Original Research Article

TJ Tumori Journal

Tumori Journal
1–9

DOI: 10.1177/0300891618792473
journals.sagepub.com/home/tmj
SAGE

Cancer precision medicine today: Towards omic information in healthcare systems

Norbert Maggi[1], Roberta Gazzarata[2], Carmelina Ruggiero[1,2], Claudio Lombardo[3,4] and Mauro Giacomini[1,2,5]

Abstract

Introduction: This article focuses on the integration of omics data in electronic health records and on interoperability aspects relating to big data analysis for precision medicine.

Methods: Omics data integration methods for electronic health record and for systems interoperability are considered, with special reference to the high number of specific software tools used to manage different aspects of patient treatment. This is an important barrier against the use of this integrated approach in daily clinical routine.

Results: The correct use of all three levels of interoperability (technical, semantic, and process interoperability) plays a key role in order to achieve an easy access to a significant amount of data, all with correct contextualization, which is the only way to obtain a real value from data for precision medicine.

Conclusions: The proposed architecture could improve the potentialities of data routinely collected in many health information systems to form a real patient center information environment.

Keywords

Electronic health records, electronic medical records, precision medicine, personalized medicine, interoperability

Date received: 19 February 2018; revised: 14 June 2018; accepted: 13 July 2018

Introduction

In 1999, personalized medicine was introduced as relates to cancer, focusing on the drugs targeting for each genetic profile.[1] At the time, the human genome project,[2,3] aiming to obtain an accurate sequence of the majority of the euchromatic portion of the human genome, had been carried out for several years. The project, sponsored by the US federal government, had been launched in 1990 and initially focused on two main aspects; that is, the mapping of the human and mouse genomes for the study of inherited diseases and genome assembly and the sequencing of organisms with smaller genomes. The success of both aspects was the basis of the human genome sequencing. In order to carry out these aspects of the human genome project, the International Human Genome Sequencing Consortium, which involved 20 centers in six countries, was formed. Moreover, the Human Genome Organization was established, in parallel with the human genome project in 1989. It is grounded in Geneva, Switzerland, and the founding council initially had 42 scientists from 17 different countries and later 220 members. Its mission was promoting fundamental genomic research, fostering scientific exchange in genomics with particular emphasis on scientifically developing and emerging countries, and globally supporting the ethics of genetics and genomics. In

[1]Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genova, Genova, Italy
[2]Healthropy s.r.l., Savona, Italy
[3]Sos Europe s.r.l., Genova, Italy
[4]Organisation of European Cancer Institutes–European Economic Interest Grouping, Brussels, Belgium
[5]Centre of Excellence for the study of molecular mechanisms involved in cell-to-cell communication: from basic research to the clinic (CEBR), Genova, Italy

Corresponding author:
Mauro Giacomini, Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genova, Via all'Opera Pia 13, Genova 16145, Italy.
Email: mauro.giacomini@dibris.unige.it

addition, a project on mapping human genes was launched in 1999. The project involved 10 of the largest drug companies and was focused on the identification of several hundred thousand chemical signposts to explore regions of the human DNA. This mapping project complemented the human genome project and focused on single nucleotide polymorphisms; that is, slight genetic variations between human beings: single nucleotide polymorphisms may make some people more predisposed to specific diseases and explain different response to the same drug by different individuals.

In 2001, the International Human Genome Sequencing Consortium[4] and Celera Genomics[5] provided a first description of the human genome, focused on gene identification, polymorphism, and other aspects. In the following years, further detailed features were clarified.[3]

In 2004, the Personalized Medicine Coalition was formed in the United States. This nonprofit organization of companies, health care providers and payers, patient groups, industry organizations, academic institutions, and government agencies addressed the changes required by personalized medicine. Specific requirements as relates to health care institutions, diagnostic and therapeutic business models, reimbursement policy, and regulatory oversight were considered. Moreover, electronic medical records (EMR), interoperability, decision support systems, and other aspects of information technology received attention.

The availability of improved genomic tools has improved the understanding of diseases at the molecular level, originating the development of molecularly targeted therapies and improved diagnosis criteria.

In 2008, the President's Council of Advisors on Science and Technology (PCAST) published the report "Priorities for personalized medicine," which summarized the results of a review with input from industry, physicians, patients, scientists, and US government agencies. The report provides a picture of the potential core of personalized medicine to reshape healthcare provision and economics. This report provides a comprehensive, helpful, and relevant definition: "'Personalized medicine' refers to the tailoring of medical treatment to the individual characteristics of each patient. It does not literally mean the creation of drugs or medical devices that are unique to a patient but rather the ability to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventive or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not." PCAST remarks that recent advances in genomics show a great number of possible genome-related new molecular markers for disease presence, individual risk of disease, and different response to treatment by different patients. In addition to the possibilities of improving patient care and disease prevention, personalized medicine has the potential to impact healthcare costs and the development of new medical products, identifying in advance patients who will benefit from a specific treatment and patients who are likely to experience adverse effects. These aspects can bring about cost savings for healthcare. Moreover, size, duration, and costs of clinical trials could be reduced. PCAST reckons that genomics-based molecular diagnostics is likely to speed up personalized medicine progress. Many examples can be identified in which personalized medicine influences clinical decisions and contributes to improve healthcare provision.

In 2015, President Obama announced that he was launching the Precision Medicine Initiative, defined as a new research effort to revolutionize the way in which health is improved and disease can be treated, taking into account individual differences in people, genes, environment, and lifestyle. Examples were given on the way in which diseases such as cancer could be treated, improving survival chances and reducing adverse effects. Relevant investments were made since then in order to support research, development, and innovation in this respect.

The term "personalized medicine" was used at the beginning of century, following the human genome project achievements. The term precision medicine attracted attention after the US National Research Council published a report on precision medicine in 2011 in which they proposed guidelines for modernization of disease taxonomy tacking into account molecular information rather than classification based on symptoms.[6]

Recently there has been a shift from personalized medicine towards precision medicine. In 2005, a PubMed query showed that there was one paper mentioning precision medicine and 74 mentioning personalized medicine, whereas in 2016, 3020 mentioned precision medicine and 1857 mentioned personalized medicine. Personalized medicine was mentioned increasingly starting from 2015, after the launch of the precision medicine initiative (Figures 1 and 2).

There is much overlap between the terms "precision medicine" and "personalized medicine." According to the US National Cancer Institute, personalized medicine is an older term that could be misinterpreted, implying that treatments and prevention could be developed for each specific individual, whereas precision medicine focuses on the identification of approaches for specific patients according to genetic environmental and lifestyle aspects, with special reference to pharmacogenomics. The council's preferred term is "precision medicine" rather than "personalized medicine"; however, the two terms are often used interchangeably.

In Europe, several workshops on different aspects of personalized medicine were carried out since 2010, and this led to a conference on perspectives in personalized medicine, which was organized by the European Commission in 2011. The first European policy document in this field was published in 2013.[7] It focused on omics

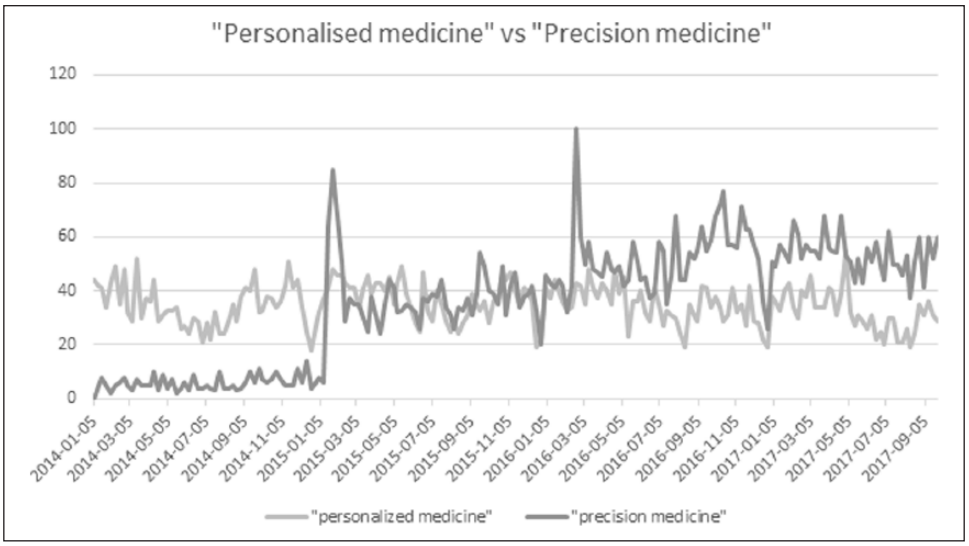

Figure 1. Comparison of frequency of search for the terms "personalized medicine" and "precision medicine." Data from Google Trends.

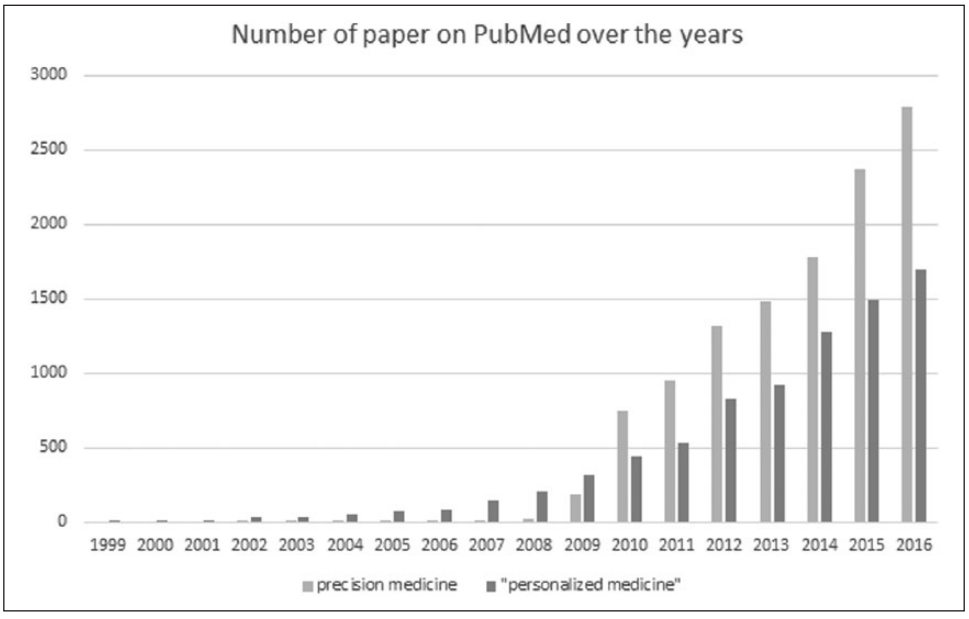

Figure 2. Number of papers indexed in PubMed over the years.

technologies and examined EU healthcare systems in this respect. In 2015, health ministers discussed ways to advance personalized medicine and member states were encouraged to promote education, training, and professional development for health professionals.

In Europe, the Organisation for European Cancer Institutes, which has set up the Education and Training Working Group[8] involving Alleanza contro il cancro (ACC), the Italian Network Comprehensive Cancer Centres, and the European Association for Cancer Research, have focused on novel topics in precision medicine with special reference to genetic heterogeneity and pharmacoepigenetics. Since 2015, precision medicine for cancer meetings have been organized.

At the same time, several initiatives were undertaken in EU member states. In the United Kingdom, the Academy of Medical Sciences organized workshops and conferences. The German Academy of Sciences Leopoldina published a report on individualized medicine and the German Ministry of Education and Research set up an action plan for individualized medicine. The French National Alliance for Life Sciences and Health issued its genomic medicine 2025 plan. Moreover, activities in the field were presented at a personalized medicine conference in 2016 in Brussels.[9]

The Organisation for Economic Cooperation and Development is carrying out a project on the potential of emerging technology for health, in which the convergence of technology especially biotechnology, information technology, and nanotechnology on precision medicine are analyzed. This topic has a key position within Europe Horizon 2020, US National Institutes of Health, Canadian Institutes of Health Research, and China.

The heterogeneity of cancer and the human genome variations between individuals are the basis of a more personalized cancer treatment. Current genome technologies make it now more feasible to match treatment to cancer features, selecting optimal drugs and drug dosage for each specific patient, improving therapy outcomes. Personalized cancer care, however, still has limitations as relates to the understanding of cancer biology and the identification of molecular targets related to tumor progression. New directions in cancer therapeutics include the development of targeted therapies to interrupt critical molecular pathways, molecular profiling of tumors, development of gene expression signatures related to specific drugs, development of vaccine therapies, and immunology approaches.[10]

Molecular diagnosis of cancer is based on a great number of genes that are known to be involved. The recent developments of next-generation sequencing (NGS) together with biomedical informatics data analysis make it possible to obtain a greater amount of information about the molecular biology of tumors.

In addition, The Cancer Genome Atlas (TCGA), a project of the Center for Cancer Genomics at the National Cancer Institute, terminated in 2017, that collected, selected, and analyzed human tissues for large-scale genomic alterations, has achieved a high impact on the taxonomy of human tumors. Building on the success of TCGA, further projects have been launched to integrate the data generated so far. From a precision oncology perspective, these advances will be useful for cancer patients.[11]

Next-generation sequencing

DNA sequencing and genomics derive from the combination of molecular biology and of the chemistry of nucleotides, which are linked together forming the building blocks of DNA. In the 1970s, several methods were proposed for determining nucleotide sequences in DNA simply and rapidly. In 1975, Sanger and Coulson[12] proposed a method based on the use of DNA polymerase to transcribe a particular region of the DNA. The use of this method made it possible to determine the sequence of the genome of the bacteriophage φX174.[13] Other methods for determining nucleotide sequences in DNA were described at the time. Barnes[14] described a method based on partial ribosubstitution allowing the determination of 100 nucleotides. Maxam and Gilbert[15] described a method based on specific chemical degradation of DNA that can be applied to double-stranded DNA, permitting sequencing of at least 100 bases. Sanger et al.[16] in 1977 described a technique that is more rapid and accurate than that described in reference 12 and only required the commercially available DNA polymerase.

At the time, DNA sequencing achieved the determination of a few hundred nucleotides at a time. In the following decade, the enzyme-based method[16] and the chemical degradation method[15] were used to a great extent. Both methods were based on radiolabeled DNA fragments and based on manual and monotonous procedure. Subsequently efforts have been made to automate several steps in order to make the acquisition of DNA sequence information more rapid.[17,18] The development of capillary array electrophoresis[19,20] was a breakthrough in DNA sequencing leading to the production of a commercial single capillary sequencer (ABI Prism 310) and in 1998 the MegaBACE 1000 with 96 capillary sequencers become commercially available.[21] This sequencer was the first commercial high-throughput system. Other sequencing strategies have subsequently become available. Since 2006, a great number of methods and techniques became available. This led to technology platforms different from the Sanger method, for massively parallel analysis with reduced cost. At the time, the costs of human genome sequencing dropped by 50,000, with respect to the costs at the time of the human genome project.[22] The method developed by Sanger in 1977, also called chain-termination method, determined the sequence of nucleotide coping DNA, which was being analyzed many times obtaining fragments of different lengths that terminated at different points, which were labeled by fluorescent nucleotides that marked the end of the fragments. This method is still used for sequencing fragments of DNA; however, some difficulties of this method were the fact that it required considerable time and effort and the use of radioactive materials. In 1987, Applied Biosystems adopted capillary electrophoresis, which improves speed and accuracy of the method. The instruments based on capillary sequencing and related software, together with Sanger sequencing technology, were the basis for the human genome project that was completed in 2001. After this, many efforts were made to increase speed and accuracy and reduce costs by high-throughput and parallel analysis.

The term NGS appeared then, with reference to highly parallel sequencing methods for genome analysis.[23,24]

Some aspects of NGS approaches are outlined below.

The first NGS sequencing platform on the market was the Roche 454GSFLX titanium sequencer, which was released in 2005 and discontinued in 2016.[25] It was developed using pyrosequencing technology,[26] which is based on the measurement of the release of inorganic pyrophosphate during the polymerization of DNA molecule, in which by-product pyrophosphates are released at each single nucleotide addition. These pyrophosphates are converted into

visible light by means of a sequence of enzymatic reactions.

In 2006, Solexa launched the Genome Analyser, and was acquired by Illumina the following year. This technology differs from the 454GSFLX owing to the amplification strategy used and because the dye-labeled nucleotides are added simultaneously, whereas in the 454GSFLX they are added sequentially.

In 2007, the Sequencing by Oligonucleotide Ligation and Detection (SOLiD) genome sequencer was released by Applied Biosystems. This platform uses DNA ligase to detect and incorporate bases in a very specific way instead of DNA polymerase.

From 2008 to 2009, the improvement of technologies led to the development of third-generation sequencing. These platforms complement the previous technologies with improved features, such as single-molecule template, lower cost per base, easier sample preparation, considerably faster run, and improved data analysis.[27] In this respect, several platforms have been released, such as Ion Torrent and PacBio sequencers.

The 454 pyrosequencing approach was adopted and modified by the Ion Torrent platform, in which the H^+ ions that are released at every nucleotide addition are used instead of the pyrophosphates,[28] causing a change in pH that can be detected by semiconductor and field effect transistor technology. The Ion Torrent sequencer was released in 2011.

PacBio technology was also made available in 2011. It is based on a nanophotonic tool called zero mode waveguide (ZMW). ZMW allows polymerization of DNA in real time. The sequences are read through fluorescently labeled nucleotides by measuring the bursts of light that are released during the polymerization reaction, like in the 454 and Illumina sequencers.

An interesting alternative to the PacBio has been developed. In 2014, the minION desktop sequencing instrument was launched. This device does not require DNA synthesis or DNA hybridization, but it is based on the translocation of a DNA strand through a pore, which is suspended within a membrane. During DNA translocation, a modulation of an electric current that passes through each pores takes place, and this originates a shift that characterizes the base or bases that are in the pore at a specific time.[29] This device only needs a USB port or a mobile phone or tablet and data analysis can be carried out by an Internet connection and cloud applications. The potential of the minion sequencer is high, mostly because of its small size and limited equipment cost, especially when fast and reliable sequencing is necessary when resources are restricted.[30] Even though nanopore sequencing has been so far mainly applied to small genomes, a new approach has been developed that allows portable de novo sequencing of human genomes, which may allow quick point of care diagnosis of cancer and cancer progression monitoring.[31]

Genomics, which deals with the determination of the genomic sequence of a specific organism, is the basis of functional genomics, which is the study of the function of the genes in one genome, and of comparative genomics, which is comparing the genes of different organisms and of structural genomics, which is the study of the 3D structure of proteins, obtaining clues to their function.

Proteomics focuses on the study of protein structure and function, and specifically on several protein function aspects, such as expression profiling, which identifies proteins as a result of the expression to a stimulus, protein–protein interaction networks, environmental aspects, and others. A more recent development originating from genomics is metabolomics, which focuses on chemical processes involving metabolites, which are produced by cellular processes.

Transcriptomics focuses on the RNA transcripts produced by the genome, aiming to understand the way in which altered expression of genes may contribute to complex diseases such as cancer.

Genomics in EHR

Clinical decision support by health care providers uses electronic records. EMR are electronic versions of paper charts summarizing a patient's medical and treatment report within a clinical setting, whereas electronic health records (EHR) incorporate information that is intended to be shared among setting and healthcare systems.

The data in EHR are various, and include drug prescription information and descriptions of medical aspects that underline prescriptions. A health record can be regarded as an information repository on the health status of a patient as machine-readable data (or metadata). EHR data have been classified as ancillary clinical data and clinical text.[32] Administrative data and ancillary clinical data are structured, readily available for insertion and for extraction from databases. Administrative data consist of demographic patient information or financial data. Ancillary clinical data consist of information from laboratories, medical imaging, and clinical drugs. Clinical text consists of unstructured data such as admission and discharge notes, treatment plans, and daily observation notes (Figure 3).

At present, no current standard for EHR is established for the integration of cancer omics data within an EHR,[33] but the incorporation of omic information into an EHR is regarded as important.

Interoperability is critical for precision medicine, especially for big data analysis. The Institute of Electrical and Electronic Engineers defines interoperability as the "ability of a system or a product to work with other systems or products without special effort on the part of the customer" and states that it "is made possible by the implementation of standards." The Health Level Seven International

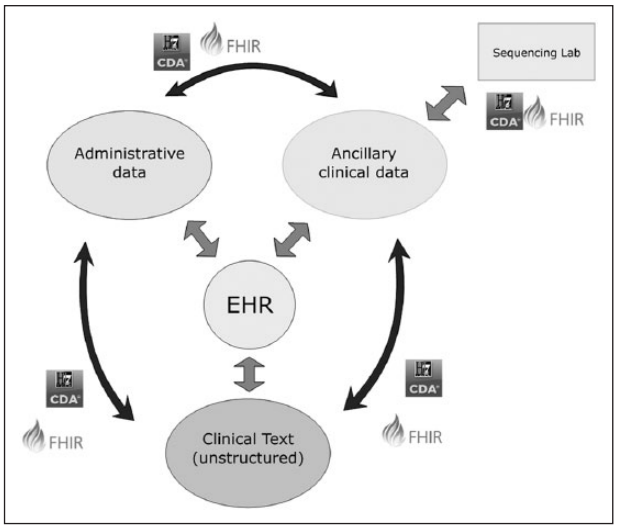

Figure 3. Overview of electronic health record (EHR) components and their mutual interactions. FHIR: Fast Healthcare Interoperability Resource.

(HL7), a nonprofit standards developing organization, which provides a widespread framework for exchange, integration, sharing, and retrieval of electronic health integration, proposed many different frameworks to support format interoperability. Since the beginning of the 2000s, the Clinical Document Architecture (CDA) (Release 2)[34] provides a widely used structure to share clinical data among health organizations.[35] In some cases, it has also been used for connecting omics data.[36] More recently, HL7 proposed the Fast Healthcare Interoperability Resource (FHIR) standard for interoperability. FHIR allows the use of omic information together with other EHR data in order to identify the most effective treatment for a specific patient, which is the aim of precision medicine.[37]

Interoperability can be divided into three layers: technical, semantic, and process interoperability.[38] All three levels of interoperability are interfering: semantic interoperability requires technical interoperability while process interoperability requires semantic interoperability.

Technical interoperability is "the ability to move data from one system (A) to another (B). It defines the degree to which the information can be successfully transported between systems." Service-oriented architecture (SOA) is the design strategy most adopted to support technical interoperability between real-time applications implemented within large-scale distributed environments. The main reason for the diffusion of the SOA paradigm is that it proposes a highly feasible approach to promote the easy integration and alignment of new and existing solutions into a cohesive architecture, all with minimal impact to service consumers with a resulting highly reduced economic cost.[39,40] For these reasons, this approach was successfully adopted in distributed healthcare architectures.[41,42]

Semantic interoperability "ensures that both systems understand the data in the same way: the information sent is unaltered in its meaning." Unlike technical interoperability, which is realized with common technologies in all information technology sectors, semantic interoperability depends on the specific application field. In health informatics, this interoperability layer is guaranteed by the adoption of standards to manage both syntax and semantics, produced by different international initiatives. The standardization efforts produced different data models (e.g. HL7 version 3 Reference Information Model [RIM] and Informatics for Integrating Biology and the Bedside [i2b2][43]) to manage syntax while semantics is guaranteed by the adoption of different vocabulary representations (e.g. SNOMED CT,[44] LOINC,[45] International Classification of Diseases [ICD]). Some of these standards were adopted to support semantic interoperability in different solutions to manage multicentric clinical trials.[41,42,46]

Finally, process interoperability "enables business processes and organizations housing systems A and B to work together." It defines the degree to which the integrity of workflow processes can be maintained between systems. This includes maintaining/conveying information such as user roles between systems. The process interoperability requirement is satisfied when a process is compliant with standards that allow it to reach its own objective, irrespective of the propriety, location, version, and design of the IT systems used. To address this need in e-health, the Healthcare Services Specification Project (HSSP) was promoted.[47] The HSSP was formed in 2005, by HL7 International and the Object Management Group, in order to define health industry SOA standards that promote effective interoperability among applications and distributed and heterogeneous devices that belong to independent socio–health system organizations. An example of a solution to support process interoperability in multicentric clinical trials is presented in reference 48.

Another important aspect is that the infrastructure must guarantee high availability to the involved centers and must be easily scalable to manage the increase of involved hospitals. Cloud computing represents a suitable solution to support these needs. In addition, security and privacy issues must be considered, which can be managed by access control services, audit services, and trust services.[41,49]

Figure 4 shows the infrastructure proposed to support all the mentioned needs to manage multicentric approaches.

In this respect, clinical decision support tools can interact with clinicians for decision-making and helping them to comply with guidelines, improving healthcare outcomes. A key point in omics-based clinical decision support is achieving a standardized vocabulary for omics variants. This aspect is being addressed by the HL7 genomics work group.

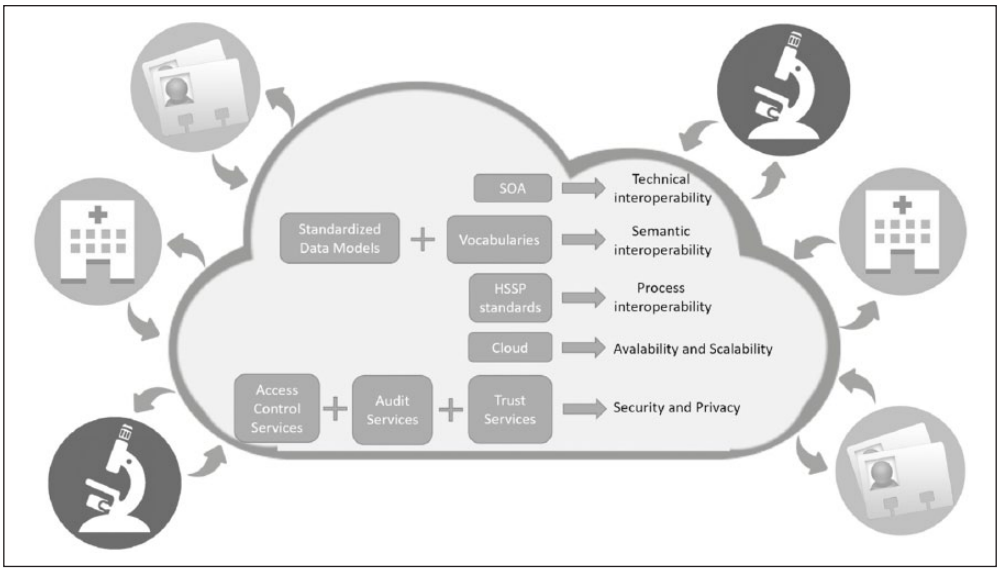

Figure 4. The proposed infrastructure to manage the multicentric approach. HSSP: Healthcare Services Specification Project; SOA: service-oriented architecture.

Impact of precision medicine on health systems

Precision medicine for oncology is important for healthcare systems. However, its implementation in the clinic means solving problems such as genomic results analysis, targeted treatment, and the implementation of therapies based on genomic tests. Precision cancer medicine implementation has evidenced the need for integrated information technology tools for genomic data sharing and interpretation.[50] Precision medicine is able to increase the impact of existing treatments by improving the effectiveness of treatment for a given patient and by increasing awareness of the risks of serious side effects. In recent years, several initiatives related to precision medicine have taken place in various countries such as the United States, China, and Australia, and in Europe (Denmark, England, France, Finland, Germany, the Netherlands). In addition, various government bodies are in synergy to support these initiatives through the formation of international consortia such as the International Consortium for Personalized Medicine (www.icpermed.eu) or the European alliance for Personalized Medicine (www.euapm.eu). In Italy, an oncogenomics program (ACC genomics) has been started with the aim of raising the quality of care to cancer patients by detecting all the genetic alterations of tumors that may constitute a target for new molecular drugs. The program was implemented by the Italian Ministry of Health through ACC, the largest Italian cancer research network.

In addition, several private companies are entering the market, such as Foundation Medicine (which started providing analysis services in Italy in June 2018), which has already analyzed thousands of patients through its commercial platform, providing an effective strategy to help identify both common mutations and rare mutations of a patient's tumor, to allow the identification of targeted therapeutic options.

In this context of continuous development of precision medicine in which a large amount of information (big data) is collected and analyzed, it is necessary that the confidentiality of sensitive information is guaranteed. Moreover, if the samples are used for subsequent research, the property rights on the samples and the validity of the consent that has been provided must be clarified. Anonymization techniques can be successfully used to guarantee citizens' rights.

Conclusions

Omics data is a leading aspect of personalized medicine. Rapid advances in high-throughput technologies of biomedical informatics aspects such as database storage and big data management, integration of omics within EHR, and interoperability have the potential to increase the weight of precision medicine within clinical practice.

Widespread use of informatics and decision support tools plays a key role in enabling care providers to use omics at the point of care. At present, the high number of specific software tools used to manage different aspects of patient treatment is an important barrier against the use of this integrated approach in daily clinical routine. The correct use of all three levels of interoperability (technical, semantic, and process interoperability) can help enable easy access to a substantial amount of data with correct contextualization, to obtain real value from data for precision medicine. The proposed architecture could improve the potentialities of data routinely collected in many health

information systems to form a patient-centered information environment.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Langreth R and Waldholz M. New era of personalized medicine: targeting drugs for each unique genetic profile. *Oncologist* 1999; 4: 426–427.
2. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004; 431: 931–945.
3. Stein LD. Human genome: end of the beginning. *Nature* 2004; 431: 915–916.
4. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860–921.
5. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001; 291: 1304–1351.
6. Katsnelson A. Momentum grows to make 'personalized' medicine more 'precise.' *Nat Med* 2013; 19: 249.
7. European Commission 25.10.2013 SWD (2013) 436 final. *Commission staff working document: use of '–omics' technologies in the development of personalised medicine*, https://ec.europa.eu/research/health/pdf/2013-10_personalised_medicine_en.pdf (accessed January 2018).
8. Lombardo C, Albanese D, Belardelli F, et al. Training and mobility: a priority for the organisation of the European Cancer Institutes: how a national mobility initiative could enhance EU cooperation in cancer research contributing to the development of a European research area: the example of the Italian Comprehensive Cancer Centers' Network "Alleanza Contro il Cancro." *Tumori* 2018; 94: 147–153.
9. Nimmesgern E, Benediktsson I and Norstedt I. Personalized medicine in Europe. *Clin Transl Sci* 2017; 10: 61–63.
10. Schilsky RL. Personalized medicine in oncology: the future is now. *Nat Rev Drug Discov* 2010; 9: 363–366.
11. The TCGA Legacy. *Cell* 2018; 173: 281–282.
12. Sanger F and Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 1975; 94: 441–448.
13. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 1977; 265: 687–695.
14. Barnes WM. DNA sequencing by partial ribosubstitution. *J Mol Biol* 1978; 119: 83–99.
15. Maxam AM and Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA* 1977; 74: 560–564.
16. Sanger F, Nicklen S and Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74: 5463–5467.
17. Smith LM, Sanders JZ, Kaiser RJ, et al. Fluorescence detection in automated DNA sequence analysis. *Nature* 1986; 321: 674–679.
18. Swerdlow H and Gesteland R. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nucleic Acids Res* 1990; 18: 1415–1419.
19. Mathies RA and Huang XC. Capillary array electrophoresis: an approach to high-speed, high-throughput DNA sequencing. *Nature* 1992; 359: 167–169.
20. Pang HM, Pavski V and Yeung ES. DNA sequencing using 96-capillary array electrophoresis. *J Biochem Biophys Methods* 1999; 41: 121–132.
21. Dolník V. DNA sequencing by capillary electrophoresis (review). *J Biochem Biophys Methods* 1999; 41: 103–119.
22. Goodwin S, McPherson JD and McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 2016; 17: 333–351.
23. Kircher M and Kelso J. High-throughput DNA sequencing: concepts and limitations. *Bioessays* 2010; 32: 524–536.
24. Levy SE and Myers RM. Advancements in next-generation sequencing. *Annu Rev Genomics Hum Genet* 2016; 17: 95–115.
25. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005; 437: 376–380.
26. Uhlen M. Magnetic separation of DNA. *Nature* 1989; 340: 733–734.
27. Munroe DJ and Harris TJ. Third-generation sequencing fireworks at Marco Island. *Nat Biotechnol* 2010; 28: 426–428.
28. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011; 475: 348–352.
29. Goodwin S, Wappel R and McCombie WR. 1D Genome Sequencing on the Oxford Nanopore MinION. *Curr Protoc Hum Genet* 2017; 94: 1811–1814.
30. Lu H, Giordano F and Ning Z. Oxford nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics* 2016; 14: 265–279.
31. Loose MW. The potential impact of nanopore sequencing on human genetics. *Hum Mol Genet* 2017; 26: R202–R207.
32. Jensen PB, Jensen LJ and Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 2012; 13: 395–405.
33. Warner JL, Jain SK and Levy MA. Integrating cancer genomic data into electronic health records. *Genome Med* 2016; 8: 113.
34. Dolin RH, Alschuler L, Boyer S, et al. HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc* 2006; 13: 30–39.
35. Muller ML, Uckert F, Burkle T and Prokosch HU. Cross-institutional data exchange using the clinical document architecture (CDA). *Int J Med Inform* 2005; 74: 245–256.
36. Bellazzi R, Masseroli M, Murphy S, Shabo A and Romano P. Clinical Bioinformatics: challenges and opportunities. *BMC Bioinformatics* 2012; 13 (Suppl 14): S1.
37. Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R and Wang MD. -Omic and Electronic health record big data analytics for precision medicine. *IEEE Trans Biomed Eng* 2017; 64: 263–273.
38. Greiner U, Legner C, Lippe S and Wende K. Business interoperability profiles: relating business interoperability issues to technical interoperability solutions. *Enterprise*

Interoperability II: New Challenges and Approaches. 2007; 865–877.
39. Gazzarata R, Vergari F, Cinotti TS and Giacomini M. A standardized SOA for clinical data interchange in a cardiac telemonitoring environment. *IEEE J Biomed Health Inform* 2014; 18: 1764–1774.
40. Vasilescu E and Mun SK. Service oriented architecture (SOA) implications for large scale distributed health care enterprises. *1st Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare, Conference Proceedings*, 2006.
41. Kondylakis H, Claerhout B, Keyur M, et al. The INTEGRATE project: delivering solutions for efficient multi-centric clinical research and trials. *J Biomed Inform* 2016; 62: 32–47.
42. Alonso-Calvo R, Paraiso-Medina S, Perez-Rey D, et al. A semantic interoperability approach to support integration of gene expression and clinical data in breast cancer. *Comput Biol Med* 2017; 87: 179–186.
43. Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc* 2010; 17: 124–130.
44. Donnelly K. SNOMED-CT: The advanced terminology and coding system for eHealth. *Stud Health Technol Inform* 2006; 121: 279–290.
45. McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem* 2003.49:624–633.
46. Alonso-Calvo R, Perez-Rey D, Paraiso-Medina S, Claerhout B, Hennebert P and Bucur A. Enabling semantic interoperability in multi-centric clinical trials on breast cancer. *Comput Methods Prog Biomed* 2015; 118: 322–329.
47. Kawamoto K, Honey A and Rubin K. The HL7-OMG Healthcare Services Specification Project: motivation, methodology, and deliverables for enabling a semantically interoperable service-oriented architecture for healthcare. *J Am Med Inform Assoc* 2009; 16: 874–881.
48. Gazzarata R, Giannini B and Giacomini M. A SOA-based platform to support clinical data sharing. *J Healthc Eng* 2017; 1–24.
49. Gazzarata G, Gazzarata R and Giacomini M. A standardized SOA based solution to guarantee the secure access to EHR. *Conference on Enterprise Information Systems/International Conference on Project Management/Conference on Health and Social Care Information Systems and Technologies, Centeris/Projman/HCIST* 2015; 64: 1124–1129.
50. Nadauld LD, Ford JM, Pritchard D and Brown T. Strategies for clinical implementation: precision oncology at three distinct institutions. *Health Aff* 2018; 37: 751–756.