INNOVATIONS JOURNALS

Research Article

Increasing the Value of Real-World Crowdsourcing Health Data with e-MetaBio, a Novel Patient-Centric IT Infrastructure

Desislava Ivanova,[1] **Panagiotis Katsaounis**,[1] **Konstantinos Votis**[1]

[1]Metabio, Thessaloniki, Greece

Address correspondence to Panagiotis Katsaounis (katsaounis-ceo@metab.io).

Sources of support: None. Conflict of Interest: The authors are employed by Metabio, and the described technology is created and implemented for commercial purposes.

Submitted: Nov 14, 2023; First Revision Received: Jan 22, 2024; Accepted: Jan 25, 2024

Ivanova D, Katsaounis P, Votis K. Increasing the value of real-world crowdsourcing health data with e-MetaBio, a novel patient-centric IT infrastructure. *Innov Dig Health Diagn Bio*. 2024; 4:15–24. DOI: 10.36401/IDDB-23-14.

ABSTRACT

Introduction: Digital health and evolutionary medicine create new insights of mediation and health treatment plan support, introducing crowdsourcing and patients' real-world data records, so as to promote the development of high-quality healthcare accessible to everyone. Within the scope of its activities Metabio's team has developed an interoperable unified method and technology for crowd-generated databases, creating a user-friendly platform for data collection, processing, and distribution among stakeholders within the global healthcare system in real time. **Methods:** In this paper we describe standard methodologies, requirements, issues, and challenges for the design and deployment of an advanced IT infrastructure for longitudinal structured patient-related data records, based on a patient-centric model of operation, as well as the difficulties for the development of disease-specific user-prefixed interface for real-world data collection. **Results:** Through a dynamic real-time (DRT) e-consent module and digital rights management protocols, the overall platform enables patients to monitor and manage their disease-related conditions, as well as for healthcare providers and/or research entities to have access to valuable biomedical patient data, not recorded so far. **Conclusion:** The project introduces novel perspectives for future evidence-based practices, promoting research and development and improving current healthcare systems, by using crowd-generated data sources that bring a much higher degree of accuracy and value for the entire healthcare system.

Keywords: crowd-generated data, innovation technology, interoperability, data harmonization, e-consent, IT infrastructure, biotechnology, biobanking, biosample

INTRODUCTION

Globalization is a term used often in different fields to describe the elimination of geo-political boundaries and the creation of international collaboration among countries, for the common good.[1] Within the scope of healthcare-related activities around the world, many countries have set up new principles and disciplines to promote health needs, establishing new patterns of actions and approaches of operational systems and involving diversity of stakeholders.[2] Entities and patients/donors, as components in the global health system, create a huge network that consists of transnational actors aiming to improve healthcare and the polylateral arrangements that establish the requirements within which these actors operate, considering governance, economics, and final delivery.[3] Both practical and managerial implications should be adapted in the operational healthcare system, focusing on real-time data monitoring and management.[4]

Health security and quality overcome several issues and challenges to achieve the goal of progress in the era of globalization, including thematic areas such as strategic planning and policies formulation, intelligence, regulation requirements, coordination and collaboration among entities, accountability and managing of health resources, and final decision making per patient, leading all to the need of robust data collection across the overall healthcare system.[5] According to the World Health Organization (WHO), among the urgent health challenges for the upcoming decade is the capitalization on technological advancements.[6]

For the purpose of ensuring the quality of health over years, evolutionary medicine has been established, creating new clinical insights and pathways of mediation for future fast treatment approaches, under several circumstances of specific diseases and predictable outcomes.[7] High-throughput genomic technologies and evolutionary history of humans, including data records, such as

family history and genetic factors of significance, have led to the mapping and creation of disease-specific networks that have revealed new insights in clinical research and modern medicine management for unique human diseases.[8]

Bioinformatic tools, analytics, and the ability to link genomic information to phenotype, not only promote advanced research and health improvement, but also create large multidimensional data cohorts that, except for the many aspects that it encompasses, creates many biases, as well.[9] Currently, healthcare systems are forced to adapt all these changes and new perspectives, dealing with big data technologies and multisourcing electronic medical records and other patient-related data from ex situ entities, leading to the need for novel patient-centric learning paradigm and artificial intelligence (AI) technologies for data integration, processing, analysis and final dissemination.[10]

The effectiveness of data final dissemination is not always possible, considering the very important fact of significance, the "translational gap," occurring due to diverse disciplines and practices among the source and the final receiver, increasing the risk for bias.[11] The impact of such bias is more visible in rare diseases and chronic conditions, being one of the biggest challenges in healthcare universally, due to their low prevalence.[12] This progressive burden within global public health creates a huge gap, having both social and economic impact, based on the fact that very few reliable data have been available for basic research and treatment plan set up.[13] Despite their low prevalence, most of them are chronic conditions that decrease patients' quality of life, because of the lack of treatment approach, but also patients' psycho-social well-being, as many symptoms, including anxiety, depression, fear, and isolation, are reported to influence the course of their medical condition.[14]

On the other hand, the healthcare system must manage the ad hoc models created under specific circumstances, such as the COVID-19 pandemic, within their organizational system, including medical-research team conduction and random sample collection for research, that not only give rise to communication and dissemination issues among research participants, but also create cumulative biosample and un-linked data and meta-data collections that are not valuable for basic research and subsequently for novel treatment approaches.[15–17]

Considering all the above, in a multidimensional and multisourcing environment for data collection on patients and biosamples, interoperability and harmonization within the whole network should be the basis for the development of federated international unified research infrastructure for all participating stakeholders.[18] Semantic modeling for interoperability and exchange information among entities based on international data sharing have been used for proper data interpretation, using common standard terminologies, vocabularies and ontologies. Still these models should be updated by linking the harmonized data to their senses and adding semantic value to categorize and use the disseminated data, so as to answer specific research questions of significance.[19,20]

Within the global healthcare system, the most important players are the patients themselves, carrying valuable information about their unique medical conditions; however, many difficulties in patients' engagement are reported, based on self-related, staff-related, system-related and community-related barriers.[21] The practical implementation of patients' engagement, based on major obstacles, such as resource and subject limitation for research coordinators, on the one hand, and participant knowledge and comprehension about research-related activities that might be of benefit for their condition, on the other hand, create the current selection bias for specific trials and increase the gap in clinical practice.[22] Despite that, legal and ethical issues arise around the protection of data privacy and patients' rights that might lead the crowd participation to failure.[23]

The future sustainability of a high-quality universal healthcare system, providing advanced healthcare services, seems to be based for the next years, on developed software services and platforms for data collection, processing, and procurement, accelerating the initial pharma research and development, by enrolling patients in a patient-centric model of healthcare approach.[24] Citizen science and crowdsourcing environments, by voluntary patient enrollment in the scientific process, gathering real-world data for evidence-based strategies and treatment management approaches, seems to be of huge potential within the crisis of the COVID-19 pandemic and its outcomes management.[25] The sudden pandemic outburst revealed important already existing friction points within the global healthcare system, suggesting big steps and changes into the future, including novel "smart" technologies, analysis of big data in real time, and patient enrollment in the healthcare system.[26]

New technologies drive fundamental changes not only in patient-centric treatment strategies and personalized medicine, but also in pharmaceuticals, decreasing the time and cost for basic research and introducing an advanced business model, based on digitalization, as well as patients themselves providing methods and technologies for self-medication support and disease management.[27] Within the vision of the global strategies for future development of the healthcare system, based on WHO insights and planning procedures, the acceleration of high-quality healthcare accessible for everyone is one of the most promising values that digital health and new technologies for crowdsourcing and big data management can add to the global network.[28]

Although promising, digital transformation is not a simple procedure, in fact it is reported that more than 70% of the projects addressing the adoption of novel

technologies into the healthcare system failed to meet their final goals.[29] Despite that, patients' engagement in the development of the digital environment specifically designed to collect, process and disseminate data is even more challenging, not only to develop and apply standard policies that cover national and international initiatives, but also to intrigue patients and meet their perspectives, in order to empower their participation in the whole process.[30,31]

MATERIALS AND METHODS

Overview

Within the scope of its action, Metabio has developed a crowdsourcing platform for real-world data (RWD) collection, creating evidence-based patients' pushed data records that can be stored, analyzed and disseminated within the implicated entities in real time. Prior research for already existing crowdsourcing platforms has shown two main weaknesses in current systems: (1) the lack of streamlined data records from multiple sources per patient and (2) the lack of linkage of data delivered from crowdsourcing platforms to the electronic medical record of a patient, in association with analytical data from potential biosample donation. The scope of this project was to create a user-friendly environment for patients' enrollment in the healthcare system, as well as to raise patient awareness and support toward biospecimen donation and involvement in research-related activities. We aimed to develop an advanced disease-specific platform that not only collects outsourcing patient-related data that could be monitored and managed from their healthcare providers and create disease-specific patient profiles, but also enrolls patients in their disease management and gives them total control over their donations and related data.

Project Design and Development

The first steps in this project were to establish the general framework of activities and requirements for overall architecture development, as well as set all the desirable data types that will be recorded from the patient's interface. Moreover, the project plan includes the design of operating systems, appropriate software and digital technologies implementation mode for overall platform development and maintenance, including data integration method and structure. Roles and tasks were distributed between the biomedical and technological team, creating an interactive working group, consisting of experts in molecular biology, genetics, healthcare management, information technology, engineering, and data protection. Investigation and extensive search in databases and prior crowdsourcing platforms contributed to the final platform model design, taking into account all the potential phases of significance from initial design to final implementation and use, as described in the following sections.

Strategies and Modality Requirements

Based on previous extensive research on disease-specific, healthcare-related, and patient-related databases and platforms, our team decided to create an innovative crowdsourcing system for efficient medical research and health self-management improvement, by recording data delivered from patients themselves, not recorded anywhere before, and integrating them within their biosample usage chain, so as to create multidimensional data records per patient, as a complement to their electronic medical records. For that purpose, huge amounts of relative data records were collected, evaluated, categorized, and integrated in implementable disease-specific workflows, prioritizing chronic neurodegenerative conditions and rare diseases that have no cure and meet the credentials to increase the burden within the current and future healthcare system. The overall procedure of research, data collection, and linkage with standard reference ranges, appropriate units and options took several months, so as to create prefixed ready-to use lists per category for the final users, meeting all desirable standards.

The main purpose and modality were to create a patient-friendly platform that would combine effectiveness with decreased time and effort cost for the final user. On the other hand, we aimed to record as much as possible valuable data per category, increasing the variability of information and data sets, essential for a patient's profile construction and future utilization in research. Thus, planning and recording of the initial phase were based on data sets and information within a user-friendly mode and standard modular constructions, used in Google and Amazon.

Standardization and Interoperability

Multisourcing data sets had to be available and useful for all implicated entities on international levels, promoting the interoperability and harmonization within the entire data chain. For this purpose, we gathered and categorized the potential standard ontologies used universally for clinical data collection, selecting and combining the more commonly used, including HL7, openEHR, LOINC, SNOMED, ICD-11, etc. In addition, the platform should be able to link and integrate these data records into different electronic health record (EHR) systems and health information technologies (HIT). For that purpose, the management system was designed to be based on Fast Healthcare Interoperability Resources (FHIR) specifications, which allows interoperability and integration of health data, readily available via RESTful APIs in the cloud. The process of the final crowdsourcing platform had to be able to collect data from patients in real time and establish a chain of characterized, relational timeline data per donor.

Security and Compliance

One of the most important issues and challenges during the project design and development was to set all the proper conditions, ensuring that the entire system will

meet all the requirements for data protection and accessibility, based on national and international standards. The security infrastructure has been carefully designed to provide the different types of end users with a unique web portal and the ability of single-entry point to the data collected, setting a high level of encryption for all data sets. All procedures and data access had to be based on strong user authentication credentials, so as to control not only access, but also analysis, storage, and dissemination of data. Methods and approaches for security assurance have been estimated for both data protection and software development, following strict guidance and methodologies according to the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) regulations.[32,33] The overall system was designed from the beginning to prohibit direct access to, or copying or exporting personal and unidentified/de-identified data. Among the methodologies followed were anonymization and/or tokenization of servers, pinning methodologies and integration of encrypted protocols.

Data Management and Transfer

As a patient-centric environment, the new platform had to give patients total control over their donation and associated data, by giving them the opportunity to monitor and manage the overall usage chain, from donation to final shipment and future utilization. Based on that concept, the platform was structured using improved protocols for the management of digital rights (digital rights management [DRM]) and consent management technologies. More specifically, we aimed to give the patients the opportunity to create their own consent status and change it anytime. On the other hand, as the data would be linked to their electronic medical records and biosamples-related data, all implicated entities should have the ability to be informed of any changes in real time. Considering the preceding, we have created a dynamic real-time (DRT) e-consent module, enabling the system with patients' personal preference over their donation and associated data, including the ability for real-time authorization to all stakeholders. From the beginning, the back-end of the system was designed based on a combination of role-based access control (RBAC) and attribute-based access control (ABAC) to implement the logic of electronic consent and the rights of users to read and modify any of the data and meta-data related to a target individual. The basic idea was to create a novel methodology and system that would not only collect RWD from another crowdsourcing platform but would also reinforce the development of a valuable chain of data sets, incorporating crowdsourcing data in patient-associated data records, both from healthcare providers and research-related entities, including biobanks. All the involved methodologies had to be able to maintain patients' rights and the ability of anonymization, at the same time as to disseminate these data among the entire network of involved entities, providing increased value to crowdsourcing data and promoting research and development in healthcare and personalized medicine.

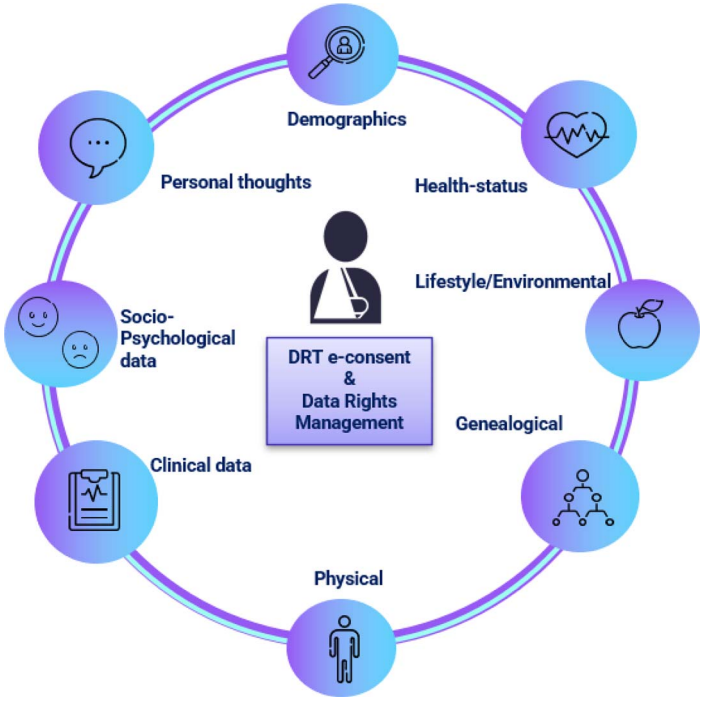

Figure 1. Patients' platform for real-world data records based on digital rights management and consent management technologies. DRT e-consent: Dynamic Real-time Tiered e-consent.

RESULTS

Overview of the Entire System and Architecture

The new platform for patients (Fig. 1) provides an electronic system for recording RWD, associated with the development of a patient's disorder, and collects the data in real time from patients' entries and places them in data collection sets in a longitudinal way (along the time axis). Thus, the system is designed not only to connect the series of data that were created, but also to restore the connection with the different data sets collected in chronological order, without disturbing the anonymity status of the donor.

The platform is part of an umbrella system, a web platform for multisource longitudinal biomedical data collection, storage, and distribution within the biosample usage chain, providing streamlined services for biobanks, hospitals, research entities, and industry in general (Figure 2). The system acts as an umbrella to all integrated platforms for data collection from distinct sources, including pre-acquisition data from healthcare facilities (EMRs, EHRs), research data assigned to biospecimen cohorts, and biospecimen management and analyses data from biobanks. The overall architecture of the entire system is shown in Figure 3 and has the potential to

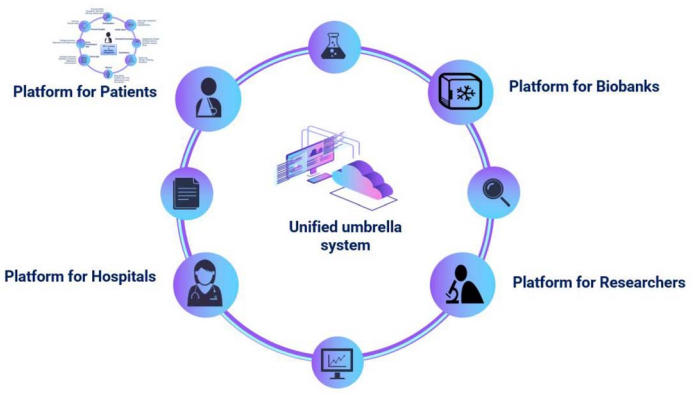

Figure 2. Unified umbrella system for multisource data collection, processing, and dissemination in real time within the biosample usage chain.

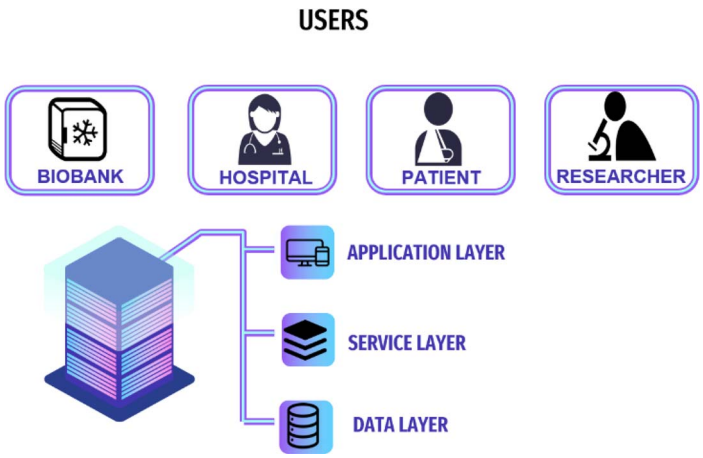

Figure 3. Information technology architecture of the entire system.

expand into other areas and health services. As part of the unified umbrella system, the innovative crowdsourcing platform has been integrated into the system architecture, providing an additional source for data collection. More specifically, the platform complements the entire system, by patient-pushed health status data collections in real time, through an advanced consent management system.

The system is separated into four main layers:

Data Layer – includes the technical infrastructure that allows the proper definition, storage, and exchange of data, following ontologies and standards for harmonization and interoperability. Various data models, databases, and servers needed have been developed in this layer.

Service Layer – includes the main services provided by the platform, such as the collection and maintenance of patient-generated disease-specific information and big data integration.

Application Layer – includes the applications targeting the end-user roles.

Security and Privacy Layer – ensures security and privacy in all the preceding architectural layers. In addition, this layer accredits the secure exchange of information between the system and existing/legacy systems used by the associated biobanks and hospitals. This layer will also ensure compliance of the platform with local, national, and international legislation and ethics. The Security and Privacy Layer is responsible for handling (1) the authentication and encryption to the stored/transmitted data, (2) the permissions of each end-user for accessing/modifying the data stored in the Data Layer, (3) the data sharing agreements among the different actors, and (4) patient consent status via the e-consent module. JSON Web Token (JWT)-based authentication for all server requests is implemented.

The entire chain of data collection, analysis, storage, and distribution, including patients/donors, hospitals, biobanks, and research infrastructures, is operated and managed through DRM technologies, authorizing the different types of stakeholders and users on the utilization of biosamples and related data, according to the consent that patients have been provided.

Real-Time Dynamic E-Consent and Custodianship

The platform allows the only and main custodian of the biosamples as well as the related data (i.e., the donor) to form, according to his or her preferences, the usage chain described previously, the relationships between the different users and the level of access to both the biosamples and data, and the information he or she wishes to receive from their use, the type of research applied by users upon their donated biosamples, in real time and in a dynamic way for the entire usage chain. The platform provides the donors/patients, with the appropriate configuration of the access rights, the capability to control their medical records, and the biosample usage through the Dynamic Real-rime Tiered Electronic Consent form. The preference for the patient's consent status translates into the determination of access, supply, and management rights to data and biosamples in a network of healthcare providers, biobanks and research entities.

The different stakeholder requests, the correlations with patient data and biosamples, and their interactions are defined by this configuration. The system allows the patients to control flow, supply (i.e., the provision of the data and/or biosamples of the donor to one or more users), and define the management and research of their data and biosamples. The DRT e-consent system establishes a process, by which it manages the accessibility of data, biosamples, as well as the application of research thereon and the communication of research between patients and research institutes. The solution sets out the conditions for the general and specific obligations of hospitals, biobanks, and research institutes.

Patient User Interface/User Experience (UX)

The application layer provides the end users with a specifically designed user interface (UI), meeting the needs for each user case (i.e., patients-donors). Patients can be involved into the crowdsourcing system in two distinct ways:

Table 1. List of data type collections

Category	Types of Data
Demographics	Personal details, occupation, education, ethnicity, marital status
Health status	Vitals signs, symptoms, allergies, hospitalizations
Lifestyle/Environmental	Supplements, dietary, smoking, exercise, activities, calories, sleep
Genealogical	Family tree, siblings, offspring, hereditary
Physical	Body details, usability, pain level, menstruation cycle (for women)
Sociopsychological data	Feelings, emotions, depression level, mania level
Clinical data	Standard treatments, alternative treatments, medications, immunizations
Personal thoughts data	Free text, personal diary
e-Consent	Consent type, relationship status, access management, research, research outcome

1. *Case 1*: Entering a healthcare provider (HP)/hospital for caring or diagnostic purposes. In this case, patients are provided with a unique ID, username, and password. They should first sign and form their consent status for data and biospecimen donation. The alphanumeric ID entered to the HP database as well as diagnostic results tests and treatment information are automatically fed into the system. Once successfully registered, they are able to log in and explore the interface.
2. *Case 2*: Directly joining the community via the application. Patients are provided with a unique ID, username, and password. They log in to the system independently and without the involvement of an HP and record their data. Patients registering in this fashion use the e-consent form and do not provide any kind of biospecimen but only their data.

Based on the preceding, two patient types could be obtained: (1) Donor Patient and (2) Community Patient. All data recordings are prefixed options related to the patient's profile and the disease he or she is suffering from, and patients select from prefixed scroll-down option menus, within acceptable limits.

Table 1 presents the list of available modules for data collection types recorded from the interface. Patients' interface is an easy operated environment that allows patients to monitor and manage their own conditions and compare themselves with others. It is enabled with graphs and diagrams, giving the opportunity for patients to create their health status representation by entering their own data that can be viewed per day, per week, per month, and per year.

UI is designed based on Akveo NGX Admin templates that, among others, meets the best criteria for efficiency, performance, customization, update, and customer support, making this routinely based patient experience unique and easy to follow up.[34]

The interface records real-world, environmental, lifestyle, health-status, genealogical, bodily, socio-data, psychological, clinical, and treatment data. The platform implements and disseminates these data for research purposes and healthcare treatment monitoring through only a single system, setting multiple users' choice, selection criteria, related to all types of data associated with them, in real time for multiple units in different geographic locations. It also implements the control of the availability of biosamples and the data related thereto, in real time and following the consent criteria set by the donors. These data allow the creation of annotated patient profiles, creating disease-specific cohorts available for novel treatment approaches and decision making.

DISCUSSION

Digital health services that support self-management, control of well-being, and treatment, although a promising and evolutionary intervention in healthcare systems, lack sustainability, because they are influenced by both actors and aspects, including human factors, technical factors, and extrinsic/intrinsic healthcare ecosystem characteristics.[35] To simplify the research value chain and promote the interaction among stakeholders in novel research projects, many research-relative activities are currently structured in a manner of outsourcing and implicate a large group of people, through an open call, that provide valuable insights within the scope of investigation.[36] Citizen-generated information and crowdsourcing methodologies, although associated nowadays with digital technology evolution and being the basis for many projects, where multiple stakeholders are required, has been used several centuries ago, recruiting large number of volunteers to participate in a wide spectrum of activities (i.e., the development of "The Oxford English Dictionary").[37] Nowadays, crowdsourcing is related to innovation and development of advanced medical research, promoting an open source decentralized model of action and contributing to the development of collaborative frameworks for preclinical and clinical research, based on collective intelligence.[38] Crowdsourcing has been demonstrated as the most promising solution for the global crisis of drug development, integrating large-scale data sourcing, including patients themselves, and covering major areas of interest in current and future models of research.[39] Through crowdsourcing, individuals contribute to research and development, provide valuable information and insights for disease progression and outcome, support the evaluation of initial research hypotheses, and interact actively with healthcare professionals and researchers for their own benefit.[40]

Project design and development for crowdsourcing seems to be extremely difficult and although meeting the requirements, many projects fail to effectively cover all the

multidimensional aspects arising from such frameworks of action, considering the successful engagement of the target number of individuals, for example, or establishing the roles and rights of each contributor.[41] In addition, the implementation and data integration from crowdsourcing demand the development and deployment of technological support and systems for collection, analyses, and interpretation of diverse multisource data, so as to provide longitudinal patient care that lacks in current systems and structured data sets of scientific significance that could be used from multiple health-related sectors.[42]

Although many applications for crowdsourcing are available today, covering more frequently the areas of health promotion, health research, health care, and less frequently problem-solving surveillance or monitoring, much information associated with crowd workers' characteristics and/or logistics are missing from most of the reports.[43]

Our prior research in crowdsourcing systems for data sharing among hospitals, academia, industry, and patients has shown significant friction points in data federation and "open science" initiatives, most associated with strict privacy laws and legal issues, increasing the need for efficacy and accuracy in future solutions, as well as the need for establishment of global standards and interoperability.[44] One of the most popular crowdsourcing and real-world platforms with an 18-year-old experience that focuses on personalized networking within healthcare is "PatientsLikeMe."[45] Similarly, other platforms serving as a pool of patient data collections, using citizen science and providing access to a large pool of diverse medical information, have been developed during the years, including Humanscape, Wisdo, RxPx, PicnicHealth, carecircle, Mango Sciences, Alike, antidote, MyHealthTeam, and Savvy Cooperative.[46–55] These platforms, some of which use advanced technologies, serve as patient networks for health monitoring and allow pharmaceutical companies, healthcare providers, or researchers to use the data sets to improve the existing healthcare system and provide a specialized effective treatment plan per patient.

Based on that concept, the aim of this project was to design and develop not only a platform for static data collection and share among stakeholders, but a multidimensional platform, part of an unified system with common electronic ontology and system of operation, that is able to acquire, maintain, and transfer data among different sites in real time, connecting patients' RWD data to their EHRs from their associated healthcare providers and to biomedical/analytical data associated with donated biosamples, thus creating a whole profile of a given condition per patient, representing the entire cycle of the evolution of a given disorder. The platform is specifically designed to operate through a system using a holistic approach and method, enabling multiple systems to be combined into one, addressing the lack of interoperability and ontological inconsistencies in separate systems. In addition, patients are created from the hospital, representing a real-world patient cohort, validating the data records.

In this paper, we present the strategies and methodologies for the design and development of an interoperable system for longitudinal crowd-generated data sets from patients themselves, based on advanced consent management technologies that enable the collection, storage, and dissemination of these data among implicated entities in real time, following national and international regulations and standards. Through the phase of project planning, the platform should meet specific criteria, so as to overcome issues mentioned previously and fill the current gap in crowdsourcing and healthcare ecosystems in general. Challenges within the project implementation were met both during the construction of the back-end and the creation of users-friendly functional frontend, meeting the needs of the end-users. In Table 2 below the main specifications for both front-end and back-end are introduced, covering challenges in each step of the data collection, storage, final distribution and integration processes.

First, the IT architecture of the platform must be planned and predicted outcomes must be estimated from the beginning, to ensure that the platform would not only be functional for data collections, but would also enable the secure transactions within the network and integration of crowdsourced data into already existing systems and/or new ones, according to standard policies, as well as ethical and legal requirements. The compatibility with other systems and legacies, alongside with the ability to put data along the time axis, creating standardized unique collections per patient, demanded the implementation of specific methodologies and protocols during the deployment phase. On the other hand, the front-end should be customized for patients' needs, representing a real-world disease-specific environment with easy-to-use functionalities.

Succeeding the deployment of the entire platform, further testing for improvements and final revision were planned and executed, so as to release the first version of fully equipped and functional platform for crowd-based data collection. The main advantages for both patients and the entire community within the healthcare system are summarized in Table 3, taking into account the importance of crowdsourcing and serious gaps in current systems, described extensively in this article. The final platform could serve as an added value source for patients' data records, developing a real-time interactive environment for unique patient cohorts and data entries, available for each authorized research-related and/or healthcare-related entity that joins the network. These longitudinal data sets and cohorts could accomplish already existing clinical studies and research analyses, giving an insight from the real-world environment of the patients' well-being and support research results and hypothesis-driven research

Table 2. Requirements and challenges within the whole process of project design and deployment, including the final aspects for both back-end and front-end

Item	Specifications/Requirements
Back-end	
Security	Both for data and software protection Encryption of data elements Authentication and permission (access control, manage control, distribution control) Firewalls establishment and protection from data breach Backup and data retrieval
Operational processes and workflows	Implementation of protocols and standard methodologies for digital platforms (i.e., DRM) System control and maintenance set up Establishment of software development methodologies Creation of environment for the project development Login/Monitoring tools Set up deployment processes and pipelines
Consent and permission	RBAC and ABAC set up permissions Development of DRT e-consent module Strategies for consent status incorporation within operational systems of the involved entities
Compliance	GDPR and HIPAA standards guidelines and principles Ethical and legal policies analysis and incorporation
IT Architecture	Diagrams and layers of the entire system set up Services and functionalities of each layer Design, construction, testing, improvements Software building applications Data acquisition model
Front-end	
Data types	Extensive research in disease-specific databases Collection, categorization and linkage of disease/patient-related data per category Prefixed dropdown lists design and deployment Standard units and reference ranges set up
e-Consent	Customization of consent module Patients' guidance through the consent options
UI/UX	User-friendly environment Fit-for-the purpose options for data records Visualization of patient's records Customization of usability (i.e., themes, right-handed/left-handed)
Authentication	Unique credentials per user Options for password retrieval and / or account deletion

ABAC: Attribute-Based Access Control; DRM: Digital Rights Management; DRT e-consent: Dynamic Real-time Tiered e-consent; GDPR: General Data Protection Regulation; HIPAA: Health Insurance Portability and Accountability Act; IT: information technology; RBAC: Role-Based Access Control; UI: user interface; UX: user experience

activities. In addition, healthcare providers and practitioners have access to valuable clinical data and treatment outcomes, not recorded so far, that increase the ability to establish standards for future decision making and more effective personalized treatment approaches.

The innovation of the tool enables the patients to self-monitor and maintain not only their health-related conditions, but also the entire chain of their data usability, creating a patient-centric network within the healthcare system. The self-reported data from these patients on the progression of their disease help to better understand the variables of clinical cases, their etiology, and the unique progression of the disease of certain groups of people from different demographic areas with different genetic backgrounds and hereditary traits. The platform could boost the development of novel strategies to alleviate current burdens in healthcare, including chronic conditions, rare diseases, and potential pandemic outburst, as well as to upgrade current inefficient systems for biomedical data collection and distribution.

The transfer of biomedical data is achieved in real time, establishing a chain of interactions, based on advanced technologies and methodologies. These crowdsourced data sets bring a much higher degree of accuracy and value for the entire system, as they are ontologically unified and synchronized, structured into a system that can work harmoniously with existing systems in an interoperable manner.

The successful implementation of a digital health environment for global healthcare improvement, though,

Table 3. Perspectives and advantages provided from the crowdsourced platform both for patients and the entire health-related community

Patients/Donors	Health-Related Factors
• Dynamic, real-time tiered e-consent module • Self-reported disease-specific tool for monitoring and management of their disease • High protection of their health-related and personal data and/or authorization for future utilization • Actively participating in their treatment plan • User-friendly customized environment for comparability with other patients • Remote interaction with healthcare providers and practitioners in real time • Enrollment in research community as equals	• GDPR and HIPAA-compliant solution • Harmonized and ontologically unified patient-reported data sets • Data transferring and export according to standard security protocols and methodologies • Ability to map patients' profiles (linkage with biosamples-related data and EHRs) • Off-site monitoring and treatment management in real time • Patients' follow-up and re-consent options • Establishment of novel standard policies and strategies for a high-quality healthcare system

EHRs: electronic health records; GDPR: General Data Protection Regulation; HIPAA: Health Insurance Portability and Accountability Act.

must deal with current privacy and ethical principles, that alongside the lack of public trust for the proper handling of data and transparency, decreases the effective collaboration of different stakeholders and slows down the entire process.[56] Through the implementation of the current project, we met several controversies in legislation and regulations that need to be reviewed in the future. Despite that, the most significant challenge of an advanced system was to engage patients and gain their trust, especially considering patients who might not have access to such technological tools and/or lack effective education. More educational webinars, workshops, and open calls among patients' communities might be essential for patients' education and encouragement to become an active player within the entire chain.

CONCLUSIONS

In conclusion, crowdsourcing platforms have the potential to be improved and enabled with state-of-the-art technologies, empowering the different types of end users with unique opportunities for exploration, research, monitoring, and treatment management. Future goals of the system are to engage patients on a large scale and incorporate the platform with a substantial list of disease-related conditions, with the collaborative contribution from healthcare providers and research entities, creating a strong network that would make the difference in future global healthcare. Distributed ledger technologies and consensus algorithms, such as the blockchain, can strengthen information security and data integrity in the public networks of patients collecting crowdsourcing health-related data.

Acknowledgment

This manuscript is part of a registered international Patent Cooperation Treaty (PCT/GR2020/000056).

References

1. Globalisation. Council of Europe. Accessed Aug 17, 2022. www.coe.int/en/web/compass/globalisation
2. Global health challenges. British Medical Association. Accessed Aug 16, 2022. www.bma.org.uk/what-we-do/working-internationally/our-international-work/global-health-challenges
3. Hoffman SJ, Cole CB. Defining the global health system and systematically mapping its network of actors. *Global Health*. 2018;14:1–19.
4. Svensson A, Hedman E. Operations knowledge management in health care. In: *Operations Management – Emerging Trend in the Digital Era*. IntechOpen, 2021.
5. Debie A, Khatri RB, Assefa Y. Successes and challenges of health systems governance towards universal health coverage and global health security: a narrative review and synthesis of the literature. *Health Res Policy Syst*. 2022;20:1–17.
6. The 13 biggest threats to global health, according to WHO. Advisory Board. Accessed Aug 17, 2022. www.advisory.com/daily-briefing/2020/01/15/who-health-challenges
7. Gluckman PD, Low FM, Buklijas T, et al. How evolutionary principles improve the understanding of human health and disease. *Evol Appl*. 2011;4:249.
8. Benton ML, Abraham A, LaBella AL, et al. The influence of evolutionary history on human health and disease. *Nat Rev Genet*. 2021;22:269–283.
9. Hood E, Jenkins KP. Evolutionary medicine: a powerful tool for improving human health. *Evol Educ Outreach*. 2008;1: 114–120.
10. Kazmierska J, Hope A, Spezi E, et al. From multisource data to clinical decision aids in radiation oncology: the need for a clinical data science community. *Radiother Oncol*. 2020;153: 43–54.
11. Brownson RC, Eyler AA, Harris JK, et al. Getting the word out: new approaches for disseminating public health science. *J Public Heal Manag Pract*. 2018;24:102–111.
12. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 2019;28:165–173.
13. Delaye J, Cacciatore P, Kole A. Valuing the "burden" and impact of rare diseases: a scoping review. *Front Pharmacol*. 2022;13. DOI: 10.3389/FPHAR.2022.914338
14. Uhlenbusch N, Löwe B, Depping MK. Perceived burden in dealing with different rare diseases: a qualitative focus group study. *BMJ Open*. 2019;9:e033353.
15. White BAA, Eklund A, McNeal T, et al. Facilitators and barriers to ad hoc team performance. *Proc (Bayl Univ Med Cent)*. 2018;31:380.
16. Ranasinghe S, Pichler H, Eder J. Report on data quality in biobanks: problems, issues, state-of-the-art [Preprint]. *arXiv*. Published online Dec 17, 2018.
17. Govindaraj GM, Krishnakumar P, Scaria V, et al. Building on an ad hoc covid-19 response to enhance community-based care for vulnerable children in Kerala, India. *NEJM*

Catalyst. Published online Nov 4, 2020. DOI: 10.1056/CAT.20.0543.
18. González-García J, Estupiñán-Romero F, Tellería-Orriols C, et al. Coping with interoperability in the development of a federated research infrastructure: achievements, challenges and recommendations from the JA-InfAct. *Arch Public Health*. 2021;79:1–18.
19. de Mello BH, Rigo SJ, da Costa CA, et al. Semantic interoperability in health records standards: a systematic literature review. *Health Technol (Berl)*. 2022;12:255–272.
20. Rinaldi E, Stellmach C, Moses N, et al. Harmonization and standardization of data for a Pan-European cohort on SARS- CoV-2 pandemic. *NPJ Digit Med*. 2022;5:1–13.
21. Chegini Z, Janati A, Babaie J, et al. Exploring the barriers to patient engagement in the delivery of safe care in Iranian hospitals: a qualitative study. *Nurs Open*. 2020;7:457–465.
22. Kowalski C, Post S, Seufferlein T, et al. Barriers and facilitating factors for research involvement in cancer centers: a survey of colorectal cancer center coordinators in Germany, Austria, and Switzerland. *Cancer Control*. 2018;25. DOI: 10.1177/1073274818765475
23. Sims MH, Hodges Shaw M, Gilbertson S, et al. Legal and ethical issues surrounding the use of crowdsourcing among healthcare providers. *Health Informatics J*. 2019;25:1618–1630.
24. Singhal S, Patel N. Future of US healthcare: What's next for the industry post-COVID-19. McKinsey & Company. Accessed Aug 16, 2022. www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-future-of-us-healthcare-whats-next-for-the-industry-post-covid-19
25. Bowser A, Parker A, Long A. Citizen science and COVID-19: the power of the (distanced) crowd. Wilson Center. Jun 22, 2020. Accessed Aug 16, 2022. www.wilsoncenter.org/blog-post/citizen-science-and-covid-19-power-distanced-crowd
26. Schiavone F, Ferretti M. The futureS of healthcare. *Futures*. 2021;134:102849.
27. Deneux F, Sehlstedt U. The future of healthcare. Arthur D. Little. Feb 2017. Accessed Aug 19, 2022. | www.adlittle.com/en/insights/viewpoints/future-healthcare
28. Dhingra D, Dabas A. Global strategy on digital health. *Indian Pediatrics*. 2020;57:356–358.
29. What is holding healthcare back from digital transformation? TechTarget, Inc. April 27, 2022. Accessed Aug 16, 2022. healthitsecurity.com/features/what-is-holding-healthcare-back-from-digital-transformation
30. Shanley LA, Parker A, Schade S, et al. Policy perspectives on citizen science and crowdsourcing. *Citizen Science: Theory and Practice*. 2019;4:30.
31. Oehrlein EM, Graff JS, Harris J, et al. Patient-community perspectives on real-world evidence: enhancing engagement, understanding, and trust. *Patient*. 2019;12:375–381.
32. Guide to the General Data Protection Regulation (GDPR). Information Commissioner's Office, 2022.
33. Health Information Privacy. United States Government; U.S. Department of Health and Human Services; Office for Civil Rights. Accessed Oct 13, 2021. www.hhs.gov/hipaa/index.html
34. ngx-admin. Akveo, LLC. Accessed Aug 24, 2022. akveo.github.io/ngx-admin/
35. Schlieter H, Marsch LA, Whitehouse D, et al. Scale-up of digital innovations in health care: expert commentary on enablers and barriers. *J Med Internet Res*. 2022;24:1–11.
36. Bücheler T, Sieg JH. Understanding Science 2.0: crowdsourcing and open innovation in the scientific method. *Procedia Comput Sci*. 2011;7:327–329.
37. Mazumdar S, Wrigley S, Ciravegna F, et al. Citizen science and crowdsourcing for earth observations: an analysis of stakeholder opinions on the present and future. *Remote Sens*. 2017;9:87.
38. Tucker JD, Day S, Tang W, et al. Crowdsourcing in medical research: concepts and applications. *PeerJ*. 2019;4:1–17.
39. Callaghan CW. Crowdsourced 'R&D' and medical research. *Br Med Bull*. 2015:115:67–76.
40. Lichten C, D'Ioppolo R, Rebecca A, et al. Citizen Science: Crowdsourcing for Research. *Healthcare Improvement Studies Institute*. Published online May 23, 2018
41. Franzoni C, Poetz M, Sauermann H. Crowds, citizens, and science: a multi-dimensional framework and agenda for future research. *Industry Innov*. 2021;29:251–284.
42. Lee P, Abernethy A, Shaywitz D, et al. Digital health COVID-19 impact assessment: lessons learned and compelling needs. *NAM Perspectives*. Published online Jan 18, 2022. DOI: 10.31478/202201C
43. Créquit P, Mansouri G, Benchoufi M, et al. Mapping of crowdsourcing in health: systematic review. *J Med Internet Res*. 2018;20:e9330.
44. Hulsen T. Sharing is caring—data sharing initiatives in healthcare. *Int J Environ Res Public Health*. 2020;17.
45. PatientsLikeMe website. Accessed Jan 8, 2024. www.patientslikeme.com
46. Humanscape website. Accessed Jan 8, 2024. humanscape.io/en/index.html
47. Wisdo Health website. Accessed Jan 8, 2024. wisdo.com
48. RxPx website. Accessed Jan 8, 2024. www.rxpx.health
49. PicnicHealth website. Accessed Jan 8, 2024. picnichealth.com
50. carecircle website. Accessed Jan 8, 2024. www.carecircle.org
51. Mango Sciences website. Accessed Jan 8, 2024. mangosciences.com
52. Alike Health website. Accessed Jan 8, 2024. www.alike.health
53. Antidote website. Antidote Technologies, Inc. Accessed Jan 8, 2024. www.antidote.me
54. –MyHealthTeam website. MyHealthTeam, Inc. Accessed Jan 8, 2024. myhealthteam.com
55. Savvy Cooperative website. Accessed Jan 8, 2024. www.savvy.coop
56. Al Knawy B, Mckillop MM, Abduljawad J, et al. Successfully implementing digital health to ensure future global health security during pandemics: a consensus statement. *JAMA Netw Open*, 2022;5:e220214–e220214.