate the entire reception process for the patient but also to allow the flow of information between services. To make the booking process as simple and convenient as possible, patients received a reminder notification one day before their appointment as well as a check-in code that they could easily enter at the check-in desk at the arrival area at the hospital. The new check-in procedure led to significantly decreased waiting times and higher efficiencies regarding the admission of patients.

## **MyCUF**

In response to changing expectations, Salinas and his team thought about new ways to leverage digital technologies to better serve patients and make relationships with patients more personal, accessible and convenient. Consequently, JMS introduced its new digital health service platform, MyCUF. This patient-focused portal provided patients with their own personal area, which was accessible through the websites of the CUF health units and via the MyCUF app. The app was launched in the first quarter of 2014 to respond to the increasingly mobile consumers. MyCUF allowed patients to access a set of functions and personal information about their activities and medical records over the past three years at CUF hospitals and clinics. Users also had the possibility to schedule appointments and exams, consult waiting times at the nearest health unit, browse and retrieve results analysis, imaging reports, consult the status of requests for authorizations of surgery, and make online payments [Exhibit 7]. Overall, the goal of this new digital service was to provide patients with a holistic view of all their interactions with CUF health units and empower them to manage their own health more effectively. The app had more than 310,000 users in 2017, an increase of approximately 48% compared to 2016. From January to December 2017, more than 385,000 appointments were registered on the MyCUF web version [Exhibit 8]. Moreover, MyCUF received the Portugal Digital awards for best digital engagement and digital transformation in 2016 (José de Mello Saúde, 2018).

## **Managing relationships**

The automation of administrative processes, such as patient registration and appointment bookings, lead to increased efficiencies in processes and workflows. Moreover, increasing amounts of patient data was captured and shared between services, which allowed for increasing improvements of health services. With these new digital efforts, it was now possible to simplify processes while at the same time strengthening the relationship with patients. The relationship with patients extended to any interaction the patient had with JMS, including making an appointment and finding out what came out from the latest lab tests. While all interactions and touchpoints were part of the patient journey at JMS, the most important stage was still the human interaction and time a patient spent with a doctor, which was a highly unique experience based on trust and confidentiality. It was not a new phenomenon that the doctor-patient relationship played a central role in improving the patient's health outcome and their perception of healthcare. Salinas underlined: "Building and sustaining a relationship of trust is not something we invented two years ago, it has been and will be key in managing healthcare." He added: "Patients with positive experiences with their clinicians build more trust in us as a healthcare provider, which strongly correlates with the perceived quality and continuity of care" (Grade, M., Salinas, R., personal communication, April 10<sup>th</sup>, 2018).

## **The paradox of choice**

Considering the importance of building trust with doctors, Salinas found that "patients should have the autonomy to choose their preferred doctor, the one they trust". He added: "Typically, patients want to see doctors that are most experienced". While the degree of experience was relevant in choosing a doctor, however, patients often faced uncertainty in knowing doctors and building trust because they had too many doctors to choose from. In 2017, one hospital in



Lisbon alone had more than 130 primary care doctors<sup>2</sup> [Exhibit 9]. When accessing the MyCUF app to schedule an appointment for consultation or examination, a patient was able to choose between the different hospital units and the specialty of the doctor that the patient was requesting to visit [Exhibit 10]. Then, based on this information, the patient was given a long list of all available doctors at the requested hospital unit [Exhibit 11]. Given this long list, the cognitive load of choosing among a few hundred doctors with little information about them was very high. Therefore, patients often relied on peer recommendation, from family and friends, therefore, engaging in limited search for alternatives. Even though the MyCUF app also offered a feature showing the most popular doctors of each hospital [Exhibit 12], these doctors were less likely to be available due to their popularity. Thus, patients' choice of doctor was also strongly influenced by the doctor's availability. Moreover, patients often sought a second opinion and therefore switched doctors, which led to additional consulting times and sometimes inefficient use of resources. Ultimately, this led to difficulties in developing long-lasting doctor-patient relationships and imbalances in assigning doctors. From the hospital management perspective, weak relationships could also result in patients' churn. Management at JMS understood that a higher probability of developing and maintaining trustful relationships between doctors and patients would benefit all stakeholders. However, for management, without any medical background, it was not an easy task to extract an understanding from doctor-patient relationships, which were based on very personal and confidential interactions.

## **Solution development: Intelligent- doctor patient matching**

### **Project scoping**

It was early February 2017, when Rui Diniz, Vice President of JMS, proposed a problem that could grow into an experimentation with data analytics. In several meetings, he talked about

---

<sup>2</sup> Primary care doctors Include General Practice and Internal Medicine doctors

the need to understand the relationship between doctors and patients, and the impact of the relationship on patients and their health. His general idea was to improve doctor-patient relationships with the use of existing data, though initially use cases were not clear. Diniz was convinced that machine learning offers opportunity for a more sophisticated, proactive approach to better match doctors and patients with the objective of building long-lasting relationships and attain better health outcomes simultaneously. With a small team in the Business Intelligence unit, internal analytics capabilities and resources were too limited to approach this problem of wide scope on their own. As such, they required knowledge from outside their organizational borders and people with strong skills in both data science and project scoping management. Benefiting from their close cooperation with science and research institutions, they welcomed a university-driven experimentation in data science, a program called Data Science for Social Good Europe (DSSG) organized by Nova School of Business and Economics. JMS got a place in the DSSG Europe 2017 summer fellowship edition, and a three-member development team of aspiring data scientists, was allocated to the project. In addition, the team had a support of one technical mentor, one project manager with expertise in project management, and one academic supervisor from Nova School of Business and Economics, who initiated the collaboration and led the project. As the project group decided to take on the challenge provided by JMS, they first needed to define the scope of the project. For this, they decided to use a project scoping methodology developed by the Data Science for Social Good team at the University of Chicago (Ghani, 2016). The methodology defines three basic steps: Firstly, defining goals, secondly, informing actions, thirdly, receiving and working with data, and lastly applying an analytical approach.

### *Step 1: Defining goals*

Defining a clear goal was not an easy task, mainly because key decision makers at JMS had not explicitly defined analytical goals for challenges the organization was facing. Also, for the

initiative to improve patient-doctor relationships with the use of existing data, goals were initially vaguely defined. Initial use case, proposed by Rui Diniz, was to focus on the hiring process of doctors to build an effective doctor infrastructure to increase the likelihood of long-lasting doctor-patient relationships. This challenge was primarily an HR management related challenge directed to operational processes and the assistance in hiring family doctors for a target population. This initial use case was analysed by the DSSG team, who then came up with an alternative problem statement that was actionable on the short-run: “Maximizing the likelihood of finding the right primary care doctor while minimizing the cognitive load for patients of choosing among many alternative doctors” (DSSG, 2017).

### Step 2: Informing action

Once the problem and goal were more clearly defined, the DSSG research fellows had to think about actions that had to be taken to make goals achievable. They identified two main applications; firstly, guiding patients in the moment of choosing a doctor, and secondly, allocating patients to doctors. Within these applications, they then developed use cases with different target populations based on the patients’ available past interactions with doctors. By offering two applications, the team aimed to demonstrate that their solution could function in two key areas of business analytics: firstly, customer experience, by offering personalized recommendations to patients [see Exhibit 13, application 1], and secondly, operational excellence, by assisting the decision makers at JMS to make better decisions when looking at the performance of doctors and understanding which doctors would correspond to the target population [see Exhibit 13, application 2]. The second application was designed to inform the hiring process of doctors who were the preferred choice of the target population.

### *Step 3: Receiving and working with data*

While the project was mainly problem-centric, the accessibility and relevance of data was a very critical component to achieve the stated goal and apply the identified use cases. Once the use cases were developed, the research fellows had to see firstly, what data they needed to develop their model, secondly, what data the organization made available and thirdly, if this data answered the questions they needed to address.

#### *JMS data management*

Over the years, vast amounts of clinical patient data were collected in various databases of different systems of the organization. The electronic health records (EHR) system stored and secured data such as patients' demographics, medical history, laboratory results, clinical findings, medication lists, and imaging results in an electronic format. Moreover, the Business Intelligence unit, responsible for data collection and management, captured structured diagnostic documentation using ICD-9 system, which was mainly used for billing and other transactional purposes. Other data was held in unstructured formats, such as clinical letters found as a text file. Systems and medical devices, such as MRI, came from different vendors and each had their own definition for data structures, which implied that each system stored and processed data in their own way. There was a total of eight different operational databases that were not connected with each other. Hence, the exploration of available datasets to achieve valuable insights was a daunting task for the Business Intelligence unit due to the variety of data, structured and unstructured, from various sources. The Business Intelligence Unit had started to design a data lake with the purpose to store and merge data from all databases. However, this was a long process and to this point data was still siloed to a large extent; each application had its own database and there was no possibility for analytics across applications and systems. Consequently, it was very challenging to meaningfully use and extract knowledge from the available data (Grade, M., Salinas, R., personal communication, April 10<sup>th</sup>, 2018).

### *Accessible data sources*

The databases provided by JMS to the DSSG team were developed and maintained by SAP, Deloitte and Glintt. They consisted of 34 tables holding approximately 500 million records, recorded between 2012 and 2017. The Business Intelligence unit provided anonymized socio-demographic information of patients and doctors and interactions between patients and doctors based on transactions. Additionally, this data included details about medical insurance, patients' expenditures and medical treatments received by patients. As required by Portuguese law, any information that could expose a patient by name and medical history had to be excluded, which implied EHR (medical records). The Quality Assurance unit provided ICD-9 codes of the conditions for which inpatients were admitted. These were not available for outpatients, which resulted in incompleteness of the information for outpatients. In summary, the DSSG data scientists obtained a large-scale data set with over 72 million consultation history recordings on the interaction and socio demographic data of 1.5 million patients and 3500 doctors from 16 CUF hospitals and clinics across Portugal (DSSG Europe, 2017).

### *Working with data*

The data science team extracted a subsample of approximately 280,000 interactions between 226,000 patients and 314 primary care doctors in 12 CUF hospitals (Han, Zejnilovic, Barros, 2017). Since the obtained data was not organized for applying machine learning models, they needed to manage, analyze and interpret the available data records. The application of analytics required the transformation of the available, unstructured datasets from JMS' data warehouse into standardized data and useable information that could be relayed back to the end-user, the hospital management. The technical mentor of the project said: "Once we received the data, we realized that there are a lot more challenges than we thought" (Han, Q., personal communication, March 22<sup>nd</sup>, 2018). Each of the hospitals maintained their own transactional customer database. For this project it was the first time that patient data was assembled from

multiple hospitals. As the data were not standardized across the different entities, data analysis was not possible instantly and the team had to explore more with it. For example, one patient had two different IDs from two different hospitals. As such, the team had to refer to another, separate database for which they needed to link other data sources to uniquely identify one patient. Moreover, they had to access another database provided by the Human Resources department with a list of all registered doctors, their roles, their location and utilization of medical supplies. Besides having to work with several different databases, the actual data the team received was not necessarily information on clinical doctor-patient interactions. Instead, the data was billing and transactional information for legal and financial purposes, which accounted as a receipt every time a patient received a treatment. Thus, the data source only included information on the reason of a patient's visit, the cost and the expected payment for the claimed services. To make meaningful use of this kind of data, the data scientists had to further clean, match and validate the data. One research fellow said: "We spent almost half of the summer trying to figure out how to use this information and extract an understanding of patient-doctor trust" (Martínez, Í, personal communication, March 22<sup>nd</sup>, 2018). After a lot of cleaning and matching, they were eventually able to transform the 34 tables from JMS' data warehouse into 6 tables in form of materialized views<sup>3</sup> that transformed the existing business structure into one which was suitable for developing and running machine learning models. When exploring with the data, the team found statistical evidence regarding the need for guidance in patient' doctor selection. Less than 40% of patients consistently visited the same primary care doctor, while more than 60% patients changed their primary care doctor at least once in the period between 2012 and 2017 (DSSG Europe, 2017).

---

<sup>3</sup> Materialized View is used to reduce the load on database servers when tables contain millions of records to create machine-learning understandable data set that conform to a clearly defined schema and make it computationally easier to develop recommendations (DSSG, 2017).

#### Step 4: Analytical approach

Once the data was transformed to represent the data in a format amenable to perform analytics, the data scientists needed to develop a framework to address the challenge and modularize the problem of choosing a doctor. When determining the modelling approach, they had to decide whether to formulate the problem as a combination of a classification problem, classifying the doctor-patient relationship as good or bad, and a grouping based on similarities, or as a recommender system, anticipating which doctor a patient is most likely to choose and ranking the doctors based on predictions. The latter proved to be most viable in the context of addressing the information overload problem and guiding patients in their choice of doctors by providing a ranked list of relevant doctors. Once the model was chosen, the team faced the challenge of how to leverage past patient-doctor interactions to anticipate future interactions. Using statistical models, they started exploring the relationship between doctors and patients by predicting each patient's outcome. They modelled the visitation history as implicit feedback from patients to indicate their trust in doctors, since explicit feedback, such as ratings or explicit opinions on experiences of visits, was not available. The team also proposed a quantitative trust parameter based on the number of interactions (episodes) while considering that trust could decay over time, meaning that patients value more recently visited doctors and switch from doctors they have negative past experiences with. This trust measure was incorporated as implicit feedback to represent the patient's preferences towards the doctors. To tackle the paradox of choice and information overload problem, the idea was not to present a long list of possible doctors but rather a subset with top  $k$  (3 to 5) doctor recommendations ranked by relevancy according to the trust measure. The recommender system was formulated as a learning to rank task to present a list of doctors to each patient to optimize the prediction precision. The result was a hybrid recommender system in form of open source, one that involved both collaborative-filtering and content-based prediction techniques. The system united the benefits of both approaches by incorporating the patient and doctor metadata

(content-based) into the traditional collaborative-filtering model, which explored the interaction history between doctors and patients. The hybrid setting allowed to offer patients with different levels of consultation history [Exhibit 14] a unified solution by consistently recommending relevant doctors ranked by the predicted trust scores. The cold start problem was therefore eliminated by complementing meta data from patients who had no prior interactions to find the doctors visited by patients with similar metadata (DSSG Europe, 2017) [Exhibit 15].

## **Results and evaluation**

To understand how good the system was at predicting how likely a doctor was a good match for a given patient in upcoming years, the data scientists performed temporal cross-validation. This method allowed them to evaluate how well the trained model performed on new data, to predict the near future. However, there was no reference point or benchmark to understand the value of the recommender system. Therefore, they needed to create a heuristic baseline model [Exhibit 16] that simulated a simple decision process when choosing a doctor. Once the baseline model was developed, the performance of the hybrid recommender system was compared against the baseline model using hit-rate@ $k$  measure. This measure reported the percentage of the population for which the patients have chosen one of the recommended  $k$  doctors. The results showed that the hybrid model outperformed the heuristic baseline model by approximately 20% (DSSG Europe, 2017). For nearly 70% of patients, the recommendation algorithm was able to recommend at least one relevant doctor when presenting patients with three recommendations as compared to 50% when recommending lastly visited or most popular doctors using the baseline model (Han, Zejnilovic, Barros, 2017) [Exhibit 17]. The results showed that the recommendation algorithm provided significant value to recommend hundreds of thousands more patients with their preferred doctors compared to the baseline model. Hence, the system proved to be an effective tool for guiding patient's choice of primary care doctors. The results also demonstrated that the system was scalable; the results shown for 226,000

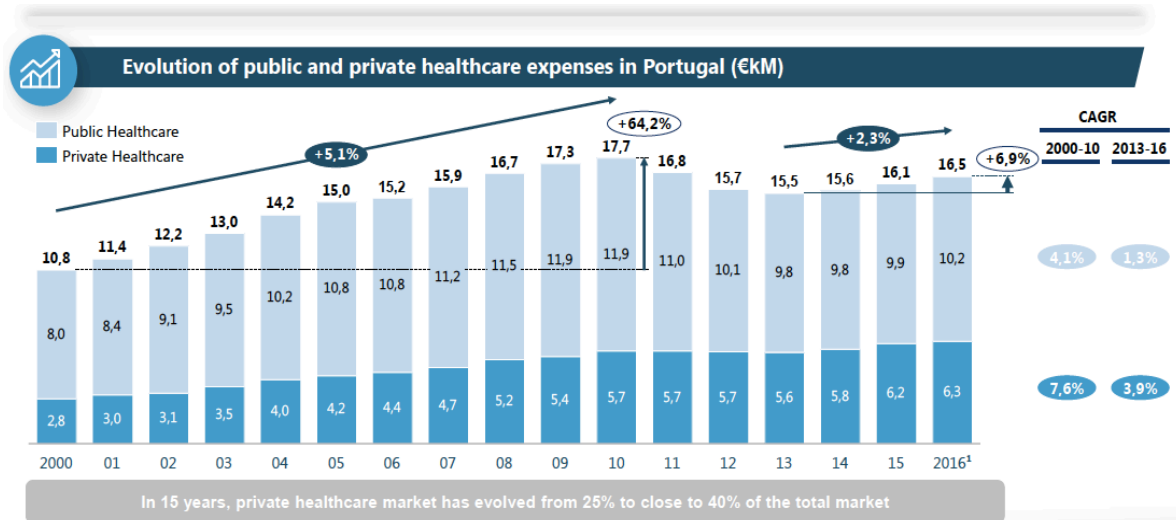


patients could be scaled up for the total of 1.5 million registered patients at the hospital group. “This level of scalability implies that the increase of amount of data collected could lead to even more precise and personalized recommendations of the system “, one research fellow concluded.

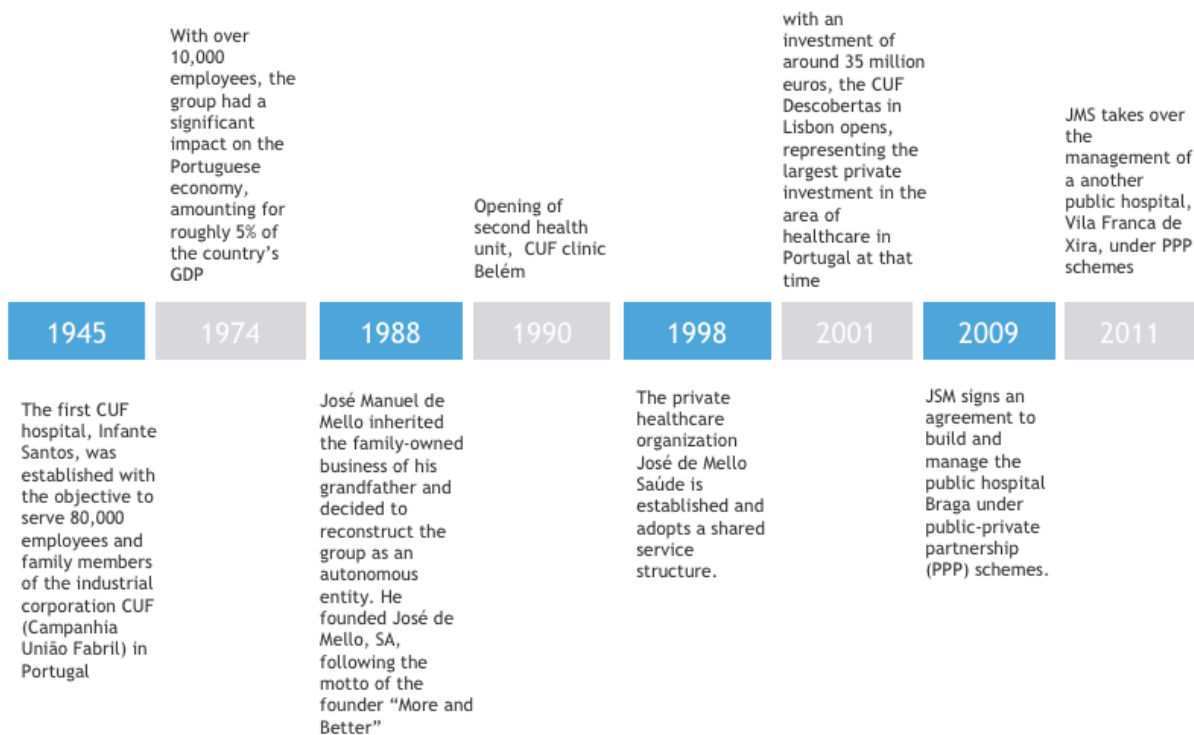
### **What’s next?**

As the results showed, intelligent-doctor-patient matching presented a novel use case for the potential of analytics to manage operations more effectively, improve customer experience and build long-term connections with patients. Besides these potentials, Rui Diniz understood that the value of advanced analytics would only be achieved once it was fully integrated into the organization as a long-term project. He suggested to Rui Salinas and Miguel Grade to explore how to bring the value of the project to JMS. Grade who took the new position of the Marketing Intelligence Manager, was keen to see this project implemented. With the changes in management structure, the question was how to proceed. Should they go into a testing and experimentation phase and if yes, over which channels (MyCUF, traditional mailing service, at the front-desk)? How could he find a way to convince people, both leadership and medical staff, about the value of intelligent doctor-patient matching? Would patients trust recommendations that are based on machine learning algorithms? How would the new General Data Protection Regulation (GDPR) from the European Union that is starting to be enforced on May 25, 2018 change the prospects of doing personalized doctor-patient matching? These were some of the essential questions facing Grade and Salinas. To understand the big picture, they needed to know what challenges they would encounter and what capabilities and resources they needed to develop when implementing the recommender system and maintaining the use of data-driven decision making on a large scale.

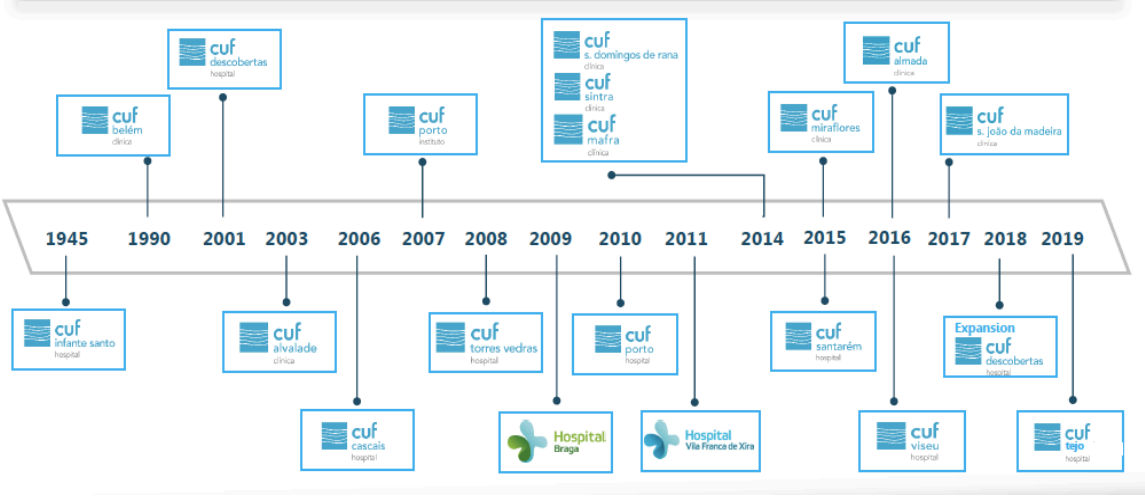
**Exhibit 1:** Healthcare expenditures in Portugal (2000-2016), in *José de Mello Saúde Investor presentation, February 2018*



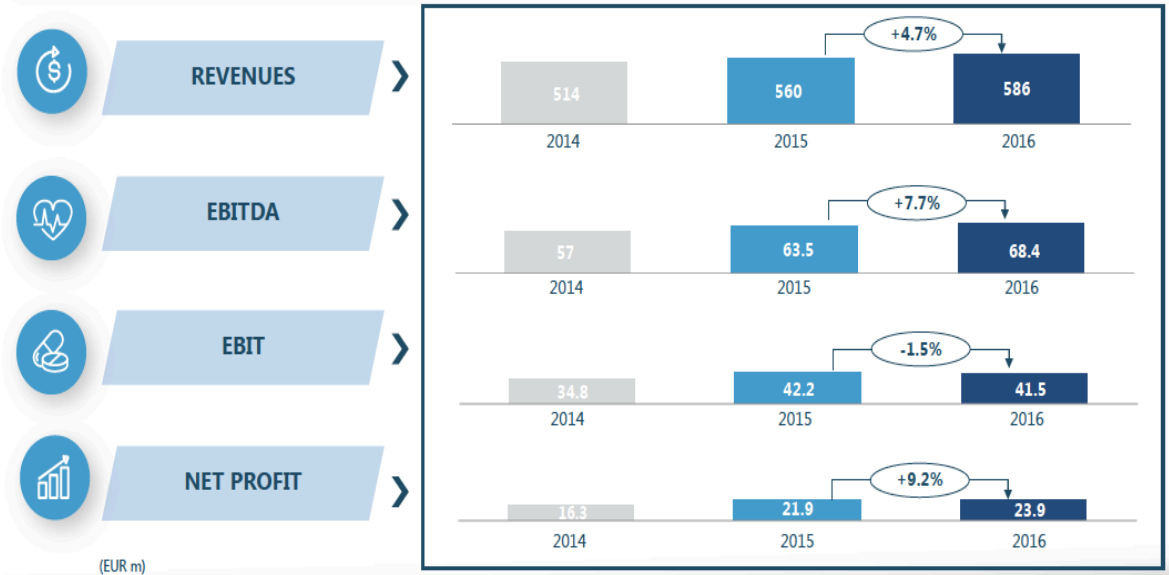
**Exhibit 2:** Timeline of important events in the history of José de Mello Saúde, in *Vida Magazine “José de Mello Saúde, 70 years of health in Portugal”, June 2015*



**Exhibit 3:** José de Mello Saúde Health Units (1945-2017) and expansion plans until 2019, in José de Mello Saúde Investor presentation, February 2018



**Exhibit 4:** José de Mello Saúde financial Results (2014-2016), in José de Mello Saúde Investor presentation, February 2018



José de Mello Saúde financial Results (2016 and 2017), in *José de Mello Saúde Annual Report 2017 (in Portuguese)*

## Resultados consolidados

### Demonstração dos Resultados

(Milhões de Euros)	2016	2017	Var.	Var. %
Proveitos Operacionais	586,3	637,4	51,2	8,7%
Custos operacionais*	(517,9)	(565,4)	(47,5)	-9,2%
<b>EBITDAR</b>	<b>79,4</b>	<b>83,8</b>	<b>4,4</b>	<b>5,6%</b>
Margem EBITDAR	13,5%	13,2%	-0,4%	
<b>EBITDA**</b>	<b>68,4</b>	<b>72,0</b>	<b>3,6</b>	<b>5,3%</b>
Margem EBITDA	11,7%	11,3%	-0,4%	
Amortizações e Provisões	(26,8)	(29,4)	(2,6)	-9,6%
<b>EBIT</b>	<b>41,5</b>	<b>42,6</b>	<b>1,1</b>	<b>2,5%</b>
Margem EBIT	7,1%	6,7%	-0,4%	
Resultados Financeiros	(8,9)	(10,5)	(1,6)	-18,0%
<b>EBT</b>	<b>32,6</b>	<b>32,1</b>	<b>-0,5</b>	<b>-1,7%</b>
Impostos	(8,4)	(8,8)	(0,4)	-4,3%
<b>Resultado Líquido</b>	<b>24,2</b>	<b>23,3</b>	<b>-0,9</b>	<b>-3,8%</b>
Resultado Líquido Atribuível aos Interesses que não Controlam	0,3	0,5	0,2	67,3%
<b>Resultado Líquido Atribuível aos Acionistas da JMS</b>	<b>23,9</b>	<b>22,8</b>	<b>-1,1</b>	<b>-4,6%</b>

\*Total menos Amortizações e Provisões

\*\* Resultados Operacionais mais Amortizações e Provisões

**Exhibit 5:** José de Mello corporate values, mission, vision and strategy (2014-2016), in *José de Mello Saúde internal presentation “Building the Future, Respecting the Past”, 2018*

THE CULTURE AND IDENTITY OF JOSÉ DE MELLO SAÚDE  
IS BASED ON STRONG VALUES



#### VALUES

- Respect for dignity and well-being of patients
- Human development
- Competency
- Innovation

#### MISSION

Promote the delivery of healthcare services with the highest levels of knowledge, respecting the primacy of the environment and human life, through the development of intellectual capital in constant search for the best.

#### VISION

To be Portugal's leader in the delivery of healthcare with outstanding quality, supported by an integrated network of high-performance units in both the private sector and the public sector, with growth options in selected international markets.



### EXCELLENCE IN SERVICE

- Development of centers with clinical excellence
- Customer relationship management
- Humanizing care
- Constant service level improvements

### EXCELLENCE IN HUMAN TALENT

- Transmission of group values
- Performance evaluation and rewards
- Attentively managing and challenging the professional career of everyone
- Fostering a culture of accountability, demand, accuracy with the sharing of knowledge and teamwork

### EXCELLENCE IN OPERATIONS AND SYSTEM

- Ongoing development of innovation and planning capacities
- Continuous process improvement
- Systematic increase in productivity
- Strong focus on clinical and information technologies
- Strict cost controls

JOSÉ DE MELLO SAÚDE HAS 5 STRATEGIC PILLARS



#### CLINICAL DIFFERENTIATOR

Consistent clinical performance of excellence across the entire CUF network, with differentiated offerings in flagship hospitals



#### OUTSTANDING CUSTOMER EXPERIENCE

High satisfaction rates across the network, maximizing the digital commitment as a vehicle for convenience and efficiency



#### EFFICIENT AND CONSISTENT OPERATIONS MANAGEMENT

Capture network synergies, achieving high EBIT levels in line with the most efficient players in the market



#### GROWTH AND VALUE GENERATION

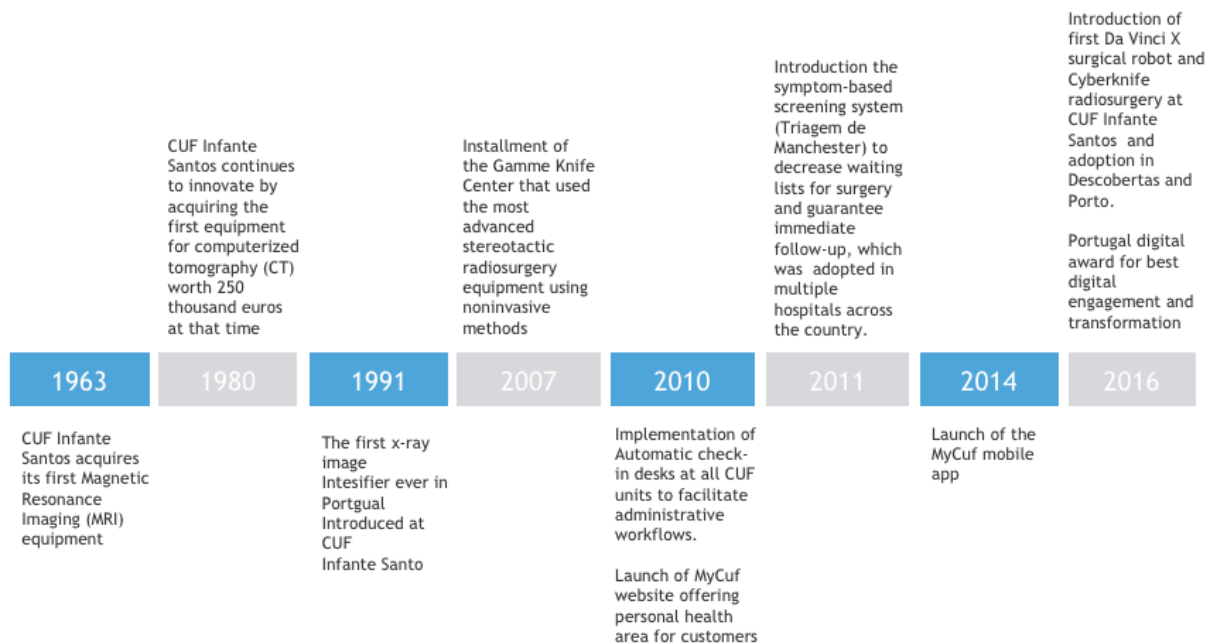
Consolidation of the leader position in the Portuguese market, capturing growth opportunities that generate maximum value



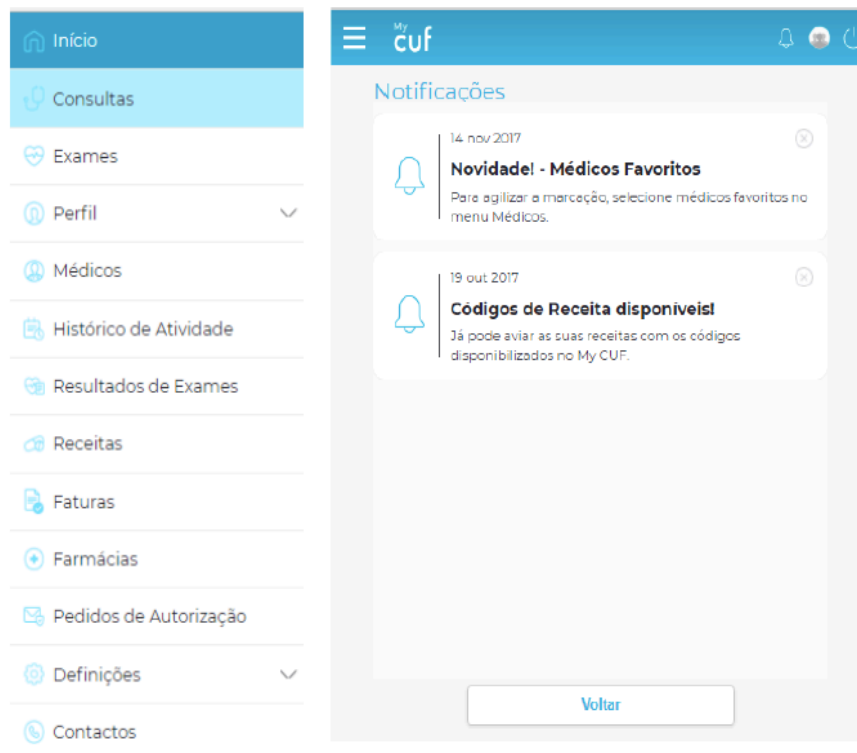
#### FOCUS ON HUMAN TALENT

Strengthening the value proposition of JMS by attracting, training and retaining professionals of excellence

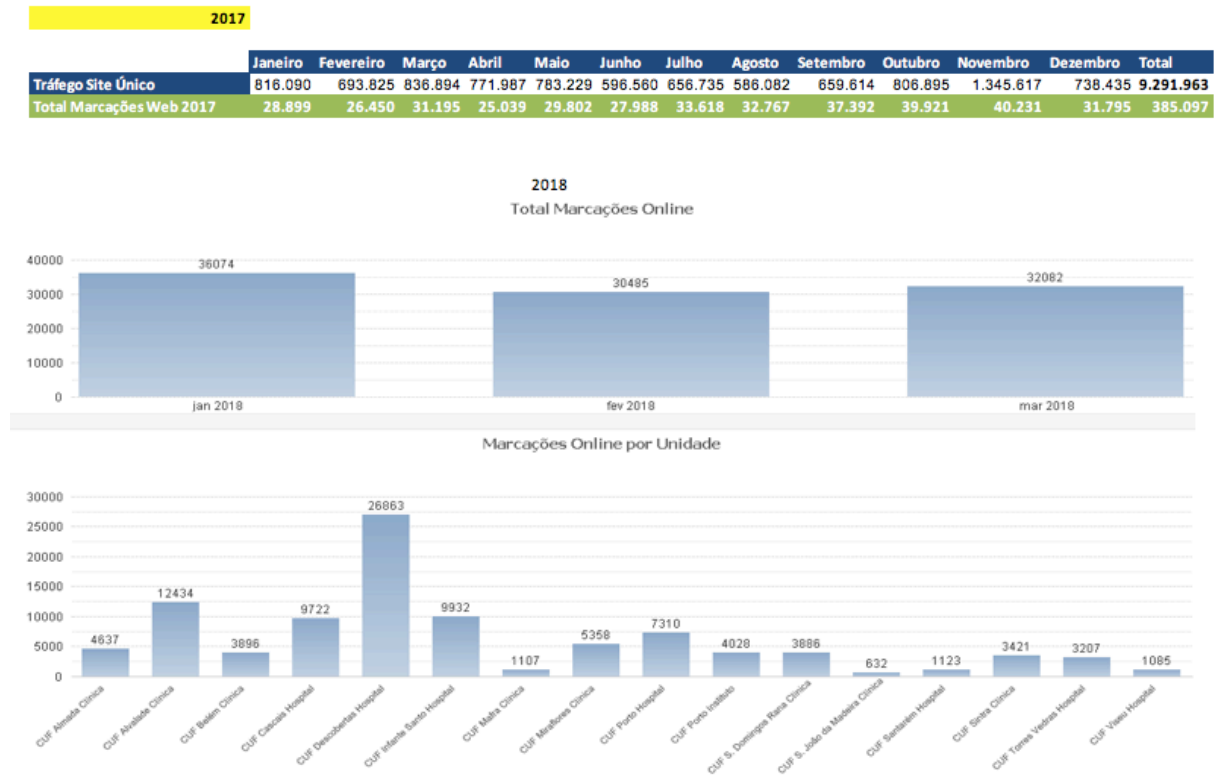
**Exhibit 6:** Timeline of important clinical and information technology adoptions at José de Mello Saúde, in *Vida Magazine “A Tecnologia ao Serviço da Medicina”*, August 2016



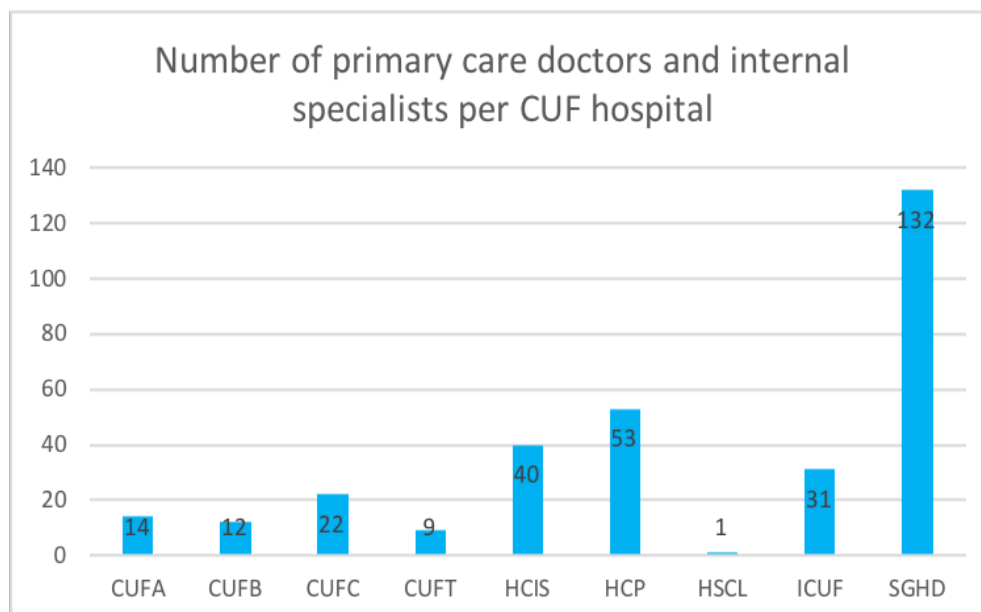
**Exhibit 7:** MyCUF app functions, in *José de Mello Saúde MyCUF*, 2018



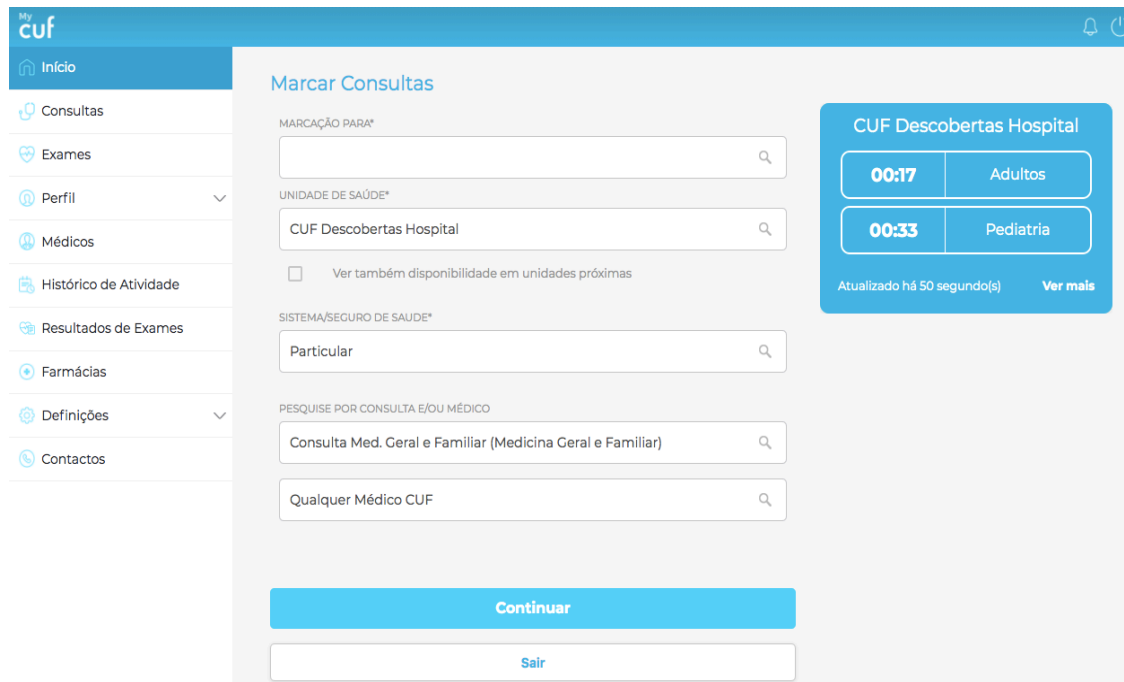
**Exhibit 8:** Number of appointments on MyCUF in 2017 and 2018 (per hospital unit), in *José de Mello Saúde internal database 2017*



**Exhibit 9:** Number of doctors at CUF hospitals, in *José de Mello Saúde internal database, 2017*



**Exhibit 10:** MyCUF app appointment booking function, in *José de Mello Saúde MyCUF*, 2018

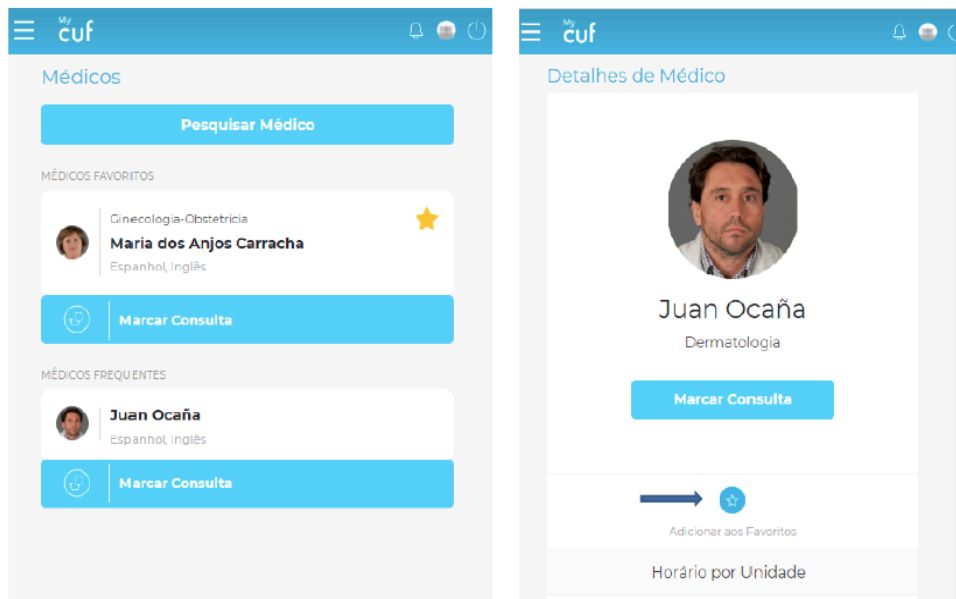


**Exhibit 11:** MyCUF example list of primary care doctors, in *José de Mello Saúde MyCUF*, 2018





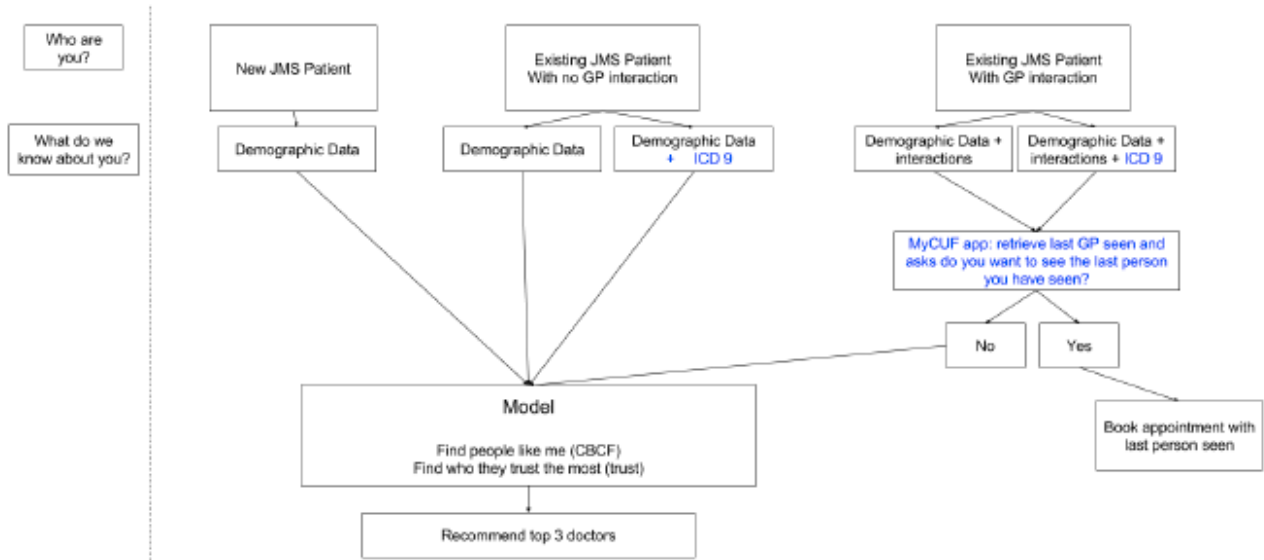
**Exhibit 12:** MyCUF example of “popular doctors” feature, in *José de Mello Saúde MyCUF*, 2018



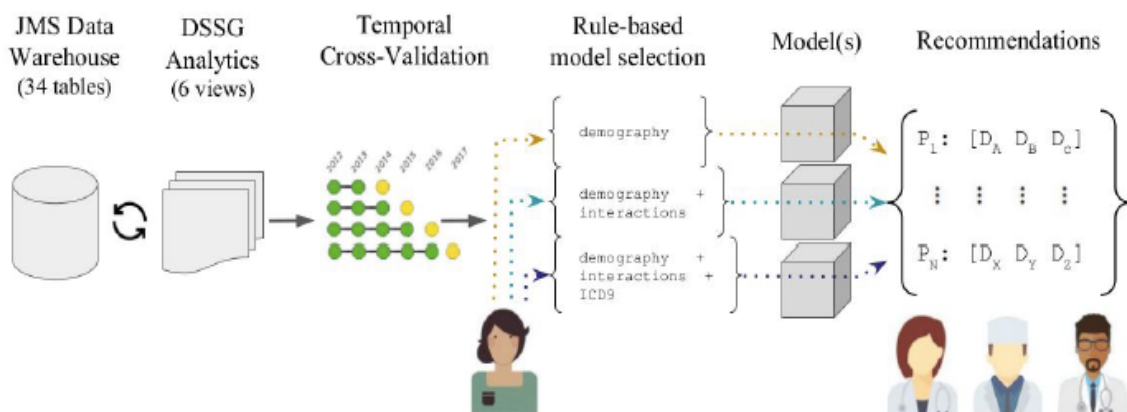
**Exhibit 13:** Doctor-patient matching applications and use cases, in *DSSG Intelligent doctor-patient matching project report*, October 2017

Application 1 – Recommending doctors to patients (target user – a patient)	Application 2 – Allocating patients to doctors (hiring doctors) (target user – hospital management)
<p><b>Use Case 1.1</b> For <b>existing patients</b> who have already seen a primary care doctor, but may want to switch, recommend top D most suitable doctors.</p> <p><b>Use Case 1.2</b> For <b>existing patients</b> who have not yet seen a primary care doctor, but have seen other specialists, recommend D doctors who have been most frequently visited by P most similar patients.</p> <p><b>Use Case 1.3</b> For <b>newly registered patients</b> who have not yet seen any doctor, recommend D primary care doctors who are the most popular overall.</p>	<p><b>Use Case 2.1</b> Allocate the types of doctors to a group of patients  Based on the characteristics of the most popular doctors, inform the <b>hiring process</b> for new doctors (to serve a general population, or for a specific segment of the market)</p> <p><b>Use Case 2.2</b> Allocate patients without a family doctor to available medical doctors.  <b>Marketing</b> department may want run a campaign and offer family doctors to patients who do not have one.</p>

**Exhibit 14:** Recommendations based on available patient data, in *DSSG Intelligent doctor-patient matching project report, October 2017*



**Exhibit 15:** Doctor-patient matching model development pipeline, in *DSSG Intelligent doctor-patient matching project report, October 2017*



**Exhibit 16:** Baseline heuristics for doctor recommendations, in *DSSG Intelligent doctor-patient matching project report, October 2017*

**Algorithm 2: Baseline heuristics for a single doctor recommendation**

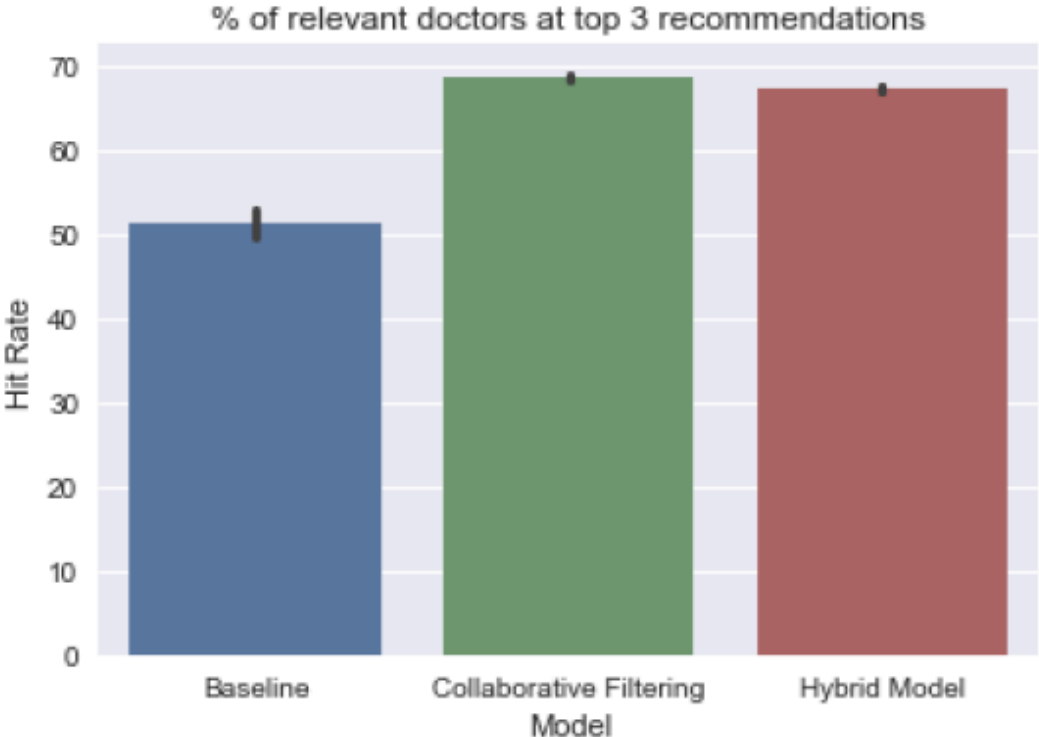
- a. Recommend the most frequently visited doctor.
- b. Given same visit frequency to multiple doctors, recommend the most recently visited doctor.
- c. If difference in frequency or recency not discernible from the data (given the defined time granularity, e.g. yearly visit count), randomly select one of the doctors who was visited most frequently and recently.
- d. If no previous doctor visits, recommend the K most popular doctors, (here we define popular doctors based on the number of patients visited in the same time period)

Top K=3 baseline model decisions for the patient, application of algorithm 2

2015	<ol style="list-style-type: none"> <li>1. Dr. C or Dr. D</li> <li>2. Dr. C or Dr. D</li> <li>3. Dr. B</li> </ol>	<p>With a year-wise time granularity, Dr. C and Dr. D appear to be visit at the same point in time (though the exact date and time may differ). Therefore, choose randomly between them, as rules 1 and 2 of the baseline model algorithm cannot be met. Dr. B is the next-most-recently visited doctor.</p>
2016	<ol style="list-style-type: none"> <li>1. Dr. D</li> <li>2. Dr. C</li> <li>3. Dr. B</li> </ol>	Dr. D is the most frequently visited.
2017	<ol style="list-style-type: none"> <li>1. Dr. D</li> <li>2. Dr. C</li> <li>3. Dr. B</li> </ol>	(Same as 2016)
<u>Year</u>	<u>Baseline model recommendation</u>	<u>Reason</u>
2012	<ol style="list-style-type: none"> <li>1. Dr. A</li> <li>2. Dr. X</li> <li>3. Dr. Y</li> </ol>	<p>Dr. A is the only doctor visited so far.            Dr. X is the most popular doctor overall.            Dr. Y is the 2nd most popular doctor overall.</p>
2013	<ol style="list-style-type: none"> <li>1. Dr. B</li> <li>2. Dr. A</li> <li>3. Dr. X</li> </ol>	<p>Dr. A and Dr. B are both visited once, but Dr. B was visited more recently.            Dr. X is the most popular doctor overall.</p>
2014	<ol style="list-style-type: none"> <li>1. Dr. B</li> <li>2. Dr. A</li> <li>3. Dr. X</li> </ol>	(Same as 2013)

**Exhibit 17:** Comparison of algorithms’ performance (for patients with previous interactions), in *DSSG Intelligent doctor-patient matching project report, October 2017*

Hit rate comparison for the baseline, collaborative filtering, and hybrid recommender model



## Case References

Data Science for Social Good (DSSG) Europe (2017). *Data Science for Social Good Summer Fellowship*. Retrieved from <https://dssg.uchicago.edu/europe>

Data Science for Social Good (DSSG) Europe (October 15, 2017). Project Report: *Intelligent Doctor-Patient Matching at José de Mello Saúde*

Ghani, R. (October 2016). *Scoping Data Science (for Social Good) Projects*. Retrieved from <https://dssg.uchicago.edu/2016/10/27/scoping-data-science-for-social-good-projects/>

Grade, M., Salinas, R. (April 10<sup>th</sup>, 2018). *Personal interview*, José de Mello Saúde, Carnaxide, Portugal

Grow José de Mello (2017). Retrieved from <http://grow.josedemello.pt/>

Han, Q.(March 22<sup>nd</sup>, 2018). *Personal Interview*, DSSG/ Nova School of Business and Economics

Han, Q., Zejnilovic, L., Barros, P., P. (2017). *Matching Patients with Doctors Towards a Trusting Relationship in Primary Care*

José de Mello Saúde (April, 2014). *Saúde à Distância de um 'Click'*. Retrieved from Vida magazine of José de Mello Saúde, Edition April 2014

José de Mello Saúde (August, 2016). *A Tecnologia ao Serviço da Medicina*. Retrieved from Vida meagazine of José de Mello Saúde, Edition August 2016

José de Mello Saúde (February, 2018). *José de Mello Investor Presentation*

Martínez, Í. (March 22<sup>nd</sup>, 2018). *Personal Interview*, DSSG/ Nova School of Business and Economics

Nova School of Business and Economics (n.d.). Retrieved from: <http://www.novasbe.unl.pt/>

## **Teaching Note: Intelligent Doctor-Patient Matching**

### **Case synopsis**

José de Mello Saúde, Portugal's largest private healthcare organization, places high priority on innovation and clinical excellence to provide high-quality health care to their 1.5 million patients in over 19 hospitals and clinics around Portugal. Understanding the importance of high-value medical technologies from the early days, the organization evolved to become a pioneer in providing top-level medical technologies that have led to drastically improved efficiencies of operations and optimization of clinical workflows. Moreover, JMS started to leverage digital technologies by introducing the patient platform, MyCuf, to improve patients' quality, accessibility and experience of care. However, with rapidly increasing amounts of patient-related data and rising expectations of patients regarding the value of health care, JMS needed to think of more novel, proactive ways to deliver patient experience and improve patient-doctor relationships. The company sought external knowledge and skills from a team of data scientists from Nova School of Business and Economics to develop the application of an intelligent doctor-patient matching recommender system, which resulted in an experimental prototype to be implemented into the JMS environment. Rui Diniz, the Vice President of the hospital group, delegated to both Rui Salinas and Miguel Grade, the new head of Marketing Intelligence, needs find a way to implement the system, and explore whether this data-driven decision-making approach creates value to all stakeholders, including doctors, patients and JMS management.

### **Learning objectives**

The case study intends to teach students to i) examine the rationale for a traditionally managed healthcare organization to adopt analytical approaches and the value of data analytics to improve healthcare service delivery, ii) understand the approach used to adopt analytics in the

JMS healthcare environment, and iii) consider the various managerial and cultural challenges of integrating and sustaining advanced data analytics in a healthcare organization. The case can be used in intermediate level courses at Master programs for a diverse set of domains including Strategy, Leadership, Change Management, Project Management as well as Digital Transformation. The case also has the potential to be taught in a course in Data Science, taking the position of an elective in the Nova SBE Master curriculum, which can fully shed light on scoping and managing data analytics projects. Moreover, the case can also be used as supportive discussion material in a course specializing in Healthcare Management for students to explore the complexities of introducing data-driven decision making in healthcare environments.

## **Case discussion and teaching plan**

### **Block 1: The value of data analytics in healthcare**

*Question: What is driving the decision to invest in data analytics and what is the value of intelligent doctor patient matching for JMS?*

With rising expectations regarding the quality and value of health services, health care providers like JMS evolve towards a patient-centric approach with greater emphasis on patients' value and access to care. While increasing amounts of patient-related health data are stored and managed at clinical databases such as electronic health records (EHR), it becomes critical to establish ways to more effectively make use and extract value from this data. In the context of an increasingly data-reliant health care system coupled with advances in analytics technology, data analytics are becoming increasingly prevalent with the potential to increase efficiencies and foster data-driven patient-focused decision making (Ward et al., 2014). Artificial intelligence and other forms of advanced analytics have the ability to transform large amounts of complex data into valuable knowledge to make more meaningful, evidence-based decisions and better utilization of time and resources, which ultimately allows for more and

lengthier patient-doctor interactions (Cisco, 2016). In this context, intelligent-doctor patient matching works as a complementary decision-making tool while mitigating the problem of doctor information overload and physician imbalance. As such, it is an added-value service that offers benefits to different stakeholders of the organization, including patients, doctors and the hospital management. For patients, finding an appropriate and trustworthy doctor to diagnose and treat medical conditions is one of the most important decisions they have to make because it will significantly impact the patients' health. Trust plays a central role in developing and sustaining relationships with doctors as higher trust results in better health outcomes and higher patient satisfaction. However, current research related to understanding the development of trust between patients and their family doctors heavily relies on survey-based measures, which do not consider rich information about actual doctor-patient interactions that may strongly signal personal experiences and trust in doctors (Crocker, 2013). The purpose of Intelligent doctor-patient matching is to make more meaningful use of the available data by leveraging past patient-doctor interactions to predict future interactions towards long and trusting relationships. Based on predictions, the system guides patients in their choice by offering recommendations that are best suited for their individual healthcare needs. From the physician perspective, doctors can manage their time with patients more efficiently, establish long-lasting relationships built upon trust and fully understand the context and needs of each individual patient to be able to act preventatively and promote healthy lives. For hospital management, intelligent doctor-patient matching shows the potential of analytics to optimize patient-doctor relationships, which will eventually improve patient satisfaction and overall health outcomes.

**Block 2: The solution development of a project like intelligent doctor-patient matching**

*Question: What challenges can arise when scoping a project like intelligent doctor-patient matching?*



With the support of the case, students should firstly identify the different steps of the scoping process as defined by DSSG, University of Chicago (Ghani, 2016) and secondly, find problems that can occur in each of the steps of project scoping by using the examples provided by the case. *Step 1: Goals:* Defining clear goals from the beginning is probably the most critical step in the scoping process. It is vital to have a clear outline of business analytics objectives and how they may impact overarching business goals before starting with the system development. This is important in order to put value, measure the process and outcomes and have a context in which to interpret results. As the case reveals, the executive committee did not define clear goals in terms of data analytics and the decision making to be influenced. This is a common situation in many organizations and it implies that there is a need for an iterative process of problem specification and definition of a goal that can be achieved with data analytics. The lecturer could ask students in this step to identify the stakeholders, people, or parts of the organization that can be affected by the intelligent doctor-patient matching initiative. For each identified stakeholder, students can brainstorm use cases. *Step 2: Actions:* What actions/interventions do you have that this project will inform? It is the company's role to determine and allocate the resources necessary to achieve stated goals. However, it was challenging for JMS to take actions that the organization was not familiar with implementing as current business processes were not taking analytics into account. At this step, the lecture could ask students to refer back to the use cases and derive which set of actions the project could inform. The lecturer can invite students to take a look into exhibit 13 of the Case to structure their discussion on the use cases and develop potential actions. *Step 3: Data:* What data is accessible internally? What data is needed? Is the data meaningful to continue with the analytical approach? Data collection is a critical step in the scoping process. At JMS, data was not easy to be arranged due to siloed databases and data collection from different systems. Each system has different rules for data collection and storage and not all critical information, such as EHR, was accessible, which lead to difficulties in attaining the needed data for developing

the system. *Step 4: Analytical approach:* In this project scoping step, the goal is to identify which types of analysis can be performed, given the goals, actions, and most importantly, the data available. For the JMS project, the dilemma was whether it was better to predict which doctor-patient dyad is likely to repeat or sustain over time, or whether a recommender system should be developed instead. These two approaches use similar algorithms but differ in use cases. Predictive modeling means predicting a value or a category, for example, if patient P1 will see doctor D1 in the future. While a recommendation system also incorporates predictive analytics, it creates additional data about preferences; when choosing between  $k$  recommended doctors, the system learns which doctor is preferred and may use this information for further personalization. The recommender system predicts which doctor a patient may or may not choose among the doctors who are ranked according to their ranking criteria.

### **Block 3: The technology and methods of a recommender system in a healthcare application**

This discussion block will examine the theory of recommender systems as a concept of predictive analytics, specifically looking at the design, evaluation and deployment of the doctor-patient recommender system. First, students should familiarize with the concept of recommender systems. As defined by Ricci *et al* (2011), recommender systems (RS) are information processing systems based on data mining and machine learning algorithms that actively collect and interpret various types of data in order to provide customized suggestions for items (recommendations), therefore serving as a complementary tool in the decision-making process of the end user. Recommender systems generate recommendations using knowledge and data about users (user model), preferences such as ratings and explicit feedback, the available items (features), interaction patterns and transactions stored in databases. A recommender system applied in the healthcare domain is driven by individual health data. The goal of a recommender system is to provide the user with high quality, evidence-based and personalized health content, which is highly relevant for the medical development of a patient.

It intends to decrease the effects of information overload originating from the rising amount of health-related data generated in clinical and operational databases (Ziefle, 2016).

*Question: What types of recommender systems exist, and which one is developed for intelligent doctor patient-matching?*

To generate recommendations, different recommender techniques can be used. Students should firstly be able to differentiate between the main three types of recommender systems: *collaborative filtering (CF)*, *content-based (CB)*, and *hybrid recommender systems (CFCB)* and secondly, explore the application of each recommendation approach in the context of doctor-patient matching by referring to the case [see Exhibit A for detailed explanation of each system]. Based on the information in exhibit A, students should understand why the hybrid setting is best suited for developing intelligent doctor-patient matching.

*Question: How can the performance of the recommender system be measured?*

Students should explore how the performance of a system like intelligent-doctor patient matching can be effectively measured and evaluated. Training and testing of a recommender system is performed on historical data, an “offline-analysis”. Data set is split into a training set that is used to make a model, and a testing set that holds data used to test the model trained on the training data (de Wit, 2008). In testing, k-fold cross validation is commonly used to predict the likelihood of correct recommendations. Folds may be either random samples from the data, and the letter k indicates how many folds are there. In time series, the folds are combinations of years of data (temporal folds) used for training and testing the algorithm. Results can then be analyzed using classification accuracy metrics<sup>4</sup> that measure to which extend a recommender system is able to correctly classify items as relevant or not (de Wit, 2008) [see Exhibit B for

---

<sup>4</sup> Accuracy metrics empirically measure to what extend a recommendation, as predicted by a recommender system, differs from the actual choice made by the user.

evaluation metrics]. For the evaluation of the doctor-patient matching recommender system, the metric hit rate was used because it is intuitively simple to measure how many recommendations are relevant (doctors a patient is likely to visit). Hit rate is a particularly useful measure in this case as patients tend to not visit many different doctors, about two on average.

*Question: How can you evaluate that the recommender system gives better recommendations than random recommendations?*

Students should think critically how it can be evaluated that recommendations mediated by the recommender system are better when compared with non-mediated (random) recommendations. Therefore, they need to understand that when setting up the system, it needs to be compared against something else as a reference point to understand the value of the system. The most effective random recommendation method to compare an AI based system against is the baseline model that applies common sense by using heuristics to create predictions for a dataset, which can be used to measure the baseline model's performance as compared to the machine learning algorithm of the recommender system. The desired goal is that the recommender systems outperforms the baseline model to provide evidence that the former performs better, i.e. it performs more accurate results on the test set e.g. gives more relevant doctor recommendations than the baseline method [see Exhibit C].

*Question: Can you think of other ways to evaluate recommender systems (in terms of user acceptance and satisfaction)?*

Evaluation of the recommender system in healthcare goes beyond standard evaluation criteria, i.e. accuracy metrics such as precision and recall [see Exhibit B]. It is crucial to benchmark the system, particularly in regard to user acceptance and satisfaction (Ziefle, 2016). In this context, user satisfaction, as it is difficult to measure, is defined as the extent to which a user is assisted in coping with the information overload problem (De Wit, 2015). User satisfaction is thought

to be a driving force behind other business goals. To enable benchmarking, more comprehensive quality metrics need to be sought since user satisfaction does not only correlate with recommendation accuracy. In this regard, students should be encouraged to think about how the recommender system can optimize value delivery to users without decreasing the value of other factors important to them. *Privacy*: A widely recognized issue for recommender systems are privacy concerns regarding sensitive healthcare data. Data on patients that is stored at the health care provider's warehouse will only be accessible to the system if it guarantees anonymity of patients. Some patients may prefer that they and not an algorithm decides which doctor they visit for a medical check-up, as they feel their privacy is under threat. *Trust*: A Recommender system does not offer value to a user if the user does not trust it. Trust can be established by transparently explaining how and why the system generates recommendations. *Communication of uncertainty*: Finding ways to visualize uncertainty of users in a set of recommendations is fundamental to allow the user to evaluate options properly. This problem is linked to the risk of the consequences of a choice in healthcare as one bad choice could result in bad outcomes. Therefore, the provider of the recommender system must act responsibly in both generating recommendations and communicating them to patients (Ziefle, 2016).

*Question: What is a meaningful way to go into implementation with the recommender system at JMS?*

Students are animated to think about the next steps in terms of deployment of the work that resulted in an experimental prototype. Students should discuss whether there should be any "online" evaluation phase, which implies a testing and experimental period where users interact with a running system and actually receive recommendations in a controlled test environment (randomly assign users to different conditions). In this context, the essential task is to observe the effects of the system by testing the prediction that the recommender algorithm provides more relevant doctor recommendations than the heuristics (random recommendations). Two

main objectives can be set for the testing: Firstly, testing the effectiveness of the recommendation algorithm on patients' choice and loyalty (are doctor-patient relationships stronger and of longer duration when mediated by the recommender system? Is user satisfaction higher?), and secondly, testing the effectiveness of communicating AI-based recommendations on patient's choice and loyalty (Would it make patients more/ less likely to visit a doctor when revealing/ concealing AI-based recommendations? What is the most effective communication strategy?) [see Exhibit D]. Next, students are encouraged to think about the steps for deploying the testing environment at JMS, including the system integration across different the communication channels, such as MyCUF, phone and front desks [see Exhibit E].

**Block 4: The challenges of introducing analytics and the element of change that needs to be managed at JMS**

This block encourages students to think critically about the complexities of introducing advanced analytics inside a large, traditional health care organization.

*Question: What are the challenges of implementing data analytics in healthcare?*

Healthcare organizations, like JMS, face some unique hurdles regarding the adoption of advanced analytics. Firstly, there are managerial issues as traditional organizational structures, legacy systems and conservative management philosophies are part of an industry which is historically slow to adapt to change and innovation. The foremost concern is how the key stakeholders in the healthcare environment would embrace change and how to convince senior leadership of the organization to set data-driven decision making as a business priority (Ward et al., 2014). Moreover, the healthcare environment finds itself in a complex business context, where risk perceptions towards data security and privacy and trust in safe technology play a critical role. The highly sensitive nature of medical data in line with the strict regulatory barriers, placing more pressure on health providers to comply with data protection regulations,

such as GDPR, lead to lengthy processes involving multiple stakeholders when collecting data (Bartley, 2011). The complex nature of health data, as it is often unstructured, incomplete, non-standardized and scattered around various locations, makes it more challenging to adopt analytics tools like recommender systems in healthcare than in other domains, where datasets are consistently available and reliable for analytics. Moreover, there are significant silos, both in terms of communication and needs between different stakeholders of the organization and in terms of information flow and data integration and interoperability due to siloed databases.

*Question: What changes need to be implemented at JMS to integrate data-driven decision making in the long-term?*

When forming an ongoing analytics roadmap, JMS needs to consider three critical components: i) technology, ii) processes and iii) people along with both short-term and long-term strategies to take advantage and bring analytics into a strategic and organizational context (Bartley, 2011).

*i) Technology and infrastructure*

It is key to make decisions on IT infrastructure changes needed to store and process data for analytical purposes. This requires a significant degree of robust infrastructure to handle the exponentially growing amounts of healthcare data as well as sophisticated analytical capabilities to extract meaningful use from data. To extract value from high amounts of data from multiple sources (clinical, demographic, financial etc.) with diversified formats (structured such as ICD-9 and EHR and unstructured such as clinical handwritten notes), JMS is in need of a centralized and holistic data infrastructure (data lake) that can facilitate the collection and aggregation of data from these various sources to enable easy access and interoperability.

*ii) Organizational and process alignment*

Analytics need to be aligned with organizational strategy and existing organizational processes as well as clinical processes. In a complex environment like a hospital, where crucial decisions

have to be made and actioned at fast speed without any margin for error, any changes that can impact processes have to be considered and planned carefully. This is important for data management and governance, where data collection and delivery have to be controlled with defined rules and processes, but also in terms of organizational alignment as lasting advanced analytics projects require firm alignment with business stakeholders. As significant investments have to be made into IT, data and analytical infrastructure, commitment to invest from top-level management is needed to demonstrate positive ROI. Projects, like intelligent doctor-patient matching, have to be framed within a define business need and should have a clear definition for success, ideally quantified by KPI's, to be well positioned to drive business value (Bartley, 2011). To enable effective organization-wide decision-making, all departments and stakeholders need to communicate and collaborate effectively, and silos need to be eliminated.

*i) People and culture*

When assessing the organization's analytical capabilities, the case reveals that appropriate investments in data science are not made in-house and thus JMS is reliant on experts outside the organization for data analytics to inform decisions. Thus, the implementation of advanced analytics also implies hiring and maintaining people with the education and skills needed when performing complex analytical and data-management tasks. In addition to data analysts and other technical staff, the system integration requires human resources with project management, team development and problem-solving skills. In order to achieve a high level of acceptance among all stakeholders, analytics needs to be understood as a continuous and evolving way of working that needs to adapt as business needs and priorities change. Intelligent doctor-patient matching, as a pilot project, can deliver clear value in a short time frame, which is needed to build support and commitment from senior leadership in order to add more analytics capabilities as technology evolves and data sources grow. Thus, change management and cultural adoption are critical when implementing business analytics in order to be embraced as an organizational and cultural objective and a key component of the organizations long-term strategy.

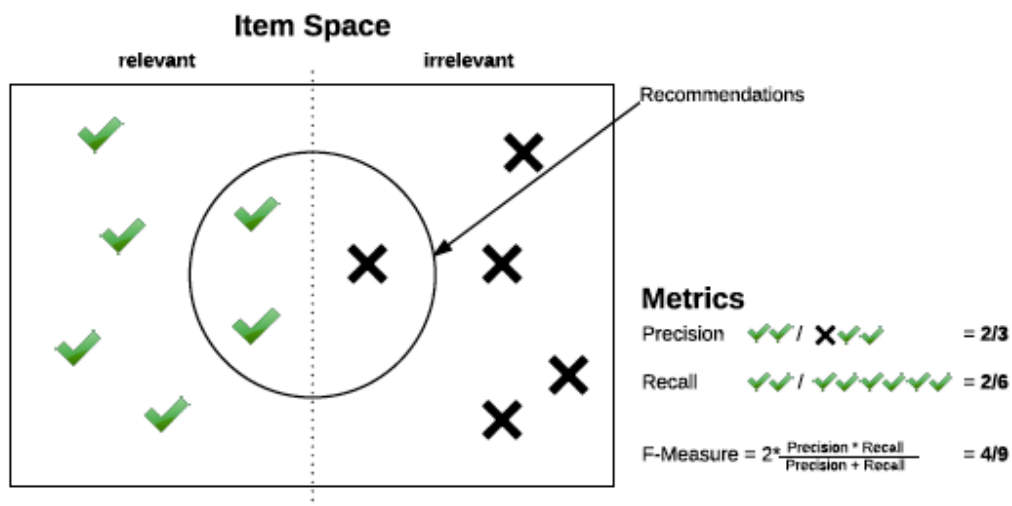


**Exhibit A:** Types of recommender system algorithms

Type	Recommendation algorithm	Intelligent doctor patient matching
<i>Collaborative filtering</i>	System gathers and analyses data about actions and behaviors of users and predicts what users like or do based on the similarity to other users. The similarity of two users is computed based on the rating or interaction history of the other similar users. Most widely used technique	System explores the interactions between patients and doctors (which patient visits which doctor how often) and can infer unknown relationships (i.e. recommend a doctor who a patient has not seen yet) based on similar coincidences across patients. However, if there is a new patient who has no previous record with any doctor, it is not possible to draw any similarities. This is called the “new-items” or “cold-start” problem.
<i>Content-based</i>	System links preferences to item attributes, i.e. it recommends items that are similar to the ones the user has liked in the past. The similarity of the items is computed based on features associated with the compared items. This system can be employed to generate personalized recommendations	System is used to recommend relevant doctors to new patients who have not yet seen any doctors in any CUF hospital. System utilizes the profiles of patients and their interactions with doctors to recommend doctors to new patients. It aims to find similarity between existing patients and new patients and between doctors selected by patients and other doctors.
<i>Hybrid</i>	system involves both collaborative and content-based prediction techniques. As it combines the techniques of both systems, it can use the advantages of one to eliminate the disadvantages of the other. For example, CF methods suffer from the cold start problem, as they cannot recommend items that have no prior ratings or interactions. This problem does not limit CB methods since the prediction for new items is based on features that are available	System unites the benefits of CF and CB by augmenting the patient attributes (gender, location, age group) and doctor attributes (gender, academic degree, working experience, specialization, age group) to the interactions between them from the CF model. When patient and doctor attributes are not presented, the model is simply reduced to a collaborative filtering model

**Exhibit B:** Visualization of three metrics used in the evaluation of a recommender system: precision, recall and F-measure, in *Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives*, November 2016

<b>Hit rate</b>	Measures whether a correct recommendation was made (value=1) or not (value=0). It does not provide a measure of how good the recommendation is (it considers hits equally regardless of their position in the list of top k)
<b>Precision</b>	Ratio of relevant items that are correctly recommended out of all recommended items (retrieved set). Precision value provides information on how well the system rejects irrelevant items from the recommended items (retrieved set). How many selected items are relevant?
<b>Recall</b>	Also called <i>Sensitivity</i> . The ratio of relevant items that are recommended out of all relevant items. Recall value of the recommender system measures the system's ability to find relevant items. How many relevant items are selected?
<b>F-measure</b>	The harmonic mean of both precision and recall and combines both values into a single metric



**Exhibit C:** Baseline heuristics for a single doctor recommendation, in *DSSG Intelligent doctor-patient matching project report, October 2017*

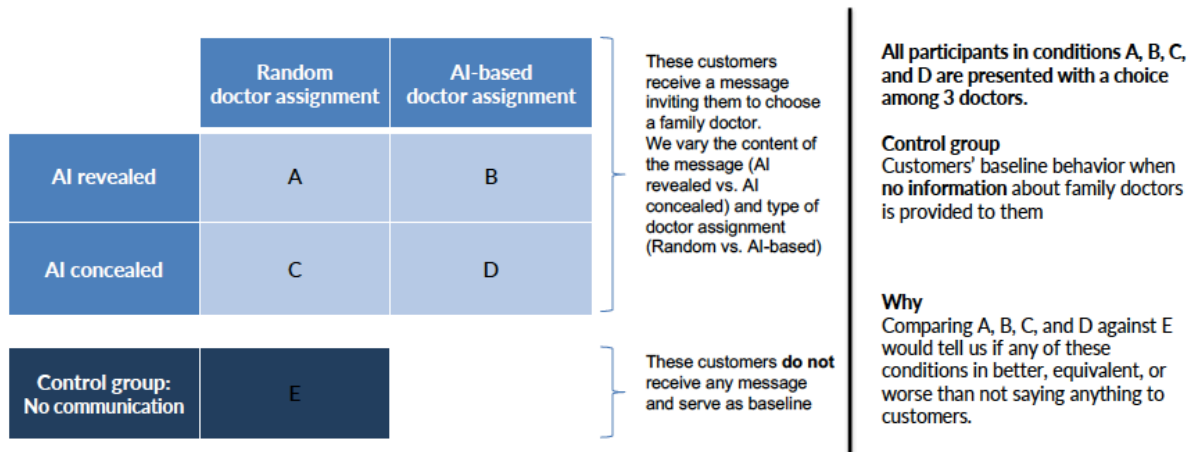
- a. Recommend the most frequently visited doctor.
- b. Given same visit frequency to multiple doctors, recommend the most recently visited doctor.
- c. If difference in frequency or recency not discernible from the data (given the defined time granularity, e.g. yearly visit count), randomly select one of the doctors who was visited most frequently and recently.
- d. If no previous doctor visits, recommend the K most popular doctors, (here we define popular doctors based on the number of patients visited in the same time period)

Top K=3 baseline model decisions for the patient, application of algorithm 2

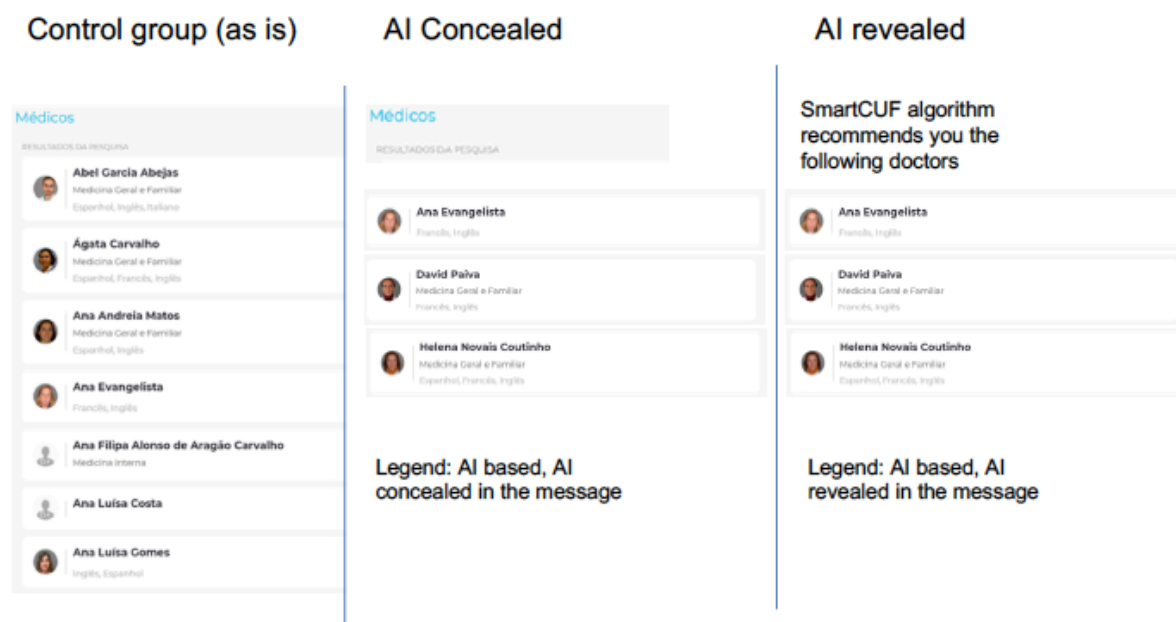
<u>Year</u>	<u>Baseline model recommendation</u>	<u>Reason</u>
2012	<ol style="list-style-type: none"> <li>1. Dr. A</li> <li>2. Dr. X</li> <li>3. Dr. Y</li> </ol>	<p>Dr. A is the only doctor visited so far.</p> <p>Dr. X is the most popular doctor overall.</p> <p>Dr. Y is the 2nd most popular doctor overall.</p>
2013	<ol style="list-style-type: none"> <li>1. Dr. B</li> <li>2. Dr. A</li> <li>3. Dr. X</li> </ol>	<p>Dr. A and Dr. B are both visited once, but Dr. B was visited more recently.</p> <p>Dr. X is the most popular doctor overall.</p>
2014	<ol style="list-style-type: none"> <li>1. Dr. B</li> <li>2. Dr. A</li> <li>3. Dr. X</li> </ol>	(Same as 2013)
2015	<ol style="list-style-type: none"> <li>1. Dr. C or Dr. D</li> <li>2. Dr. C or Dr. D</li> <li>3. Dr. B</li> </ol>	<p>With a year-wise time granularity, Dr. C and Dr. D appear to be visit at the same point in time (though the exact date and time may differ). Therefore, choose randomly between them, as rules 1 and 2 of the baseline model algorithm cannot be met.</p> <p>Dr. B is the next-most-recently visited doctor.</p>
2016	<ol style="list-style-type: none"> <li>1. Dr. D</li> <li>2. Dr. C</li> <li>3. Dr. B</li> </ol>	Dr. D is the most frequently visited.
2017	<ol style="list-style-type: none"> <li>1. Dr. D</li> <li>2. Dr. C</li> <li>3. Dr. B</li> </ol>	(Same as 2016)

**Exhibit D:** Experiment design: Doctor assignment, communicating AI-based offer and comparing against baseline behavior, in *DSSG internal presentation*

2 (Doctor assignment: Random vs. AI-based) x 2 (AI: AI revealed vs. AI concealed) + control  
 Participants are randomly selected and randomly assigned to ONE condition (A, B, C, D, or E)



Communication strategy – MyCuf web and mobile app



**Exhibit E:** Steps for developing testing environment, in *DSSG internal presentation*

- 1. Add the required tables and views to the database – for the system to function, but also for to keep the record of the offers and choices made.
- 2. Confirm that the key variables (anonymized to us) are used correctly to allow a meaningful function and presentation.
- 3. Develop the access and interaction with the system (seeing recommendations and choosing a doctor) to users in all channel (see Figure 17).
- 4. Validate that the system makes recommendations and records offers and choices.
- 5. Integrate mechanisms to collect explicit feedback about the choices across channels.
- 6. Establish the environment for re-training and fine-tuning the models.



Phone support



Hospital  
check-in desk



MyCUF app



Website

## Teaching Note References

- Bartley, A. (2011). *Predictive analytics in healthcare*. Intel White Paper - Solution Brief, 1–7. <https://doi.org/10.1017/CBO9781107415324.004>
- Cisco. (2016). *The Digitization of the Healthcare Industry: Using Technology to Transform Care*. Retrieved from [http://www.cisco.com/c/dam/en\\_us/solutions/industries/docs/digitization-healthcare.pdf](http://www.cisco.com/c/dam/en_us/solutions/industries/docs/digitization-healthcare.pdf)
- Crocker JE, Swancutt DR, Roberts MJ, Abel GA, Roland M, Campbell JL. (2013). *Factors affecting patient's trust and confidence in GPs: evidence from the English national GP patient survey*. *BMJ Open* 2013;3e002762
- De Wit, J. (2015). *Evaluating recommender systems*. *Recommender Systems Handbook, Second Edition*, 265–308. [https://doi.org/10.1007/978-1-4899-7637-6\\_8](https://doi.org/10.1007/978-1-4899-7637-6_8)
- Data Science for Social Good (DSSG) (n.d.). *Scoping Data Science (for Social Good) Projects*. Retrieved from <https://dssg.uchicago.edu/2016/10/27/scoping-data-science-for-social-good-projects/>
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y. et al. (2017). *Artificial intelligence in healthcare: past, present and future*. *Bmj*, svn-2017-000101. <https://doi.org/10.1136/svn-2017-000101>
- Kantor, P.B.; Ricci, F.; Rokach, L.; Shapira, B. (2011). *Recommender Systems Handbook*. Published by Springer Business+Science Media, New York, USA
- Ward, M. J., Marsolo, K. A., Froehle, C. M. (2014). *Applications of business analytics in healthcare*. Kelley School of Business, Indiana University. Published by Elsevier Inc.
- Zejniliovic, L. (n.d.). *The challenges and benefits of automated mediation system : case study of healthcare management*.
- Ziefle, M., Holzinger, A., Verbert, K., Felfering, A. (2016). *Machine Learning for Health Informatics, : Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives*. 9605(November), 0–24. <https://doi.org/10.1007/978-3-319-50478-0>