Towards Integration Exposome Data and Personal Health Records in the Age of IoT

[1] Savoska Snezana [0000-0002-0539-1771], [2] Blagoj Ristevski [0000-0002-8356-1203],

[3] Natasha Blazheska –Tabakovska [0000-0002-8715-3119] and [4] Ilija Jolevski [0000-0003-2262-9638]
[1, 2, 3, 4] "St. Kliment Ohridski" University - Bitola, Faculty of Information and Communication Technologies - Bitola, ul. Partizanska bb 7000, RN Macedonia
[1] snezana.savoska@uklo.edu.mk, [2] blagoj.ristevski@uklo.edu.mk,
[3] natasa.tabakovska@uklo.edu.mk, [4] ilija.jolevski@uklo.edu.mk

Abstract. Nowadays we are facing with data flood in many areas. Big data come from numerous sources such as human activities, measuring instruments and many appliances connected to computers or smart phones. One of the most challenging topics in the next decade will be how combination of genome and exposome data will contribute to reveal the risks for particular diseases. According to the medical scientists, the exposome includes all exposure environmental factors, from chemical and nonchemical agents to socio-behavioral and psychological factors as stress, diet, endogenous and exogenous factors from whole lifespan. The growing of mobile and ubiquitous computing technologies contributes in increasing the number of records regarding personal health and habits of patients. Internet of Things (IoT) includes the development of wearable measurement sensors connected with Bluetooth, which are capable to capture and store health-related data, intended to be stored in patient health records. The exposome is a healthcare and medicine concept that implies an interdisciplinary and integrated approach of many sciences domains including epidemiology, computing, environment sciences, toxicology and social science. We aim to integrate the data collected from various sensors and detectors in the patient health record to provide clinicians with more elements for better disease prognosis, diagnosis and treatment.

Keywords: Exposome, Omics Data, Personal Health Records, Health Informatics, Internet of Things.

1 Introduction

One of the emerging areas where big data can give great results are healthcare and medicine that use computer science as a tool for gaining emerging insights and analysis from a wealth of collected personal health data, medical and various omics data. With a concept of precision medicine and creating digital records for healthcare in patients' everyday life, healthcare organizations have adopted computer science and information systems as tools to collect and analyze data [14]. With technology advances and human genome sequencing, they have started to use more and more data that are stored in huge repositories in many places and in various formats, using a wide range of applications, medical and measuring instruments.

When Christopher P. Wilde explained that genome theory does not elucidate many chronicle diseases and proposed using of environmental complement to the genomics data, it was just a beginning of gaining different sets of omics data. These data sets were defined as exposome, as whole of exposures throughout human lifespan [23]. They had to be used with genome data to determine risk of diseases. Because data collected in all these manners have different formats, types, attributes, usually these data are analyzed by considering 6 V's big data characteristics: volume, velocity, variety, value, variability and veracity [17]. Hence, they demand attention of scientists from wide ranges of disciplines with purpose to be analyzed in a suitable manner to gain useful insights.

One of the biggest challenges in the next decade will be how combination of genome and exposome data (as a whole of human exposures from birth to death) will contribute to reveal the risk factors for particular disease. According to the medical scientists, the exposome includes exposure environmental factors from chemical and non-chemical agents, socio-behavioral and psychological factors such as stress, diet, endogenous and exogenous factors from whole lifespan [4]. Christopher P. Wild [23] defined three scopes of exposome: general (as social capital, education), internal (as metabolism, gene expression) and specific external (as chemical, noise).

The rest of the paper is organized as follow. In Section 2 we review some aspects related to PHR (personal health record) and EHR (electronic health record). In Section 3 we describe the concept of exposome. In Section 4 we highlighted Internet of Things (IoT) concept used as sensors for feature measurement as exposome. In Section 5 we explained the proposed relation between PHR and exposome. Finally, in Section 6 we provide concluding remarks.

2 Personal Health Records and Electronic Health Records

Several challenging areas today include relationships among healthcare and e-health, evidence based medicine, telemedicine and pervasive medicine [9]. We are aware that in the last decade population migration and peoples traveling increasing over the time. People as patients change their need for healthcare services toward using of electronic health services, which provides a medical care from everywhere. There is an arising need of medical facts for patients to help medical practitioners to make best decisions about diagnoses, treatments and patient care [1, 3], as evidence based medicine. Data can be obtained from many measuring instruments, prescribed documents and medical assessments. The best solution to support these new trends is collecting healthcare data from all patients' healthcare activities, medical examinations, laboratory results, prescription and referral, all stored in healthcare digital repositories in electronic format as EHR. Unfortunately, EHR as a repository for healthcare patients' data, are property of healthcare providers and usually are not accessible to the patients. The patients need to bring their own healthcare data with them when travelling abroad with possibility to gain a better medical assistance in the location where they are.

They have to have PHR saved and accessible to use them from anywhere and by patient-selected medical teams.

The concept of PHR nowadays is associated with web enabled technology intended to e-health and moreover it is a base for evidence-based personal healthcare. PHR enables many advantages such as improving the diagnosis efficiency, providing healthcare data management and real-time information management. For physicians, adopting PHR is very important because it provides access to the patient's medical records that might be used as a reference for diagnoses [5]. All stored data and medical information can play a positive role in medical practitioners' decision making for the patient healthcare.

The growth of using mobile and ubiquitous computing technologies contributes in increasing the number of records regarding personal health of patients. IoT includes the development of wearable measurement sensors connected with Bluetooth, which are capable to capture and store health-related data as PHR [18].

Techniques for processing data in healthcare include using of machine learning, pattern recognition, expert systems, statistics, applied mathematics and artificial intelligence [1]. But, there are many obstacles in managing such huge distributed databases with medical records because of their complexity, distribution and lack of interoperability.

Nowadays, the world is facing with global epidemics of cancer, influenza, AIDS, diabetes and obesity which demands increasing need of healthcare professionals. Patients need a more attention and education to cope with their situation. The situation also demands application of world-standards of PHR because of increased mobility of the world population and the need of evidence-based medicine that demand patients' data [18]. Since 2006, healthcare community works on making standards for PHR. For this reason, in 2012 ISO (International Organization for Standardization) introduced EHR-ISO/TR 14292 and applied the HL7 standard [12]. Although PHR and EHR differ regarding their contents, there is an increasing need of integration of these data into electronic medical records that contain all patients' health data through their live accurately, as shown on Fig.1. So, the emerging EU projects [19] intend to create a PHR whose owner will be the patient. This will be a starting point for using of PHR abroad and to give patient the possibility to have the access to their PHR stored on cloud environment [19].

When working with patient's data, data security and privacy issues have to be solved. PHR accepts data obtained from health related equipment such as accelerometers, gyroscopes, wireless devices, wristband and smart watches as IoT related devices. Data collected from these sensors can be saved in PHR and protected according to the national data privacy law, to help in patient risk detection.

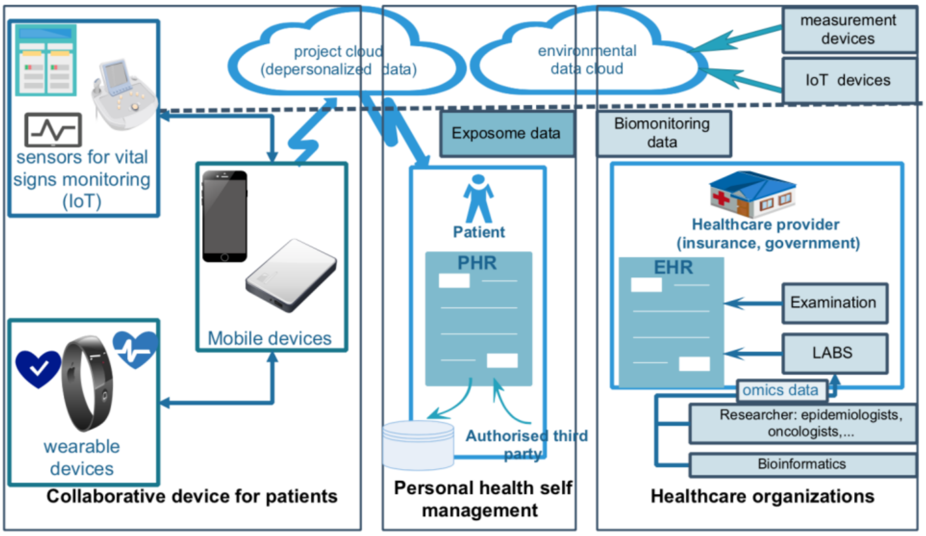

Fig. 1. Toward integration of EHR and PHR with IoT sensor's data as exposome.

3 Exposome theory

The concept exposome includes all traditional measurements in healthcare, known as bio-monitoring methods, as well as new types of measurements using sensors. The IoT concept includes a variety of wearable sensors connected to the computers or smart phones [13]. There is no limit for development sensors as measuring instruments that provide data such as spatio-temporal data with previous settings of measurement units, default values and the value ranges.

Exposome is a broader concept in healthcare and medicine. It implies interdisciplinary and integrated approaches from different domains such as epidemiology, computing, analytics, environment sciences, toxicology and social science. Exposome theory corresponds to the totality of exposure over the whole lifetime, and it is connected with usage of sensors and biological biomarkers measurements using high throughput omics technologies in order to create a prediction of effects of these exposures on human health conditions, worldwide and for the individual whole lifespan. Some epidemiologists state that exposome data have to be analyzed and connected with particular groups of patients who are exposed to factors that can have an influence on their health conditions that cannot be caused by genetics alone [23]. Precision medicine is focused on adopting treatments based on the genetic profile of tumors including biomarker but also exposome data. Large-scale studies contribute to gene environment interactions and propose epigenetics theory. They fully understand that there are complex interactions between genetics and environmental exposures that can be used to clarify the etiology of many diseases.

Nowadays, the epidemiologists remark oncometabolitics as exposome factors. Some research is dedicated to genome-wide association studies or environment wide association studies [8]. They connect genetic polymorphism with human diseases or a phenotype by defining in this way the endophenotype of each individual. To clarify the exposome-genome paradigm there are a lot of data that have to be considered and it is just the beginning of this research field [2]. Researchers have to consider matrix selection, variability of exposure and targeted and untargeted analysis to create more reliable data for exposures as exposome data [4]. Also, they have to use bioinformatics techniques and omics technologies that have the potential to discover unknown functional relationships among genes and diseases. Many initiatives provide accurate and reproducible laboratory measurements and particular groups of scientists gathered at the National Institute of Environmental Health Science (NIEHS) are dedicated these data to be stored in specific accessible databases such as METLIN (Metabolite and Tandem MS Database) and HMDB (Human Metabolome Database) [21]. The number of these omics data increases over time. The databases and repositories with human exposome data are good starting basis for providing data to the environmental health scientists as well as to the researchers that study the influence of whole exposure factors on the human genome.

Taking into account that many traditional biomonitoring methods as biological targeted analysis are available in healthcare providers' databases, created data can be used as a key exposome data for the patients. Also, data from epidemiologic studies are stored in digital repositories and they can be considered in patients' healthcare dossier, as EHR or PHR [18]. Some studies can provide a wide range of chemical exposure data for selected location in particular temporal points [4]. Some data of patient's blood and urine analysis for the presence of particular chemicals after short-time exposures can also be considered. But exposome approaches differ from these traditional biomonitoring methods because they include whole significant potential exposures, no matter they are exogenous (as pollutants, noise, UV rays, diets, drugs) or endogenous (as hormones, human or microbe's metabolites) [16]. Christopher P. Wild stated that incorporation of exposome paradigm in the traditional biomonitoring methods can improve in many ways exposure assessment [24]. The exposome theory also predicts that all measurements currently available and related to the environmental factors that have influence on human health will be considers in PHR. Also data from traditional bio-monitoring methods from medical labs, enriched with measurements of individual patient's parameters with IoT-connected devices will be available for further analysis. All these data can be used to characterize the individual behavior and improve the understanding of exposure profile on individual patient level.

Exposome studies in Europe as HELIX, HEALS and EXPOsOMICS already started to capture this type of information as CORDIS (Community Research and Development Information Service) [22]. Also, some universities founded centers for human exposome research and developed infrastructure that uses hybrid and biomonitoring approach combined for data collection. Also, they combine their techniques with untargeted omics analysis such as metabolomics, proteomics, transcriptomics and epigenomics data that can significantly help to uncover biologically meaningful exposome influence on the human genome. Metabolome and redox proteome data can also provide useful information to improve the composition of nutrients in individuals' diet on the basis of the features of the environment in which individual live [6].

4 IoT Technology and Exposome Data

As smartphones technologies become more accessible and widespread, low cost portable sensors become a “standard” for short or long-term tele-monitoring, software application development and microcontrollers. These applications provide wireless data access, usually in the cloud, where huge amounts of personal data are stored. It is a big technology revolution, because they provide many types of data, such as data for location, time, activities, habits and behaviors named as exposome data. Location can be very important factor for assessment of duration of exposure to a hazard, social environment, climate and other general external exposures, and even the microbiomics [11]. It is very important to define criteria for using sensors to collect data for further analysis. When using smartphones, data sets can be location, time, circumstances in which a person is, activities, whether a person is indoor or outdoor, its moveable positions etc. It is easy because smart phones usually have additional sensors and embedded algorithms which can help to determine whether a person is indoor/outdoor, whether is running, what are weather conditions, such as temperature, relative humidity, lightness or noise [10]. They usually have I/O (input-output) detectors, magnetism sensors and various others sensors with predefined accuracy [15]. Location can be related to the exposition to air pollutants. These data can greatly help to prediction for personal exposome. Physical activity is a very important risk factor for most diseases, and influences the environmental exposure. In that case inhalation and metabolism can be measured as energy spending through breathing rate and inhalation of the pollutants. There are many smart phones’ sensors for tracking physical activities as tri-axial accelerometers that can be used with gyroscopes. The steps also can be counted as well as metabolic equivalent. The heart rates, RR interval, which is small changes in the intervals between successive heartbeats, breathing rate, acceleration, activity level, passed distance etc. can be measured. For instance, absence of physical activities can result with obesity and diabetes.

5 Relationships between PHR and Exposome

The motivation for using software applications relied on sensor data is usually to provide information insight for healthier lifestyle or to provide health alerts. There are accessible fitness trackers and applications for diet control. These applications usually work as questionnaires that have to be filled after meals or they track dietary intake for weight loss or fitness purposes. Usually, they count burnt energy and metabolism rate for food intake [13].

Environmental and climate indoor conditions are usually measured with indoor monitoring sensors that allow access to their home measurements data remotely. Indoor and outdoor air quality can be measured, as well as temperature, humidity, mould, ventilation rate, chemical reactions producing secondary pollutants as formaldehydes and CO_2. Air quality can be measured with air sensors such as gas sensors that calculate gas concentration for metal oxide or electrochemical sensors (CO, O_3, NO_2, ammonia, hydrogen sulfide etc.) as total volatile organic compounds (TVOCs).

Humidity can be measured as a result of sensitivity of metal oxide sensors and their influence on the temperature rise rate [13]. Laser in higher-end sensors can obtain better results because of using optical methods and photodiode and lasers for indoor monitoring of very small airborne particles. It is very effective for cigarette smokes concentration detection [13]. Other sensors are specific sensors that measure micro and nano-particles. Measurement of these pollutants such as PM10 and PM2.5 are connected with different diseases, as chronic obstructive pulmonary disease (COPD).

Another type of smartphone sensor is the noise sensor that is intended to hand-held sound measurement of noise on fixed locations and noise dosimeters as personal exposure of noise. There are many devices that are not standardized and calibrated, but they are not used for working environment as well as for environmental noise measurement. Few smartphones already have the possibility for personal noise measurement calibration. The influence of noise can be very destructive for human health, especially for human hearing system as well as for some sleeping disorders. It is especially influential if its exposure is long-time and intensive noisily. Hearing system's damages can be predicted according to data gained from the measurement of the noise intensity and its duration.

Sensors for movement measurement can also measure human trajectory that can be associated with location as well as level of hazard such as air pollution, noise and exposure to UV rays. On the other side, food contamination can be connected with a level of soil contamination with pesticides and heavy metals. There are also sensors for stress detection. Exposure to radiation can be associated with appropriate measurements of environmental factors.

All these data should be stored in cloud databases as IoT data and connected with the patient's PHR as individual or as group, if the data is for outdoor sources. When all data points are available, the whole risk factors for the human health can be considered in more details. The whole gathering of exposome data as different exposure variables of one person or a group, will enable creation of more accurate machine learning models that will be able to find connections among seemingly not related exposures and actions of a person and his or her health's risks and diseases, and to suggest the adoption of corrective actions to avoid or mitigate those risks.

One case of integration of IoT data, collected from wearable sensors, and PHR is the IPA2 project Cross4all (Cross-border initiative for integrated health and social services promoting safe ageing, early prevention and independent living for all). Sensors are connected with mobile devices via Bluetooth as part of project equipment and they send data on the Cross4all cloud. The collected data have to be considered when medical practitioners have to decide about patient health condition in cross border area. Also, many traditional bio-monitoring data gained from traditional medical equipment from public and private healthcare provider databases (from patient EHR) are collected. Data gained from mobile wearable sensors that will be connected to patients for short tele-monitoring period, can be very useful to medical specialists to make decision for patient health in the era of evidence-based medicine. The project aims to create a database with PHR data for project's participants as well as to collect a large amount of data connected with human health as exposome data that should be integrated with PHR data. Connection with some environmental database that store data for air pollutants, outdoor weather and climate conditions, the presence of nano-particles, as environmental exposome data, connected with patient's location, can also provide very important data to the medical practitioners for particular types of patients with chronicle diseases. Tracking all these data can provide a better risk assessment for patient health.

6 Conclusions

Smartphones with IoT sensors have many potential data security and privacy issues. Indeed, one of the main problems is "data ownership". To resolve this issue, anonymized data can be a solution. So, there are many challenges when data are used for evidence-based medicine, as data standardization, harmonization, data privacy and confidentiality. Many procedures have to be created to support the wider use of personal sensors, as a part of e-health and IoT concepts. When they are connected with PHR, the problem is more complex because of the quantity of complex data that have to be secured and used for medical purposes only.

Another important issue is data collection. Collection, storage, processing and analysis of this kind of data demand big efforts because of the complexity, volume, variety, velocity, values, variability, and veracity of the data. Also, this type of data is very sensitive because of the possibility of abuse, misuse as well as because they contain human behavior recognition factors [2]. Another issue is how these sensors' data have to be used to highlight the influence of environmental factors to the human health and how these exposome data can be used in practice. There are many ongoing exposome projects that should answer how to balance the possibilities of new sensor technologies and more conventional and well-tested methods of bio-monitoring data collection. Exposome tools can also enable the measurement of many environmental parameters and allow improved quantification of individual and population risk factors [20]. Data analysis can also answer these questions about exposome and point out some aspects of their usage. Some semantic approaches also can help in gaining information and clinical context of evaluation and usage of exposome data [7].

However, there is a need of integration of the data collected from various environmental sensors and detectors in PHR to provide better disease prognosis, diagnosis and treatment, which we aim to achieve with the Cross4all project.

Acknowledgments

Part of the work presented in this paper has been carried out in the framework of the project "Cross-border initiative for integrated health and social services promoting safe ageing, early prevention and independent living for all (Cross4all)", which is implemented in the context of the INTERREG IPA Cross Border Cooperation Programme CCI 2014 TC 16 I5CB 009 and co-funded by the European Union and national funds of the participating countries.

References

1. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. IEEE J Biomed Health Inform 2015 Jul;19(4):1193-1208. [doi: 10.1109/JBHI.2015.2450362] [Medline: 26173222]
2. Barouki R. and all, Integration of human exposome with the human genome to advance medicine, Biochimie 152 (2018) pp.155-158
3. Dagnano S. and Macherone A., Editors, Unraveling Exposome, © Springer international publishing, 2019; Smith M.T. and all, Chapter 1: Using Exposomics to Assess Cumulative Risks from Multiple Environmental Stressors, pp.3-22
4. Dennis K.K. and all, Biomonitoring in the Era of the Exposome, Environmental Health Perspective, Volume 125, number 4, April 2017
5. Gay K. and all, Electronic Health Record Error Prevention Approach Using Ontology in Big Data, 2015 IEEE 17th International Conference on High Performance Computing and Communications (HPCC), 2015 IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), and 2015 IEEE 12th International Conf on Embedded Software and Systems (ICESS), DOI: DOI 10.1109/HPCC-CSS-ICESS.2015.168
6. Go Y.-M. and Jones D.P., Redox biology: Interface of the exposome with the proteome, epigenome and genome, Redoc Biology 2 (2014) pp.358-360
7. Fan J. and all, Semantic Modeling for Exposomics with Exploratory Evaluation in Clinical Context, Journal of Healthcare Engineering Volume 2017, Article ID 3818302, 10 pages, https://doi.org/10.1155/2017/3818302
8. Fernandez J. and all, Selenium at the redox interface of the genome, metabolome and exposome, Free Radical Biology and Medicine 127 (2018) pp.215-227
9. Househ and all, Big Data, Big Challenges: A Healthcare Perspective Background, Issues, Solutions and Research Directions, Fernando Martin-Sanchez, Big Data Challenges from an Integrative Exposome/Expotype Perspective, Springer 2019, ISSN 2195-271X, ISSN 2195-2728 (electronic)
10. IOM-Institute of Medicine (2014) Capturing social and behavioral domains and measures in electronic health records: phase 2. The National Academies Press, Washington; doi.org/10.17226/18951
11. ISO. Health informatics — Personal health records — Definition, scope and context. Switzerland: ISO; 2012. URL: https:/ /www.iso.org/obp/ui/#iso:std:iso:tr:14292:ed-1:v1:en[accessed 2019-05-10]
12. Lindlay J. and all; Why the Internet of Things needs Object Oriented Ontology, Design for Next, 12th EAD Conference, Sapienza University of Rome, 12-14 April 2017
13. Loh M. and all, How Sensors Might Help Define the External Exposome, International Journal of Environmental Research and Public Health, 2017, 14, 434; doi:10.3390/ijerphi14040434
14. Lopez-Campos G. and all, Biomedical Informatics and the Digital Component of the Exposeome, MEDINFO 2017, doi: 10.3233/978-1-61499-830-3-496
15. Morgenthaler J. Moving Toward an Open Standard, Universal Health Record. Smart-publications. 2007. URL: http://www. smart-publications.com/articles/moving-toward-an-open-standard-universal-health-record, Accesses 10.5.2019
16. Rappaport S.M. and Smith M.T., Environment and disease risk, Science 330 (2010) 460-461, doi: 10.1126/science.1192603.
17. Ristevski B, Chen M. Big data analytics in medicine and healthcare. Journal of integrative bioinformatics. 2018 May 10;15(3).

18. Roehrs A. and all, Personal Health Records: A Systematic Literature Review, Journal of Medical Internet Research, 2017;19(1):e13, doi:10.2196/jmir.5876
19. Savoska S., Jolevski I., Architectural Model of e-health System for Support the Integrated Cross-border Services, In proceedings of ISGT 2018 Conference, November, Sofia, 2018.
20. Smith M.T. and all, Using Exposomics to Assess Cumulative Risks from Multiple Environmental Stressors, *Dagnino S. and Macherone A. (editors), Unraveling the Exposome, a practical view,* 2018, Springer, pp.3-22, https://doi.org/10.1007/978-3-319-89321-1.
21. Tautenhahn R. and all., An Accelerated workflow for untargeted metabolomics using the METLIN database. NatBiotechnol 30:826-828, doi: 10.1038/nbt2348.
22. Vineis, P., et al., The exposome in practice: Design of the EXPOsOMICS project. Int. J. Hyg. Environ. Health (2016), http://dx.doi.org/10.1016/j.ijheh.2016.08.001
23. Wild C.P., Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology, Cancer Epidemiol Biomarkers Prev. 14 (2005) 1847-1850.
24. Wild C.P., The exposome: from concept to utility, Int.J.Epidemiol.41 (2012) 23-32