

Potentials and Challenges of the Health Data Cooperative Model

Ilse van Roessel^a Matthias Reumann^{b, c} Angela Brand^{a, b}

^aFaculty of Health Medicine and Life Sciences (FHML), Maastricht University, Maastricht, The Netherlands;

^bThe United Nations University – Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT), Maastricht University, Maastricht, The Netherlands; ^cIBM – Research, Zurich, Switzerland

Keywords

Health data cooperative · Big data · Electronic patient record · Mobile health apps · P4 medicine

Abstract

Introduction: Currently, abundances of highly relevant health data are locked up in data silos due to decentralized storage and data protection laws. The health data cooperative (HDC) model is established to make this valuable data available for societal purposes. The aim of this study is to analyse the HDC model and its potentials and challenges.

Results: An HDC is a health data bank. The HDC model has as core principles a cooperative approach, citizen-centredness, not-for-profit structure, data enquiry procedure, worldwide accessibility, cloud computing data storage, open source, and transparency about governance policy. HDC members have access to the HDC platform, which consists of the “core,” the “app store,” and the “big data.” This, respectively, enables the users to collect, store, manage, and share health information, to analyse personal health data, and to conduct big data analytics. Identified potentials of the HDC model are digitization of healthcare information, citizen empowerment, knowledge benefit, patient empowerment,

cloud computing data storage, and reduction in healthcare expenses. Nevertheless, there are also challenges linked with this approach, including privacy and data security, citizens’ restraint, disclosure of clinical results, big data, and commercial interest. **Limitations and Outlook:** The results of this article are not generalizable because multiple studies with a limited number of study participants are included. Therefore, it is recommended to undertake further elaborate research on these topics among larger and various groups of individuals. Additionally, more pilots on the HDC model are required before it can be fully implemented. Moreover, when the HDC model becomes operational, further research on its performances should be undertaken.

© 2018 The Author(s)

Published by S. Karger AG, Basel

Introduction

In today’s world, it is potentially possible for individuals to obtain a complete overview of their health status as a result of advances in medical technologies in the last decades. In 2003, the human DNA was fully identified for the first time as a result of the Human Genome Project.

In total, the project lasted 13 years and the costs were estimated around USD 2.7 billion dollars [1]. Current models predict that the costs for sequencing the human genome will become less than USD 1,000 within a few years [2]. Additionally, new developments in genome sequencing have made it possible to read the human DNA within 3 h [3]. Due to these improvements, it will be increasingly possible in the near future to apply DNA sequencing in healthcare settings [4]. Furthermore, in 2015, more than 165,000 mobile health (mhealth) apps were available on the market [5]. These apps are able to monitor a variety of health determinants including weight, blood pressure, sleeping patterns, nutrients intake, and physical activity levels. As a consequence of these technological advances, the amount of available personal data will expand considerably [4]. The US Institute of Medicine states that enhancements in molecular biology including genome sequencing in combination with improvements in health informatics systems such as mhealth apps and electronic patient records (EPR) have contributed to the development of a new approach to medicine. This medical paradigm is specifically called P4 medicine, which stands for personalized, preventive, predictive, and participatory medicine [6]. It studies the relation between the human genotype and phenotype [4]. All humans have a unique genetic constitution and lifestyle; therefore, they require a distinct healthcare approach. P4 medicine makes it possible to no longer administer “one-size-fits-all” medicine, but rather apply customized care [7]. It has the potential to enhance healthcare by improving disease prevention, increasing the accuracy of diagnoses, providing safer medication prescriptions, and making treatments more effective [8]. Additionally, it is likely to bring worldwide rising healthcare costs down and promote healthcare innovation [9]. Fundamental to P4 medicine is aggregated health data of individuals to evaluate the interaction between genome, environment, and behaviour. Even though large amounts of data have been produced so far, these are stored in various data silos including mhealth apps on mobile phones, EPRs, and genome sequencing companies and are inaccessible to researchers due to decentralized storage and data protection laws [4, 10]. Health data are not only indispensable for researchers to further develop P4 medicine, this information is as well of interest to a variety of other stakeholders including pharmaceutical companies, research organizations, healthcare providers, insurance companies, and commercial entities [4, 11].

The health data cooperative (HDC) model is established as an approach to make health data available to

society [4]. An HDC is a health data bank. The account holders, also referred to as members, users, or citizens, can collect and store their health-related information, for example, from the EPR, fitness or sleeping apps, heart rate monitors, and glucose meters on their HDC account. They are the rightful data owners and make the decisions regarding information sharing [12]. To date, the HDC model is no longer only a hypothetical concept but is in the phase of becoming operational [13]. MIDATA.coop is an example of an HDC. Currently, they run several projects to validate and improve the MIDATA.coop concept. One of the projects focuses on establishing an ethical and legal framework on how to handle data privacy, ownership, and consent. Another project collects data of patients who have undergone bariatric surgery. The data are stored in the individual MIDATA account of the patient. A mobile app enables the patients to follow their progress [14]. Nevertheless, more pilots are required on the HDC model before it can be fully implemented [13]. Therefore, this article focuses on the introduction of the HDC model in the European Union (EU). It aims to provide an overview of the HDC model and its features. Subsequently, it evaluates the potentials and challenges of the HDC model.

Core Principles of the Health Data Cooperative Model

An HDC is concerned with the collection, storage, maintenance, management, and analysis of health data. To become part of an HDC, an individual has to pay a one-time unit price (membership fee), which entitles the person to be a member and owner at the same time [4]. This ownership involves two different aspects. Firstly, the data are citizen owned and managed; the account holder is the only person empowered to add, adjust, or remove information and to decide when and with whom he or she wants to share personal health information. These third parties can be, among others, relatives, physicians, research institutions, public health services, pharmaceutical organizations, and technological companies [14]. Secondly, an HDC is the equal property of all its members. This cooperative approach ensures that decisions are made collectively. Each member has one vote, which, for example, can be used during elections on a new executive management. As a result of this approach, the HDC model is citizen centred, which entails that the cooperative acts in the interest of its users [4]. Another important feature is that an HDC is a not-for-profit organization [14].

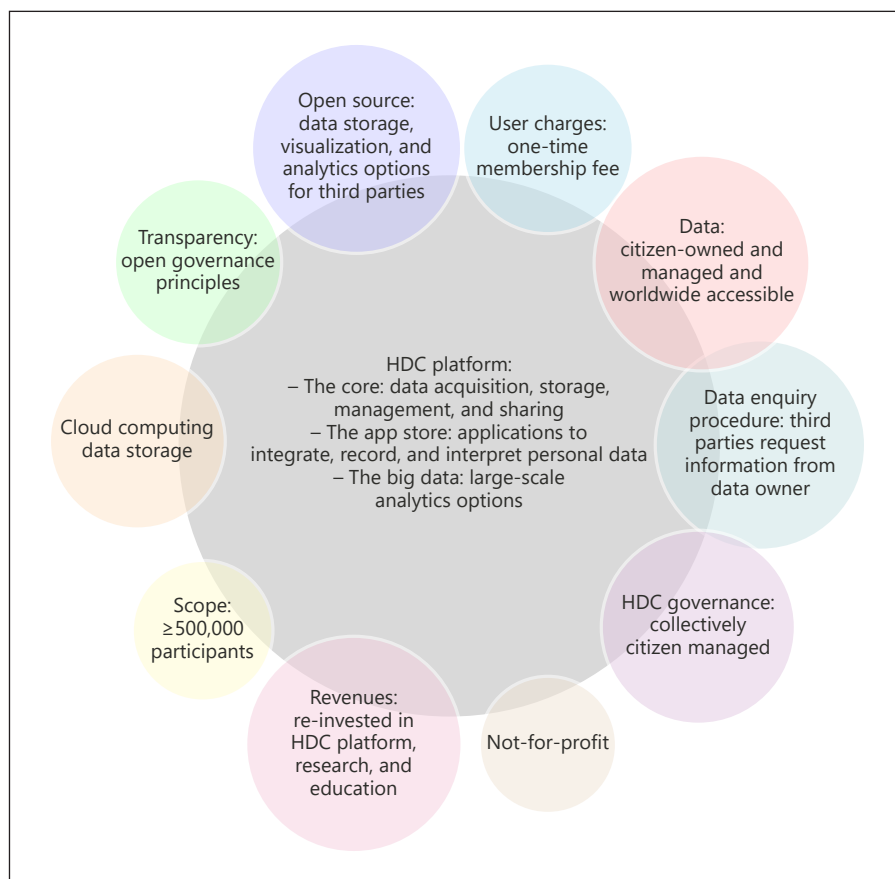


Fig. 1. Figure illustrating the components of the health data cooperative (HDC) platform and the core principles of the HDC model [4, 14].

Certain third parties, such as companies and research organizations, are required to financially compensate the HDC when members accept a data enquiry and agree to share their information [4]. Nevertheless, these earnings will be reinvested into activities such as research projects, education programs, and the HDC platform itself [4, 14]. The scope of the cooperative can differ based on the various stakeholders' needs. For example, it can be regional, national [14], or disease specific, but it requires at least 500,000 citizens to be viable [15]. Due to the fact that the HDC is accessible online, users have worldwide access to their personal account with their login details. As a result, they can update their information when and where it is convenient. A cloud computing company can provide, maintain, and safeguard a depository in the cloud for the digital data of the HDCs [4]. Moreover, the data undergo encryption procedures, which disguises the information, and only the data owner possesses the key to decrypt it. Another feature of the HDC model is its transparency about governance principles. This entails that members are informed and can engage in decision-making pro-

cesses. Furthermore, the HDC provides third parties with data storage, visualization, and analytics options to obtain, process, and extract knowledge from relevant data [14].

The Health Data Cooperative Platform

The core principles of the HDC model are integrated into the HDC platform, which is a transparent interaction platform (Fig. 1). This platform is composed of three different components and provides its users with the possibility to perform various functions. The first aspect is named the “core” and is concerned with the acquisition, storage, and management of data. HDC users can store their health data, for instance from an EPR and mhealth apps, to the core of their personal online HDC account. The user has the opportunity to store data in various formats, to organize the data into different groups such as wellness, fitness, nutrition, and EPR, and to share the personal information with relatives, the general practitioner,

research institutions, companies, or others. The second part is the “app store,” which provides the user with applications to integrate, record, and interpret the personal health data, for example, visualizing a pattern over time. The third component is called the “big data” analytics system and can be used to analyse large quantities of various types of data stored in the “core” and “app store.” For this process, personal identification information is eliminated from the data, which makes it improbable to trace the health data back to the individual. Subsequently, analytic techniques are applied to evaluate and interpret the data. This results, for instance, in new knowledge on the causes, progression, and treatment of diseases [4].

Potentials of the Health Data Cooperative Model

Digitization of Healthcare Information

The HDC platform contributes to the digitization of healthcare information. EPRs replace paper medical dossiers and can be stored on a personal HDC account. One of the benefits of digitizing health information is that it reduces the risk of medical mistakes caused by poor handwriting. Research shows that computerizing health records can decrease medication mistakes by 55%. If, in addition, the clinical decision support function of the EPR is used, which advises the doctor on patient care, the risk of medication mistakes potentially reduces by 83%. Moreover, computerizing medical dossiers is also more efficient because records are easy to access and information exchange goes effortlessly [16]. Another benefit of the platform is that it is not limited to a single data format but allows all types of data from a variety of sources to be included. This is essential because to date, no uniform data format has been established in the EU and member states have different healthcare standards [4].

Citizen Empowerment

The HDC model has a citizen-centred approach that aims to create citizen empowerment as a result of control and ownership over personal health data [4]. Members could provide their health information for research, surveillance activities, needs assessments, planning, evaluation, and other study purposes [4, 17]. Whiddett et al. [18] identify that the willingness to share personal health information increases when individuals have control over their own data and the information is anonymous. Weitzman et al. [19] suggest similar conditions for patients’ data participation in research. Anonymity, control over sharing what information and with whom, and pro-

vision of an audit trail of health information accessibility and sharing are prerequisites for individuals to share their medical information. People highly value data anonymity because it might reduce the risk of employment and insurance discrimination. A subsequent study by Weitzman et al. [17] identifies that individuals also value unlimited access to their personal data. As a result of citizen empowerment, it is likely that people become actively involved in healthcare. They might engage in decision-making processes concerning prevention, diagnosis, and treatment of diseases, start utilizing health informatics systems such as EPRs and mhealth apps, or share personal information about their risk of genetic susceptibility with relatives. Moreover, individuals could approve the use of their health information for research, contribute to the development of patient registries, or discuss patient values with scientists [20].

Knowledge Benefit

HDCs are valuable to epidemiological research because the process of data collection for certain study purposes becomes more efficient, individuals eligible for a study can be notified effortlessly, the number of study participants can be expanded, and additional information of research projects can be shared via links to websites on the HDC platform. As a result, the costs associated with epidemiological studies could be reduced [21]. Moreover, HDC members can maintain their account by regularly updating it with their newest health information [4]. This is especially the case if citizens connect the mhealth apps to their account, which continuously generate data [22]. Over time, this results in an increased amount of data, which potentially strengthens the external validity of research outcomes. For third parties, these features make it considerably convenient to conduct clinical trials or longitudinal studies [21].

Furthermore, the development of P4 medicine is interlinked with the establishment of HDCs because HDCs improve the availability of big data and offer a data analytic environment. P4 medicine could lead to cost-effective healthcare due to the fact that it stratifies diseases and patients into clinically relevant groups, provides customized care, and focuses on treating the cause of an illness rather than the symptoms. P4 medicine might also make citizens aware of the impact of lifestyle on their health. Consequently, they might undertake action to enhance their health status. Additionally, application of tailor-made therapy and advice can prevent the development of chronic diseases or improve the quality of life of patients with a chronic disease. Moreover, aggregated data make

it possible for P4 medicine to develop new multilevel biological networks. These might help to understand currently unknown causes of diseases and to identify which biological systems are disturbed in bodies during illness [6, 23]. As a result, prevention and predictive disease models might be established. Prevention is an important aspect of public health; therefore, national public health institutions might use information from these models to improve existing or develop new campaigns [24].

Moreover, the HDC model provides a procedure which makes it possible for research to disclose study outcomes. The HDC model requires third parties to follow a data enquiry procedure when they would like to obtain health information from members. If these individuals give their consent, the third party is allowed to use the requested data for the specific study purpose. After the research is conducted and the study results are obtained, this information can be send back to the participants. These individuals, the data owners, can then decide if they allow the outcomes of the study to be disclosed [4]. There are several advantages associated with revealing the study results to the participant in the trial. Firstly, it might provide the participant with a confirmation that they do or do not have a certain disease. Secondly, it can help with disease management including prevention, diagnosis, and treatment of illnesses. Thirdly, it might also guide individuals in making medical, reproductive, or lifestyle-related decisions. Another reason for sharing the study results with the participant is that it is simply their right to obtain this information because they have contributed decisive information to the study [25]. Additionally, there is a population benefit associated with the data enquiry regulation of the HDC model. This procedure enables third parties to make individual results public if study participants give their consent for disclosure. Currently, there is no such regulation; as a consequence, informative and relevant study results cannot be revealed because of the limited given consent [4]. In addition, consent management such as dynamic consent could be a suitable approach for the HDC model. This concept allows HDC members to automatically control third parties' access to their health information. Members have to indicate access permissions on their HDC account. When a third party sends an information enquiry, this is compared with the settings on the account of the individual. Consequently, based on the preferences defined by the data owner, access to the data is automatically granted or denied. In case no access permission for an enquiry is defined, it is compared to the default settings. When it is still not possible to determine the access status from the de-

fault preferences, the data owner is provided with a consent user interface, which requires the individual to select one of the consent options [26]. The advantage of this method is that it generates flexible and re-usable access rights [27].

Finally, the big data analytic options of the HDC platform provide several advantages to healthcare organizations. Firstly, illnesses are likely to be discovered at an earlier onset. Consequently, treatment might be more effective and the disease better treatable. Secondly, fraud in a healthcare institution could be traced sooner and efficiently. Thirdly, it enables researchers to analyse historical data including the length of stay, patients opting for elective surgery, patients vulnerable to medical complications, patients susceptible to methicillin-resistant *Staphylococcus aureus*, sepsis, or other hospital-acquired diseases, potential causal factors of illness progression, complication rates, and comorbidities. This supportive information provides healthcare professionals with useful knowledge and insights, which could contribute to decision-making processes regarding the healthcare organization [28, 29].

Patient Empowerment

The HDC model aims to create patient empowerment [4]. This patient care approach was first proposed in the field of diabetes in the 1990s. It entails that patients are responsible for their own health and well-being and become proactive participants in the care process. This is in contrast with the more traditional medicine approach in which doctors have the control and make the final decisions [30]. In subsequent years, this philosophy has expanded among patients due to the familiarity and affordability of computers and Internet in daily life. This revolution has given patients access to a tremendous amount of medical information, via which they become educated and familiar with medical terminology [31]. These patients are called proto-professionals because they internalize the medical knowledge and use it to take part in disease management and decision-making processes regarding their own health and that of their relatives [32]. In 2005, Andreassen et al. [33] reported that 44% of the study participants from seven member states of the EU accessed the Internet to obtain health information. E-patient is the specific term usually used to describe the younger generation, which has been raised with the Internet. Not only do they, via the Internet, have access to basic health and medical information, but also to the latest developments and various treatment modalities. Patient empowerment has several benefits, including patients asking more relevant,

accurate, and well-informed questions due to the knowledge they have obtained, they are more likely to adhere to prescribed treatment plans, their information can add to making a collaborate physician-patient decision, and the prior knowledge of the patient reduces the time necessary to explain medical information [31].

Cloud Computing Data Storage

The cloud computing structure of the HDC creates several advantages for healthcare organizations. Due to the increasing amount of available data, data management has become an essential aspect of modern healthcare. A first beneficial aspect of cloud computing data storage is that it enables healthcare professionals and citizens/patients to access their data at any time and place. This enables clinicians to be more flexible in their work as they are not confined to their office to access patients' health records. Healthcare professionals can also exchange information about a patient with their colleagues at other locations to inform or negotiate about possible medical options. In addition, cloud storage allows an unlimited number of users to access the HDC platform simultaneously and it provides a high-speed storage and retrieval procedure. A second advantage of cloud computing is that it has the potential to upgrade telemedicine. It creates the opportunity to connect patients and physicians live while being in other places. Therefore, it is no longer necessary to have a consultation in the hospital for certain medical conditions, treatment, therapy, or advice. This might decrease hospital visits for patients and home visits for doctors, resulting in reduced travel expenses and time. A third promising feature is that cloud computing tools can be used for community education. The integration of chat options, forums, blogs, and databases with articles can inform the HDC users or bring patients into contact, enabling them to share their experience and knowledge on a certain disease and treatment options. This aspect might contribute to patient empowerment as they gain more knowledge and consequently are able to become actively involved in disease management and the medical decision-making process [34]. Finally, the cloud computing structure of the HDC does not allow transferring data from the HDC platform to the digital environment of a third party. Instead, the HDC provides third parties with an internal data analytic environment, the "big data," which entails that the data is kept centralized within one place [4].

Reduction in Healthcare Expenses

Healthcare costs are considered an important point on the agenda of the EU. Since the 1960s, these expenses have

been rising steadily. In 2006, the average public and private healthcare spending was 9% as a share of GDP [35]. In 2009, a maximum of 10.5% was reached; however, in the following years, its share gradually decreased. As a result, in 2012, an average of 10.2% was reported [36]. Nevertheless, the Organization for Economic Co-operation and Development (OECD) predicts that by the year 2050 the average healthcare spending will be 13% as a share of GDP in OECD countries and yet, when a strict cost-containment is maintained, the average will be around 10%. These expenses are a great burden for the public budgets of countries [37]. The ageing European population is often considered to be one of the main causes of the rising healthcare expenditures; however, the World Health Organization (WHO) argues that this aspect only has a small contribution; less than 10%. The major reason for the increase in healthcare expenses is technological advances, contributing between 50 and 75% to the growth in costs [38]. One of the proposed solutions to reduce healthcare expenditures in the EU is to digitize the current healthcare sector. In Australia, it is estimated that this adaptation is likely to reduce the spending on healthcare by AUD 7.6 billion a year. This saving is the result of digitalization of the summary care record, decision support, patient self-management, the EPR, quality and performance management, and medication management. Digitizing the health sector does not only result in more efficient procedures in the healthcare sector but might also drastically decrease visits and admissions to the hospital and the number of required X-rays and lab tests [39]. Furthermore, the McKinsey Global Institute estimates that utilization of big data in US healthcare can save USD 300 billion annually and create 0.7% productivity growth per year. In addition, Europe's public sector administration could reduce expenditures with EUR 250 billion annually and obtain 0.5% productivity growth per year [40]. Due to the fact that HDCs are considerably involved in the digitalization of the healthcare sector, their establishment can potentially contribute to the reduction of the healthcare expenditures of EU member states.

Challenges of the Health Data Cooperative Model

Resistant Healthcare Professionals

The implementation of the HDC model will cause fundamental changes in healthcare, which can be expected to lead to resistance from healthcare professionals. A first significant shift might be the consequence of patient em-

powerment. Medical practitioners have been reluctant and opposed to this change. The main reasons for this are that they assume much health information on the Internet is inaccurate, they are concerned that patient contact will reduce, and they fear that informed patients become stubborn and problematic [31]. The second alteration is likely to result from the concept of big data analytics. Currently, most physicians have not been educated to analyse and interpret aggregated health information of their patients. Consequently, they are unable to provide personalized treatment, therapy, or advice. An increasing number of patients are interested in healthcare professionals who are familiar with these new technologies and are able to integrate them in their daily work. Therefore, a new type of healthcare specialist is suggested, called healthcare and wellness coach. This professional will be a guide in the field of data interpretation and can give advice and support in making healthy lifestyle decisions based on patients' personal data [41]. This development could result in a decreased need for traditional physicians and nurses and cause a shift to health experts experienced with big data analytics or require clinicians to be retrained.

Privacy and Data Security

Data security and privacy are important aspects of the HDC model [14]. Nevertheless, it is argued that it is not possible to guarantee absolute data security and privacy of highly personal health data [4]. The data is marked with a distinctive identifier to be able to inform citizens correctly about their diagnosis or treatment after the analysis of their anonymous personal information [42]. A possible re-identification technique, which makes use of these tags, is referred to as data linkage. It links various datasets with each other and subsequently matches a person's identity to the personal health data [43]. However, even when no identifiers are used in a dataset, it is possible to relate personal genetic data to the right individual with the surname inference strategy. This method makes use of the tradition that it is common in many countries to pass on the surname to the offspring. Genetic genealogy databases are registers of surnames and associated haplotypes. These databases are accessible via the Internet and enable users to find potential surnames with related information such as geographical locations, spelling variants, and pedigrees of persons genetically associated with the specific haplotypes. When besides the personal genetic data, additional information such as age or demographic data is available in a dataset, this makes it even more convenient to trace back the identity. Already when only a limited amount of genetic data be-

comes publicly available, this method poses a risk to the privacy of the individual, the family members, and even unacquainted persons [44]. The re-identification of personal information might cause disadvantages including social stigmatization and insurance or employment discrimination. In addition, it can possibly also have an adverse effect on research, when re-identification threatens the trust and willingness of citizens to participate. As a result, fewer individuals might volunteer to take part in studies [43].

Moreover, challenges also arise from storing sensitive personal medical data in a cloud computing environment. A characteristic of the HDC cloud computing storage is that the data is centralized, which means that the information is kept in one place. In this cloud, the privacy and security of data can never be fully assured. In addition, there are also several other unpredictable factors which might pose a threat to the confidentiality of the data stored in the cloud. These are, among others, unauthorized access by hackers, unsuccessful data separation, insufficient data encryption key management, public management interface issues, and privilege abuse [45].

Citizen's Restraint

There are several issues which might lead to restraint of citizens to participate in HDCs. Firstly, citizens' knowledge about the options of sharing their EPR is not up to date. Weitzman et al. [17] find that less than one-fourth (22.6%) of the study participants were aware that they could share parts of their EPR rather than their entire medical data file. Secondly, citizens are not always willing to provide their personal medical information to third parties. Individuals are most resistant to sharing information about mental illness, genetic disorder, tobacco, alcohol and other substances consumption, and sexually transmitted diseases. Research proposes that the resilience towards the sharing of personal health information is due to a lack of trust in a public agency's data management, distrust in appropriate data utilization, concern about unauthorized data revelation, insufficient transparency, and fear of anonymity breach. Citizens are afraid that the insensitive handling of their data might increase the risk of stigmatization or discrimination [17]. Resistance towards the sharing of health information poses a threat towards the HDC model, since the principle of data exchange is fundamental to the concept. To overcome this problem, the blockchain infrastructure could be introduced to the HDC model. This is a digital data network which enables secure and convenient data sharing. If a data owner shares personal data, this information gets en-

rypted with a private key and is sent to the public address of the receiver. The data can only be accessed with the corresponding private key of the receiver [46]. Moreover, this system is transparent about transactions [47] and provides records about data provenance [46].

Thirdly, there has been a trial with a health record bank in the USA requiring citizen to pay a one-time membership fee. This trial failed and has resulted in several lessons learned from this experience. One of the main suggestions is that it should be free to obtain an HDC account since this will be the only way to balance personal benefit and the common good within an HDC. This increases the chance that enough interested participants will be willing to become members and subsequently to establish a viable cooperative. A survey conducted in the USA shows that only 20% of the participants would purchase an account of the health record bank if their physician recommended it [48]. In addition, the HDC model does not pay the revenues of data enquiries to the study participants but reinvests this into the cooperative [4]. This might lead to disinterest from the members, as there are no incentives to maintain and update their personal health information and to contribute their data to research purposes. Therefore, another recommendation is to pay a proportion of the generated revenues to the data contributors. In this situation, active members will have an interest-bearing account and receive a small amount of money on a yearly base, similar to the saving accounts of financial banks [48]. Fourthly, the WHO reports that chronic diseases are the leading cause of disability and death globally [49] and the risk of getting a chronic illness accumulates with age [50]. Additionally, the Centers for Disease Control and Prevention (CDC) reveal that in the USA alone, since 2012, 50% of the adult population (117 million) suffer at least from one chronic disease and strikingly, 25% of all adults are diagnosed with two or more chronic illnesses [51]. People with chronic conditions might especially benefit from the HDC structure, as it supports self-management of their diseases. The HDC platform enables them to have their entire medical history in one place, to monitor their health over time, and share health information with relatives, the general practitioner, or other third parties [4]. Pagliari et al. [52] report that patients in need of episodic care or suffering from long-term conditions are most likely to utilize and benefit from health informatics tools as it provides them with up-to-date information and communication options and enables them to monitor their disease(s). Archer et al. [53] show similar findings. They identify that in particular elderly, chronically ill, or people with a dis-

ability are likely to make use of such systems. Nevertheless, these people encounter several barriers including poor contrast, unsuitable font size, difficulty navigating through the system, crowded or cluttered screens, problems with vision, fear of technology, health illiteracy, insufficient computer and Internet skills, and physical and cognitive constraints. Finally, the HDC model might contribute to health disparities as a result of digital divide [52].

Disclosure of Clinical Results

The HDC model makes it possible to disclose individual clinical results to study participants. Nevertheless, this possibility has been criticized. The main argument is that clinical study outcomes cannot always be guaranteed to be accurate [25]. Even though diagnostic testing methods provide essential information for clinical decision making, they have a risk of generating incorrect or unreliable test results [54]. When an individual receives a false positive test result, this has major implications for this person and relatives. It might lead to worry and concern or cause an individual to make unnecessary medical or financial decisions such as opting for a treatment. The opposite could also occur; an individual receives a false negative test result. This person is mistakenly informed to be healthy and consequently does most likely not obtain treatment. It might only be discovered that the diagnosis was incorrect when it is too late. A second reason not to reveal clinical results is that study participants might not know how to interpret clinically insignificant results. As a consequence, this could produce needless feelings of stress and anxiety or they might undertake baseless action. For example, this is the case when a test shows that particular genes increase the susceptibility to acquire Alzheimer's disease. As a result, the study participants could be informed about their vulnerability. However, currently, there is no treatment or prevention available for Alzheimer's disease and therefore, it might cause a feeling of impotency to these individuals and their family. The same scenario might occur when metal exposure levels are measured in individuals. Research shows an association between the exposure to metals and cancer in animals. Nonetheless, this risk has so far not been extensively studied among humans and it is unknown which levels of exposure might potentially be harmful and what the source of exposure is. If scientists communicate the personal exposure levels to the study participants, this might create uncertainty because no inferences have been made about safe or maximum exposure levels. Due to these arguments it is suggested that the research out-

comes should only be provided to study participants when they are produced by validated tests and are clinically significant. Furthermore, another point raised against disclosure is that it might create additional costs to the research institution. These extra expenses are primarily due to counselling and advice, which should be provided together with the revelation of the study outcomes to help the participants comprehend the clinical results. To not exceed the budget, there should be a cut on other aspects of the research, which might, for example, result in less study participants being included. The final counterargument entails that disclosing the research outcomes to the individuals might bring about a therapeutic misconception among study participants. This misunderstanding means that the individuals involved in the study think that the primary aim of the research is to provide them with personal clinical information and care for disease prevention and health promotion, while in fact, the study is undertaken to enhance the knowledge on a certain research topic. As a consequence of this phenomenon, study participants are not fully aware of the risks associated with certain clinical trials and at the same time overestimate the advantages. This misconception might compromise the signed informed consent because these study participants believe they have agreed to a different study purpose than is really the case [25].

Big Data

The “Big data” analytics system of the HDC platform can be used to analyse large quantities of health data. Even though big data has the potential to provide current healthcare systems with new insights and knowledge to enhance services and reduce expenses, it should be interpreted carefully [44]. The Google Flue Trend (GFT) is a recent case in which big data did not result in accurate predictions of the future. It was initially developed in 2008 as a combined winter and flue detector, but the model got updated in 2009 to predict influenza trends. In 2013, it was identified that the GFT had persistently overestimated the doctor visits for influenza-like illness for already some time and that for the same year, the model strikingly predicted more than twice the forecast of the CDC. Several factors contributing to this discrepancy were identified and they are not restricted to the GFT. Firstly, “big data hubris” is a term used to describe the incorrect supposition that big data are a replacement for conventional data collection and analysis and are not regarded as complementary. Even though big data have major potentials for science, it is regularly composed of data not produced by validated and reliable measurement

tools [55]. Secondly, algorithm dynamics, which are the changes made to enhance the model, may also cause errors in the predictions. This occurs when the adjustments are based on incorrect or inaccurate data or when there is a difference in pace at which the algorithm and the collected data are altering [44, 55]. Moreover, Andreu-Perez et al. [44] suggest some additional elements contributing to the imprecision of big data interpretation. The model could be developed from a training set with a substantial number of outliers included. These values do not contribute to an accurate prediction of future trends and consequently, the model is imprecise. Another possibility is that the number of parameters overfits the data. This means that the model captures the data too well because it does not only predict the underlying trend but also the outliers. Therefore, the model does not differentiate between a truly underlying pattern and noisy data and is not able to correctly classify any new instances. As a result, it cannot be used to make inferences about the future.

Commercial Interest

The HDC accounts of individuals are commercially valuable because of the personal health information they contain [56]. Presser et al. [57] report that currently, the legislation in the United Kingdom (UK) is not sufficient to inhibit commercial enterprises making profit from personal health data. Moreover, Blenner et al. [58] show that in the case of diabetes apps, 86.2% of the apps (56 out of a random subset of 65) feature cookies, which enable these apps to routinely collect data. This information is regularly shared with third parties, such as advertisers. However, a survey in the UK identifies that only 16% of the respondents are aware of how commercial companies utilize personal health data. In addition, merely 37% of the study participants approve the use of anonymized health information by these organizations for marketing purposes [59].

Limitations and Outlook

This article includes several studies with a limited number of study participants. Even though these studies provide some insight into the potentials and challenges of the HDC model, no general inferences can be drawn from this information. It is recommended to undertake further elaborate research among larger and various groups of people to obtain a more accurate and all-encompassing understanding of how these concepts are perceived in the population. Furthermore, currently, the HDC model is in

the phase of becoming operational [13]. More pilots are required on the HDC model before it can be fully implemented. This article presents preliminary findings on the topic showing that the HDC model could be implemented in society in the EU and beyond. Any experiences or pilots with HDCs can provide further essential information for prospective research and successful implementation.

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

- 1 Wohlsen M: Biopunk – Solving Biotech’s Biggest Problems in Kitchens and Garages. The Penguin Group; 2012, pp 1–218.
- 2 Weterstrand K: DNA Sequencing Costs – Data from the NHGRI Genome Sequencing Program (GSP) (Internet). National Human Genome Research Institute. [cited 2016 January 16]. Available from: <http://www.genome.gov/sequencingcosts/>.
- 3 Kelly BJ, Fitch JR, Hu Y, Corsmeier DJ, Zhong H, Wetzel AN, Nordquist RD, Newsom DL, White P: Churchill: an ultra-fast, deterministic, highly scalable and balanced parallelization strategy for the discovery of human genetic variation in clinical and population-scale genomics. *Genome Biol* 2015;16:1–14.
- 4 Hafen E, Kossmann D, Brand A: Health Data Cooperatives – Citizen Empowerment. *Methods Inf Med* 2014;53:82–86.
- 5 Aitken M, Lyle J: Patient Adoption of mHealth – Use, Evidence and Remaining Barriers to Mainstream Acceptance. Parsippany, USA; 2015.
- 6 Flores M, Glusman G, Brogaard K, Price ND, Hood L: P4 medicine: how systems medicine will transform the healthcare sector and society. *Futur Med* 2013;10:565–576.
- 7 Vogt H, Hofmann B, Getz L: The new holism: P4 systems medicine and the medicalization of health and life itself. *Med Heal Care Philos* 2016;19:307–323.
- 8 The Jackson Laboratory. What Is Precision Medicine? (Internet) 2016 [cited 2017 May 18]. Available from: <https://www.jax.org/research-in-action/precision-medicine-and-you/what-is-precision-medicine>.
- 9 Hood L: Systems Biology and P4 Medicine: Past, Present and Future. *Rambam Maimonides Med J* 2013;4:1–12.
- 10 Horgan D, Jansen M, Leyens L, Lal JA, Sudbrak R, Hackenitz E, Bußhoff U, Ballensiefen W, Brand A: An index of barriers for the implementation of personalised medicine and pharmacogenomics in Europe. *Public Health Genomics* 2014;17:287–298.
- 11 Miliard M: Facebook would like your health data – Social media giant reportedly eyeing new healthcare products; privacy questions remain. (Internet) 2014 [cited 2016 Jan 16]. Available from: <http://www.healthcareitnews.com/news/facebook-would-your-health-data>.
- 12 Healthbank. Healthbank – My data, my choice, my health. (Internet) 2016; p 1. Available from: <https://www.healthbank.coop/faq/>.
- 13 Horgan D: “Big Data” and the co-operative solution. (Internet) 2014 [cited 2016 Feb 8]. Available from: <http://www.euapm.eu/articles,1,39.html>.
- 14 MIDATA.coop. MiDATA.coop My Data – Our Health. (Internet) [cited 2016 Feb 8]. Available from: <https://midata.coop>.
- 15 OpenData Manchester. Open Data Cooperation – Building a data cooperative. (Internet) 2014 [cited 2016 Jan 16]. Available from: <http://opendatamanchester.org.uk/2015/04/14/open-data-cooperation-building-a-data-cooperative/>.
- 16 Menachemi N, Collum TH: Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy* 2011;4:47–55.
- 17 Weitzman ER, Kelemen S, Kaci L, Mandl KD: Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. *BMC Med Inform Decis Mak* 2012;12:1–10.
- 18 Whiddett R, Hunter I, Engelbrecht J, Handy J: Patients’ attitudes towards sharing their health information. *Int J Med Inform* 2006;75:530–541.
- 19 Weitzman ER, Kaci L, Mandl KD: Sharing Medical Data for Health Research: The Early Personal Health Record Experience. *J Med Internet Res* 2010;12:e14.
- 20 Budin-Ljøsne I, Harris JR: Ask Not What Personalized Medicine Can Do for You – Ask What You Can Do for Personalized Medicine. *Public Health Genomics* 2015;18:131–138.
- 21 Powell J, Bunchan L: Electronic Health Records Should Support. *J Med Internet Res* 2005;7:1–6.
- 22 Arora S, Yttri J, Nilsen W: Privacy and Security in Mobile Health (mHealth) Research. *Alcohol Res* 2014;36:143–150.
- 23 Schee genannt Halfmann S, N Evangelatos N, Schröder-Bäck P, Brand A: European healthcare systems readiness to shift from “one-size fits all” to personalized medicine. *Personalized Medicine* 2017;14:63–74.
- 24 Khoury MJ, Gwinn M, Glasgow RE, Kramer BS: A Population Perspective on How Personalized Medicine Can Improve Health. *Am J Prev Med* 2012;42:639–645.
- 25 Resnik DB: Disclosure of Individualized Research Results: A precautionary Approach. *Account Res* 2011;18:382–397.
- 26 Dunn MW: User-centric consent management system and method. US Patent No. 7,076,558, 2006, pp 1–45.
- 27 Asghar MR, Russello G: Automating Consent Management Lifecycle for Electronic Healthcare Systems; in *Medical Data Privacy Handbook*. Springer International Publishing; 2015, pp 361–387.
- 28 Raghupathi W, Raghupathi V: Big data analytics in healthcare: promise and potential. *Heal Inf Sci Syst* 2014;2:1–10.
- 29 Leyens L, Reumann M, Malats N, Brand A: Use of big data for drug development and for public and personal health and care. *Genet Epidemiol* 2017;41:51–60.
- 30 Anderson RM: Patient Empowerment and the Traditional Medical Model – A case of irreconcilable differences? *Diabetes Care* 1995;18:412–415.
- 31 Akerkar SM, Bichille LS: Doctor patient relationship: changing dynamics in the information age. *J Postgrad Med* 2004;50:120–122.
- 32 Caron-Flinterman JF, Broerse JEW, Bunders, Joske FG: The experiential knowledge of patients: a new resource for biomedical research? *Soc Sci Med* 2005;60:2575–2584.
- 33 Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, Voss H, Wynn R: European citizens’ use of E-health services: A study of seven countries. *BMC Public Health* 2007;7:1–7.
- 34 Zafar Z, Islam S, Aslam MS, Sohaib M: Cloud Computing Services for the Healthcare Industry. *Int J Multidiscip Sci Eng* 2014;5:25–29.
- 35 European Commission. European Economy – Projecting future health care expenditure at European level: drivers, methodology and main results. (Internet) 417th ed. Brussels; 2010 Jul. Available from: http://ec.europa.eu/economy_finance/publications/economic_paper/2010/pdf/ecp417_en.pdf.

- 36 Nuffield Trust: UK health spending as a share of GDP. (Internet) [cited 2016 Feb 21]. Available from: <http://www.nuffieldtrust.org.uk/data-and-charts/uk-health-spending-share-gdp>.
- 37 de la Maisonneuve C, Martins JO: Public spending on health and long-term care: a new set of projections. (Internet) 6th ed. Paris; 2013. Available from: https://www.oecd.org/eco/growth/Health_FINAL.pdf.
- 38 World Health Organization: Rising health care costs due to technology. (Internet) 2011 [cited 2016 Feb 21]. Available from: <http://www.euro.who.int/en/about-us/regional-director/news/news/2011/03/rising-health-care-costs-due-to-technology2>.
- 39 Royle R, Hambleton S, Walduck A: Review of the Personally Controlled Electronic Health Record. (Internet) Australia; 2013 Dec. Available from: [http://health.gov.au/internet/main/publishing.nsf/Content/17BF043A41D470A9CA257E13000C9322/\\$File/FINAL-Review-of-PCEHR-December-2013.pdf](http://health.gov.au/internet/main/publishing.nsf/Content/17BF043A41D470A9CA257E13000C9322/$File/FINAL-Review-of-PCEHR-December-2013.pdf).
- 40 Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH: Big data: The next frontier for innovation, competition, and productivity. USA: McKinsey Global Institute; 2011.
- 41 Hood L, Auffray C: Participatory medicine: a driving force for revolutionizing healthcare. *Genome Med* 2013;5:1–4.
- 42 Andreu-Perez J, Poon CCY, Merrifield RD, Wong STC, Yang G-Z: Big Data for Health. *IEEE J Biomed Heal Informatics* 2015;19:2168–2194.
- 43 Zarate OA, Brody JG, Brown P, Ramirez-Andreotta MD, Perovich L, Matz J: Balancing Benefits and Risks of Immortal Data – Participant’s Views of Open Consent in the Personal Genome Project. *Hastings Cent Rep* 2016;46:36–45.
- 44 Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y: Identifying Personal Genomes by Surname Inference. *Science* 2013;339:321–324.
- 45 Kuo AM-H: Opportunities and Challenges of Cloud Computing to Improve Health Care Services. *J Med Internet Res* 2011;13:e67.
- 46 Kish LJ, Topol EJ: Unpatients – why patients should own their medical data. *Nat Biotechnol* 2015;33:921–924.
- 47 Schatsky D, Muraskin C: Beyond bitcoin – Blockchain is coming to disrupt your industry. (Internet) 2015 Dec. Available from: <http://dupress.com/articles/trends-blockchain-bitcoin-security-transparency/>.
- 48 Yasnoff WA, Shortliffe EH: Lessons Learned from a Health Record Bank Start-up. *Methods Inf Med* 2014;53:66–72.
- 49 World Health Organization. Chronic diseases are the major cause of death and disability worldwide. (Internet) 2005, p 2. Available from: http://www.who.int/chp/chronic_disease_report/media/Factsheet1.pdf.
- 50 World Health Organization: Preventing Chronic Diseases a vital investment. (Internet) Canada; Available from: http://www.who.int/chp/chronic_disease_report/contents/part2.pdf.
- 51 Centers for Disease Control and Prevention: Chronic Diseases: The Leading Causes of Death and Disability in the United States. (Internet) 2016 [cited 2016 Apr 12]. Available from: <http://www.cdc.gov/chronicdisease/overview/>.
- 52 Pagliari C, Detmer D, Singleton P: Potential of electronic personal health records. *Br Med J* 2007;335:330–333.
- 53 Archer N, Fevrier-Thomas U, Lokker C, McKibbin KA, Straus SE: Personal health records: a scoping review. *J Am Med Informatics Assoc* 2011;18:515–522.
- 54 Douglas L, McGee D. Understanding Medical Tests and Test Results. (Internet) MSD Manual – Professional Version 2015 [cited 2015 Apr 17]. Available from: <http://www.msdmannuals.com/professional/special-subjects/clinical-decision-making-understanding-medical-tests-and-test-results>.
- 55 Lazer D, Kennedy R, King G, Vespignani A: The Parable of Google Flu: Traps in Big Data Analysis. *Science* 2014;343:1203–1204.
- 56 Nuffield Council on Bioethics: Online personal health records. (Internet) London; 2014. Available from: <http://nuffieldbioethics.org/report/personalised-healthcare-2/online-personal-health-records/>.
- 57 Presser L, Hruskova M, Rowbottom H, Kancir J: Care data and access to UK health records: patient privacy and public trust. *Technol Sci* 2015;1–31.
- 58 Blenner SR, Köllmer M, Rouse AJ, Daneshvar N, Williams C, Andrews LB: Privacy Policies of Android Diabetes Apps and Sharing of Health Information. *J Am Med Assoc* 2016; 315:1051–1052.
- 59 Ipsos MORI – Social Research Institute: The One-way Mirror: Public attitudes to commercial access to health data. (Internet) 2016 Mar. Available from: <https://www.ipsos-mori.com/Assets/Docs/Publications/sri-wellcome-trust-commercial-access-to-health-data.pdf>.