TRANSITION TECHNOLOGIES S.A.

MedStream Designer: a new software solution for patient search and analyses that supports Evidence-based Medicine and Personalized Medicine

by

Konrad Wojdan, Magdalena Ogłuszka, Rafał Ostap

Technical Report

7.12.2016

Abstract

BACKGROUND: Hospitals produce vast amounts of data that must be collected, stored and processed. Many information systems are used today in healthcare but they are not flexible and intuitive enough to meet the demands of hospital environment and modern medical standards.

AIM: The aim was to create a tailored tool for medical practitioners and hospital managers that not only enables them to exploit the potential of data stored in electronic health records (EHR) systems, analyze it and use it as a tool in the decision-making process but also supports up-to-date standards and advancements such as Evidence-based (EBM) and Personalized Medicine..

METHOD: A novel software solution has been developed. MedStream Designer (MSD) is dedicated to doctors, scientists, managers of hospitals and pharmaceutical companies conducting clinical trials. The solution allows to easily search and analyze a cohort of patients taking into account all the medical data, including their time relations, associated with the treatment. Managers can use MSD to detect adverse events and monitor quality of treatment.

RESULTS: There are several implementations of MedStream Designer (MSD) solution across Poland. The biggest one took place in Maria Curie-Sklodowska Institute of Oncology in Warsaw. The whole process of MSD implementation took about 6 months. The database of Oncology Institute covered almost 20-years history treatment of almost a million of unique patients.

Keywords: MedStream Designer, MSD, Health Information System, software, Evidence-based Medicine, Personalized Medicine

Introduction

Technology has a key role in accelerating the advancement of health care. Medical industry generates vast amounts of data that should be stored and analyzed for patient care, health care planning and clinical research.

The decision-making process in hospitals should be efficient, accurate and based on evidence to achieve the highest-quality patient care. It is crucial that physicians have the tools that help them in this process, provide with the evidence that they need but remain flexible as there are many important factors that must be considered. Choosing the right diagnosis and treatment plan is complicated. Health information systems help reduce those errors, improve efficiency and

minimize costs - this is why hospitals all over the world decide to implement them in everyday practice [1]. To provide useful information for the clinical research, the tools should be also up-to-date with the recent standards and developments such as Evidence-based or Personalized Medicine.

MedStream Designer is a unique application dedicated to the employees of the medical sector, especially for doctors and hospital managers. It is designed to support the analysis of large medical data sets. The goal was to create a tool whereby each user would be able to exploit the full potential of data collected and processed in electronic health records (EHR) systems for years. Until now, to analyze such complex data, medical centers had to buy the whole BI (Buisness Intelligence) system or a powerful tool for statistical analysis. Working with them requires expert knowledge in the field of data analysis.

We have worked together with doctors and HIS systems developers to remove all barriers in the large-scale data analysis. Thanks to MedStream Designer every physician can analyze medical records of hundreds of patients that have been treated in the clinic. The medical data linked with the patient may be freely searched and compared. Recently, an upgrade of the application has been introduced which enables to search the databases by the results of sequencing methods. The physicians are no longer limited and can explore the potential of patient's genetic information. This new feature will allow to treat the patient on more individual level and it is in line with the recent advancement in Personalized Medicine. Thus, we can conclude that the MSD supports the extraction of knowledge in a way that was not previously possible.

Existing IT systems in healthcare

Health Information System can be defined as any system that has been developed to generate, store, manage and process individual health information of patients or data used by managers and administration staff to control the activities of the hospitals or other healthcare companies. Its purpose can differ, however, usually the goal is to provide data for statistics and evidence to improve the decision-making process or research (some Health Information Systems have been designed to conduct analyses as well) and measure positive and negative outcomes of the interventions in terms of medical practice and economics.

In the 1970s, medical data was kept on paper. This approach has been partially changed in 1980s when electronic medical records were developed and used in the administrative departments of

the hospitals [2]. First health information systems were designed and developed to improve the work of special departments; e.g. Radiology Information System (RIS) [3]. RIS is dedicated for radiology departments and it has been used along with the Picture Archiving and Communication System (PACS). A significant improvement (especially in the number of steps within standard workflow) has been observed when compared with traditional approach that was based on film and paper [4].

This first approach to the health information systems has been replaced by a more general one, considering the information processing in hospitals as a whole [3]. Hospital Information System (HIS) is an example of a widely adopted integrated information system that is used for management, financial and clinical operations of the hospital [5].

Enterprise Resource Planning (ERP), Business Intelligence (BI) and many other systems used already in different industries turned out to be useful in health care management and clinical activities of the hospital [6][7]. ERP is designed to manage all the business operations and integrate the functional data in the organization but the implementation of this system is expensive and the problem of hospital's diversity must be considered. To make the ERP system effective in the hospital environment, a large group of professionals with different background must cooperate - medical practitioners should be properly trained and new procedures should be introduced [7].

It is difficult to track the history of a specific patient or population because large amount of data generated by the medical industry comes from variety of sources. To resolve this issue, BI (Business Intelligence) systems can be used [6]. Their aim is to pull data from different sources, integrate and analyze it so it can help in the decision-making process. However, these systems are not ideal - as mentioned before, they generate lots of data so the user can have problems with analysis and interpretation. Healthcare professionals need to have a proper training to use it. In addition, equipment requirements are very high which implies supplementary costs. Implementation is also long and expensive.

To conclude, many information systems have been used in healthcare but they are not flawless. Apart from benefits, there are also challenges that health workers meet while using them. Health Information systems are fragmented because they are used to achieve certain results. Data are generated and used by different public and private organizations. According to WHO [8], employees of the medical sector are overwhelmed by excessive data that needs to be processed in various and often complex ways. Data are often added to the Health Information Systems

without any critical analysis and therefore making decision based on such information can lead to mistakes that are serious in consequences. Some medical practitioners still think that information technology is depersonalizing their relationship with patients so the health information systems must be properly assimilated [7].The systems should be intuitive and possess strong usability. In this case, offering technological advances is not enough because if the healthcare professionals will not be able to easily interpret the information provided by the system, it will be useless. Additionally, IT systems must be used as a general practice because lack of information in the database can cause misleading conclusions. MedStream Designer has been created with those needs in mind. It has been developed with physicians to support them daily. MSD allows to search both numerical and descriptive data, select needed information, compare and analyze it. It also assists in identification of adverse events such as hospital-acquired infections and post-operative complications that could be overlooked otherwise. The application contains a set of predefined searching scenarios in line with the guidelines of the Center of Quality Monitoring. It is easily accessible, flexible and based on the principles of Evidence-based Medicine.

Evidence-based Medicine

In 1840 Ignaz Semmelweis correlated the mortality of women on the maternity ward with the frequency of hand washing after the examination in the dissecting room performed by the physicians. This conclusion led to implementation of the disinfection procedure and thus, reduction of death rate observed in patients on the maternity ward [9]. Ignaz Semmelweis is thought to be a protoplast of Evidence-based Medicine (EBM) - an approach with an aim of advancing the process of decision making in medical practice. It indicates the importance of using scientific evidence in clinical procedures to achieve the best efficacy and safety.

The principal of evidence-based medicine is asking properly-built clinical questions about patients, interventions, prognostic factors, exposure, prevention, comparative analysis and an outcome that should be measured or achieved [10]. Only the most accurate and relevant data should be included in the decision-making process and it is necessary to consider all positive and negative aspects of the intervention. All medical procedures must be properly planned; healthcare professionals should constantly search for the best evidence and the strength of this evidence should be evaluated.

EBM facilitates the process of choosing accurate treatment plan but the evidence can be used in inappropriate way if the tools that support the physician in the decision-making process are not

flexible enough and therefore misleading. The amount of data generated by hospitals is overwhelming, often incomplete or hard to process, which is a barrier that prevents researchers from using it for the development of new diagnostic procedures or improvement of existing ones. Patient's history is a sequence of medical events that are related to each other. The process of search and analysis should consider those cause and effect relations. Additionally, medical notes and descriptive results do not have a structured form. To fully analyze patient's medical record, this text data should also be included in the process. In MedStream Designer application, the user can ask clinical questions based on all important events from patient's medical history, determine whether the search criteria are time-dependent or time-independent, indicate if they must, can or cannot appear in the results list and evaluate search flexibility in the descriptive results or medical notes. In that way, the application adapts to the physician's needs and not the other way around.

MedStream Designer is based on EBM. It is used to provide the evidence for the physician and it shows consistent cohorts of patients based on the search criteria. The physician can examine medical history, all the procedures that have been applied and their outcomes. Thanks to the new features implemented in MSD, the user is also able to see if there are any prognostic factors for the progression of a disease. Then the application enables to compare those groups and analyze them to provide adjusted, easy to interpret and context-sensitive information that is used in the diagnostic process.

Personalized Medicine

The concept of Personalized Medicine has expanded over the recent years. The term is used to describe the process of integrating knowledge to predict the prognosis of the disease, its susceptibility and effects that can be observed during the treatment process. It has become possible thanks to the Human Genome Project which resulted in sequencing the whole human genome [11].

Genome-Wide Association Study (GWAS) is performed and used to determine whether a mutation is connected to a specific disease [12][13]. To achieve that information, sequences of genomes of many patients with clinical manifestation(s) and people without it are compared to verify which mutations can be associated with the disease. It is still difficult to choose a targeted treatment for a specific condition based on the genetic information of the individual but many

researchers contribute to make it possible. Due to advancement in genetic medicine drugs could be developed and are used today in targeted therapy (e.g., gefitinib in nonsmall cell lung cancer [14], bevacizumab, cetuximab and panitumumab in metastatic colorectal cancer [15]; omalizumab in asthma [16]). Recent studies prove that many more can be designed in the future.

Yohimbine is an indole alkaloid that exhibits high affinity for the α_2-adrenergic receptor. A study shows that it can improve insulin secretion in patients with diabetes type 2 that have a genetic variant of a gene ADRA2A which is encoding the receptor. Different variants have been compared and the author concluded that analysis of the risk variants might result in personalized therapies [17].

Cancer can be treated in many ways, depending on the type, occupied tumor area and tumor development stage. Targeted therapy works differently than nonspecific therapies, such as chemotherapy. It is focused on differences in molecular structure and function between cancer and normal cells to achieve decreased toxicity and increased effectiveness. High-throughput sequencing technologies were used to study the processes of tumorigenesis in pan-cancer, find its drivers and therapeutic agents that target them [18].

Targeted treatment must be always carefully planned and one of the most important parts is evaluating the accurate dose of the right medication. It has been reported that the responses to certain drugs have been associated with the corresponding genetic variations [19].

All the above examples show the importance of personalized medicine and its possibilities. Personalized medicine is becoming a standard practice in medical care that facilitates the process of choosing the most accurate diagnosis and specific treatment plan. IT systems that allow physicians to examine medical history of patients need to adapt to the new category of information provided by genetic testing and next-generation sequencing. The detailed, individual genetic information can help prevent adverse effects, maximize efficacy of the treatment and prevent genetic diseases. Information technology provides specific algorithms to improve those processes.

MedStream Designer enables analyzing both structured and unstructured clinical data in connection with patient's genetic tests. It can help to indicate the interdependence between clinical features and molecular basis of the disease and be used not only in medical practice but also clinical research or for education purposes.

Methods

MedStream Designer consists of five main modules: application server, web application, database, full text search server and an application that imports the data from HIS system. The crucial part of the system is the application server. It carries out all the processes of searching and analysis. It retrieves data from all essential databases. Additionally, it cooperates with the full text search server when the user is searching for a phrase. Web application communicates with the server by the means of a complex web service. MedStream Designer connects through a data import application with the Hospital Information System. This application is responsible for cyclical data synchronization.

The key function of the application is creating search scenarios - a sequence of search criteria that are time-dependent or time-independent. In medical practice, there are multiple parameters that have an impact on the decision-making process. According to Evidence-based medicine, a healthcare professional should be able to compare the patient that is treated with a respective group of patients who had similar clinical features and know how different interventions made an impact on the outcome. In the application, the search conditions refer to the respective elements of the treatment process: diagnosis, medication and medical procedures. Patients search criteria refer to the patient's history recorded in the electronic documentation system, e.g. in CliniNET, KS-MEDIS or other HIS systems.

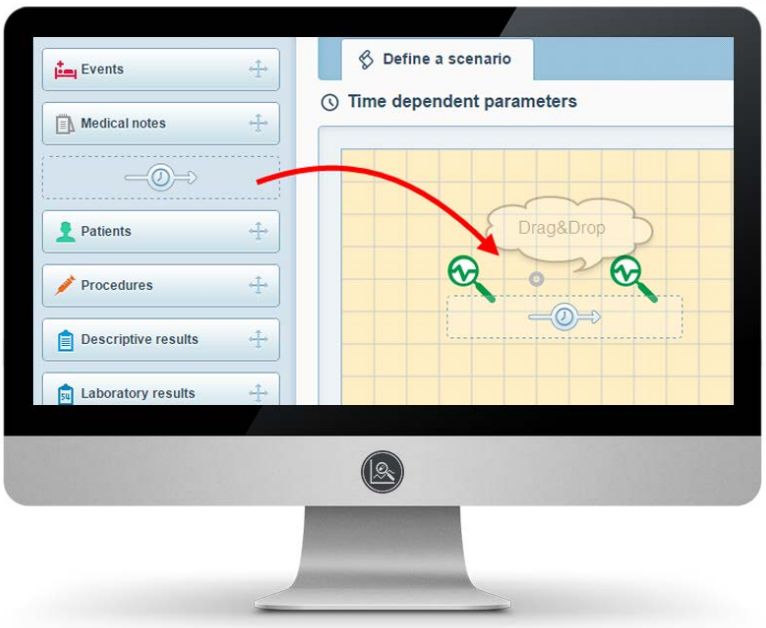

Fig. 1. Creation of a search scenario

MSD has a unique feature, which enables user to take into account time relations between medical parameters to model the treatment path. For example, MSD can be used to find patients who meet the following criteria: Patients diagnosed with diagnosis A, who was later treated with procedure B. Within two days after procedure B drug C or drug D was given. Moreover in two weeks after the procedure, he had blood test at least 3 times and at least two time the level of parameter E was exceeded. And moreover the text note assosiated with the patient has a phrase "headache". So far there was no IT system that allows physicians to easily search for patients that fits to above treatment history.

Graphical user interface based on a timeline has been created to simplify the process of searching for time-dependent elements. A timeline is a squared area on which respective parameters are placed to create a sequence of events. Beneath, there are specific areas for time-independent and additional parameters. Scenario is created by placing parameters in the specific areas based on drag & drop (Fig. 1). There are nine basic parameters: Diagnosis, Medications, Hospitalizations, Medical notes, Patients, Procedures, Descriptive results, Laboratory results and Visits.

Recently, an upgrade of the application has been designed with the possibility to search the database by the results of the High-throughput and Sanger Sequencing. Next-generation sequencing enables to quickly determine the primary structure of the DNA and RNA molecules, find mutated genes and evaluate the level of their expression. NGS offers wide range of applications which suggests that it can become the standard procedure in genetic diagnostic laboratories [20][21].It is an important advancement for the personalized medicine. It can accelerate the early detection of genetic disorders and prognostic markers - screening for the most common disorders could prevent them or improve the treatment. Healthcare professionals obtain the knowledge about the patient on the most detailed, molecular level. Along with the information about patient's history it is a tool that can greatly improve the process of decision-making during treatment.

When the parameter is properly placed, a modal window appears (e.g. Fig. 3) where user can precise the search criteria for the specific parameter (e.g. the ICD-10 code of a diagnosis, body temperature range or key phrases in text notes). MedStream Designer translates search conditions to the natural language and the system checks whether the parameters that have been selected are chronologically arranged and if the values are correctly set. Whole scenario can be viewed without a necessity to open modal window.

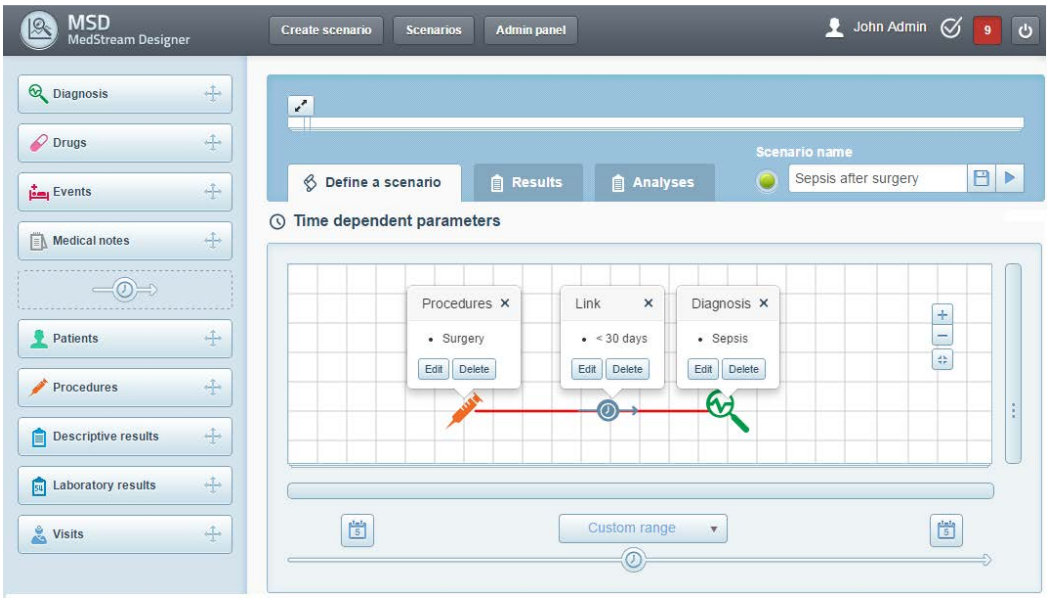

Fig. 2. Sequence of events on the timeline.

Sequence of events is a set of parameters that occurred simultaneously or one after another. These dependencies can be obtained by proper arrangement on the timeline (Fig.3). The elements in the same column are parallel. In the next column with a parameter, an event that occurred afterwards is placed. Precedence is interpreted even if the events are not placed on the same height. Search criteria can be limited to a time window and a connector can be used to precise the time relations between two parameters. Time-independent events can be additionally defined by negation - the application will search for the patients that don't meet the search criteria.

Scenarios can be saved, edited and cloned. To obtain a cohort of patients, a scenario must be launched at a time specified by the user. Depending on hardware resources, there is a limit set for simultaneously processed scenarios.

The mechanism of full text search has been implemented in MedStream Designer and it is used to search the phrases in two types of records: Medical notes and Descriptive results (Fig 3.). All the words and phrases are examined by the search engine to match search criteria. The aim is to turn text into data for analysis independently from the grammatical form of the words or phrases. The user can determine the level of search flexibility to obtain a desirable proportion between precision and overall recall. It is possible to find content based on meaning and context. Text data often contains hidden but highly important information for the treatment process. Healthcare professionals often write high-quality information in the medical notes even though it can be classified otherwise. Reading out cause and effect relationships in those notes simplifies the process of comparative analyses.

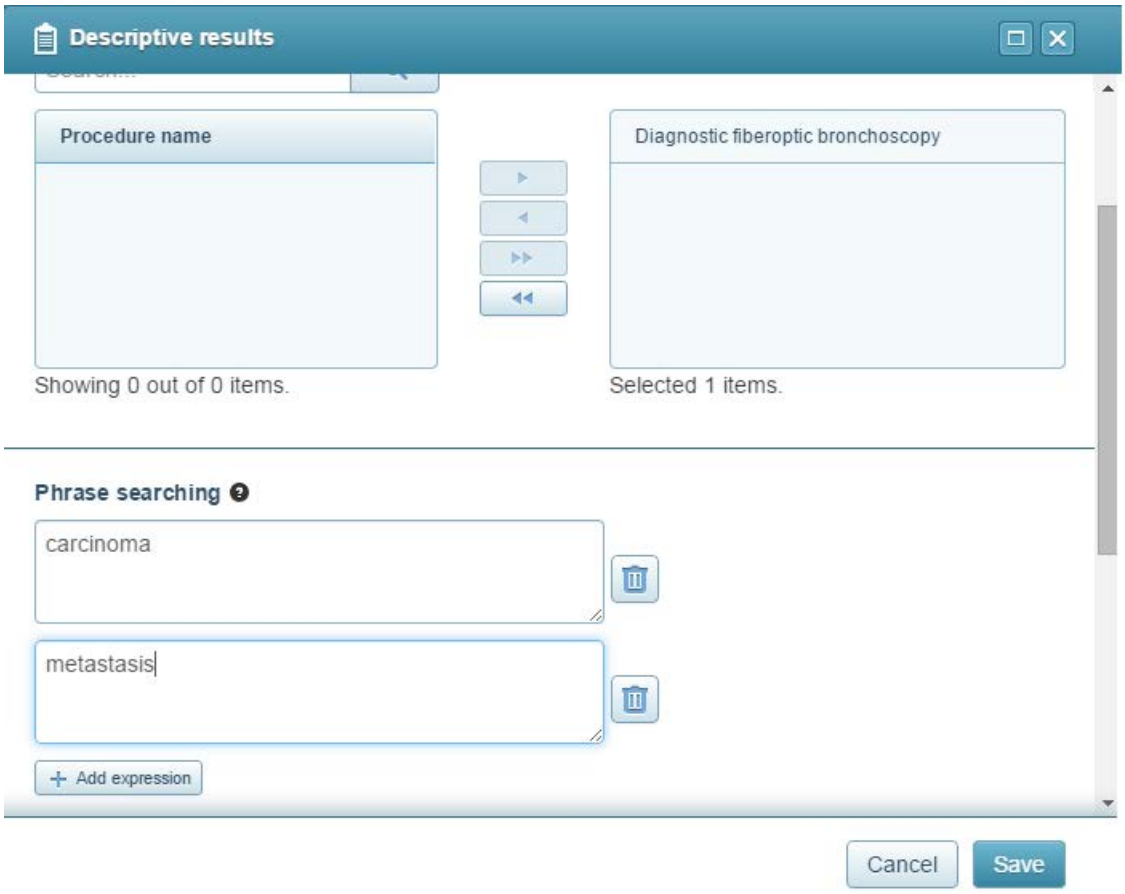

Fig. 3. Definition of phrases for the full text search mechanism.

If the scenario is properly launched, a new bookmark Results will be added to the scenario panel. If the database contains at least one patient that meets the criteria provided by the user, two new bookmarks will appear as well - Statistics and Analyses. The user can see all the patients that has been selected and their history. Results can be exported to XLS or CSV file fully or only in a limited range.

The bookmark Analyses has two aims - it is a tool to view and analyze obtained cohort of patients. Charts demonstrating relations between patients can be generated by the user but before that, a mechanism of preliminary patient analysis in respect of gender and age is launched (Fig. 4) In the Statistics bookmark, the user obtains this data arranged in various ways (histograms, table, box plot).

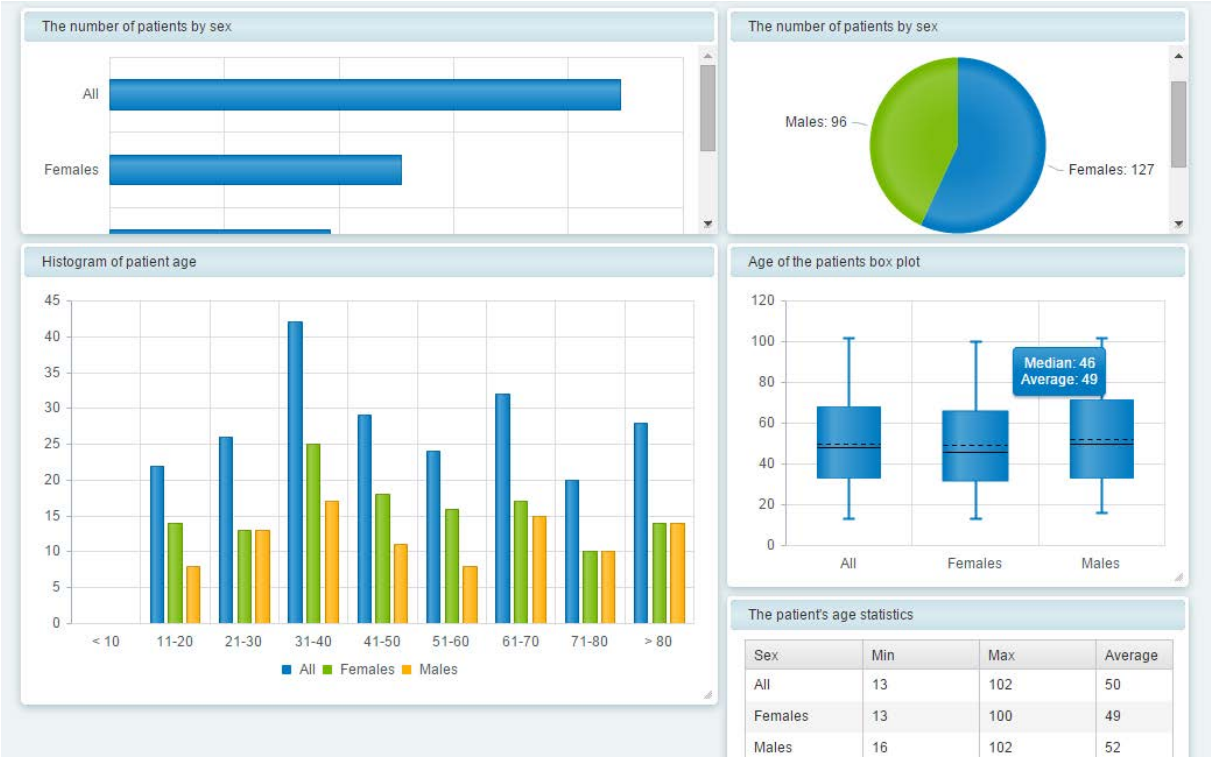

Fig. 4. Preliminary patient analysis in respect of gender and age.

The user is able to choose the parameters and a range of data that will be analyzed as well as the type of analysis that will be performed; e.g. the application can display the chart demonstrating differences in quantities of patients that underwent specific procedures or were diagnosed with indicated diseases.

Results

Currently there are a few implementations of MedStream Designer solution across Poland. The biggest one took place in Maria Curie-Sklodowska Institute of Oncology in Warsaw. The tool is implemented as a client-server infrastructure; thus, all the employees have a direct access to the MSD via any web browser from every computer working in Institute. The implementation process consisted of a few parts, where the most complicated was an implementation of ETL (and. Extract, Transform, Load) tools for data extraction from hospital information system (HIS). The whole process of MSD implementation took about 6 months. The database of Oncology Institute covered almost 20-years history treatment of almost a million of unique patients including information of: diagnosis, drugs, procedures, hospitalizations, laboratory results, medical notes etc. All that information was extracted from HIS system, cleaned, categorized and

then transferred to the newly designed data warehouse. This data are also accessible for any other tools as a structure of data warehouse and is fully described for integration with any other third party tools.

As a final step, about a hundred of Maria Curie-Sklodowska Institute of Oncology employees were trained (doctors, administration employees, managers, researchers and people involved in clinical trials recruitment process).

Conclusions

The advantages for every mentioned group of employees are different. Researchers, employees and people involved in clinical trials recruitment had been given a direct access to the wide group of patients and their full treatment history. Managers can easily verify and monitor the quality of patient's treatment (perioperative infections, prolonged hospitalizations etc.), doctors can benefit from mutual experience and knowledge (treatment regimens, unusual cases, medical notes are now available with a few clicks for everyone).

Summarizing, MedStream Designer lets hospitals take the full advantage of their EMR data. The implementation proved, that the application facilitates the decision-making process, supports Evidence-based Medicine and Personalized Medicine. It is used to find interdependence between important events from patient's history, laboratory and genetic results to provide strong medical evidence for the diagnostic process.

Bibliography

1. B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, and P.G. Shekelle, *Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care*, Annals of Internal Medicine, 2006, no.144, issue 10, p. 742-752.
2. A. W. Wu, H. Kharrazi, L.E. Boulware, C.F. Snyder, *Measure once, cut twice—adding patient-reported outcome measures to the electronic health record for comparative effectiveness research,* Journal of Clinical Epidemiology, 2013, no.66, issue 8, p. 12-20.
3. R. Haux, *Health information systems: past, present, future,* International Journal of Medical Informatics, March-April 2006, no. 75, issue 3, p. 268-281.
4. J.W. Nance Jr, C. Meenan, and P.G. Nagy, *The future of the radiology information system*, American Journal of Roentgenology, May 2013, no. 200, issue 5, p. 1064-1070.
5. Peng and S. Kurnia. Understanding hospital information systems adoption in china. PACIS 2010 Proceedings, 2010. Paper 66.

6. N. Ashrafi, L. Kelleher, and J-P. Kuilboer, *The impact of business intelligence on healthcare delivery in the usa,* Interdisciplinary Journal of Information, Knowledge, and Management, 2014, no. 9, p. 117-130.
7. M.M. Sanja, *Impact of enterprise resource planning system in health care,* International Journal of Academic Research in Business and Social Sciences, December 2013, no. 9, p.117-130.
8. World Health Organization, *Framework and standards for country health information systems / Health Metrics Network. – 2nd ed.,* 2008, p. 1-2, ISBN: 978 92 4 159594 0.
9. T.D. Noakes, J. Borresen, T. Hew-Butler, M.I. Lambert, and E. Jordaan, *Semmelweis and the aetiology of puerperal sepsis 160 years on: an historical review,* Epidemiology and Infection, 2007, no. 136, Issue 1, p. 1-9.
10. N.K. Wilton and A.M. Slim, *Application of the principles of evidence-based medicine to patient care,* Southern Medical Journal, 2012, no. 105, issue 3, p. 136-143.
11. C. Gonzaga-Jauregui, J.R. Lupski, and R.A. Gibbs, *Human genome sequencing in health and disease,* Annual review of medicine, 2012, no. 63, pages 35-61.
12. M. Wuttke and A. Kottgen, *Insights into kidney diseases from genome-wide association studies*, Nature Reviews Nephrology, 2016, no. 12, Issue 9, p. 549-562.
13. Y. Mizusawa, *Recent advances in genetic testing and counseling for inherited arrhythmias*, Journal of Arrhythmia, 2016, no. 32, Issue 5, p. 389-397.
14. F. Nurwidya, F. Takahashi, and K. Takahashi, *Gefitinib in the treatment of nonsmall cell lung cancer with activating epidermal growth factor receptor mutation. Journal of Natural Science,* Biology and Medicine, 2016, no. 7, pages 119-123.
15. Y. Ohhara, N. Fukuda, S. Takeuchi, R. Honma, Y. Shimizu, I. Kinoshita, and H. Dosaka-Akita, *Role of targeted therapy in metastatic colorectal cancer,* World Journal of Gastrointestinal Oncology, 2016, no. 15, Issue 8, p. 642-655.
16. C-H. Lin and S-L. Cheng, *A review of omalizumab for the management of severe asthma*, Drug Design, Development and Therapy, 2016.
17. A. Colmone, *Personalized diabetes therapy*, Science Signaling, 2014, no. 7, Issue 347, p. ec288-ec288.
18. A. Bhardwaj, *Personalized cancer medicines*, Science Translational Medicine, 2015, no. 7.
19. D. Voora, *Predicting the era of personalized medicine,* Science Translational Medicine, 2014, no 6, Issue 222, 2014.
20. V. Nigro and G. Piluso, *Next generation sequencing (ngs) strategies for the genetic testing of myopathies*, Acta Myologica, 2012.
21. G. Chen, Z. Yang, J.R. Eshleman, G.J. Netto, and M-T. Lin, *Molecular diagnostics for precision medicine in colorectal cancer: Current status and future perspective,* BioMed Research International, 2016.