Opinion Paper

Josep Roca*, Akos Tenyi and Isaac Cano

Paradigm changes for diagnosis: using big data for prediction

<https://doi.org/10.1515/cclm-2018-0971>

Received September 3, 2018; accepted November 21, 2018; previously published online December 11, 2018

Abstract: Due to profound changes occurring in biomedical knowledge and in health systems worldwide, an entirely new health and social care scenario is emerging. Moreover, the enormous technological potential developed over the last years is increasingly influencing life sciences and driving changes toward personalized medicine and value-based healthcare. However, the current slow progression of adoption, limiting the generation of healthcare efficiencies through technological innovation, can be realistically overcome by fostering convergence between a systems medicine approach and the principles governing Integrated Care. Implicit with this strategy is the multidisciplinary active collaboration of all stakeholders involved in the change, namely: citizens, professionals with different profiles, academia, policy makers, industry and payers. The article describes the key building blocks of an open and collaborative hub currently being developed in Catalonia (Spain) aiming at generation, deployment and evaluation of a personalized medicine program addressing highly prevalent chronic conditions that often show co-occurrence, namely: cardiovascular disorders, chronic obstructive pulmonary disease, type 2 diabetes mellitus; metabolic syndrome and associated mental disturbances (anxiety-depression and altered behavioral patterns leading to unhealthy life styles).

Keywords: big data; Integrated Care; personalized medicine; systems medicine; value-based healthcare.

Three main forces driving paradigm change in health care

Since the beginning of the 21st century, medicine has been experiencing profound qualitative changes that are driven by the convergence of three main forces that are imposing a paradigm change. Firstly, advances in biomedical knowledge, with changes in disease understanding and in prevention and treatment in the post-genomic era, have shaped the methodologies associated to systems (precision) medicine [1, 2]. Moreover, functionalities of health systems worldwide have been challenged by the epidemics of chronic conditions that are shaping an entirely new health and social care scenario [3, 4]. Last, but not least, the enormous technological potential developed over the last years is increasingly influencing life sciences and driving changes toward personalized medicine and value-based healthcare [1, 2].

Systems (precision) medicine

Since the early 2000s, two key phenomena are prompting substantial changes in both biomedical research and clinical management of chronic patients. Firstly, systems biology methodologies (i.e. ‘omics’ technologies, use of computational modeling, etc.), taking a holistic approach for solving biomedical challenges, are being progressively embedded into medical practice shaping the practicalities of systems medicine [1, 2]. This new field promises a novel approach to disease, shifting its definition from phenotypical signs and symptoms towards molecular subtypes (i.e. endotypes) of diseases, which is indispensable for precise characterization of disease relations and for the evaluation of shared mechanisms [5–9].

Simultaneously, digital health initiatives and wearable devices are facilitating access to an enormous amount of patient-related information from whole populations to personal levels, and state-of-the-art computational models and machine learning tools demonstrate a high potential for health prediction [1, 2]. Given the extremely long and

*Corresponding author: Josep Roca, MD, PhD, Hospital Clínic, IDIBAPS, Facultat de Medicina, Universitat de Barcelona, C/Villarroel 170, 08036, Barcelona, Catalunya, Spain; and Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Av. Monforte de Lemos, 3-5. Pabellón 11. Planta 0, 28029, Madrid, Catalunya, Spain, Phone: +34-932275747, Fax: +34-932275455, E-mail: jroca@clinic.cat

Akos Tenyi and Isaac Cano: Hospital Clínic, IDIBAPS, Facultat de Medicina, Universitat de Barcelona, Barcelona, Catalunya, Spain; and Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Madrid, Catalunya, Spain

expensive bench to clinics cycles of the biomedical sector, these technologies promise a fast-track approach where scientific evidence can support clinical care while simultaneously collected insights from daily clinical practice promote new scientific discoveries and optimize healthcare optimization.

The application of systems biology tools to analyze the underlying molecular mechanisms of specific diseases, as well as the complexities of disease co-occurrence can contribute to explain patients' heterogeneities. The proposed approach should lead to the generation of novel biomedical knowledge for the enhancement of patient stratification while exploring applicable strategies for improved patient management and the decrease of disease burden.

Integrated Care

Healthcare systems worldwide are undergoing profound changes to adapt to novel patient-oriented efficient services necessary to tackle the increasing epidemic of non-communicable diseases (NCDs) and their risk factors caused by unhealthy lifestyles and population aging. Facing this challenging situation implies a profound reshaping in the way we approach delivery of care, as well as cultural changes both at citizen and health professional levels. Patient empowerment is an essential part of the healthy lifestyles interventions and the prevention and management of NCDs. The roles of patients and healthcare professionals are evolving to place patients at the center of the management as an active partner of the healthcare process, both as co-designer of the interventions and manager of her/his condition [10].

The World Health Organization (WHO) [11] defines Integrated Care as “the organization and management of health services so that people get the care they need, when they need it, in ways that are user friendly, achieve the desired results and provide value for money system”. Despite known uncertainties in the characteristics of on-going deployment processes, the new models of healthcare delivery are prompted toward large-scale implementation. The term scaling-up is defined by WHO as “deliberate efforts to increase the impact of health service innovations successfully tested in pilot or experimental projects so as to benefit more people and to foster policy and program development on a lasting basis” [12]. It is acknowledged that there is a need for health-value generation of innovative information and communication technologies (ICT)-supported healthcare services. The final aim is to enhance health outcomes in a cost-effective manner such that it allows containment of overall health

costs, with a more proactive experience for patients and healthcare professionals.

The challenge imposed by the epidemic of chronic disorders constitutes one of the main driving forces that are prompting an extensive European deployment of Integrated Care services supported by ICT through initiatives like the European Innovation Partnership for Healthy and Active Aging (EIP-AHA) [13] and the European Institute of Technology-Health (EIT-Health) [14].

Digital health tools

It is acknowledged the role of ICT enabling enhanced biomedical knowledge and facilitating the change in healthcare, as already alluded to above. However, the current slow progression of adoption, limiting generation of healthcare efficiencies through technological innovation, can be realistically overcome by fostering convergence between a systems medicine approach and the principles governing Integrated Care. Implicit with this strategy is the multidisciplinary active collaboration of all stakeholders involved in the change, namely: citizens, professionals with different profiles, academia, policy makers, industry and payers.

The emerging health scenario should facilitate adoption of technologies capable of longitudinally measuring multilevel health parameters at high resolution that will be coupled with dynamic knowledge repositories and sophisticated analytics feeding decision support systems addressed to citizens (patient decision support systems, PDSS) and/or professionals (clinical decision support systems, CDSS).

A long journey toward a novel health scenario

The Catalan Reference Site, endorsed by EIP-AHA, is facilitating convergence among: (i) initiatives associated with the regional Health Plan 2016–2020; (ii) resources from main healthcare providers; and, (iii) input from different competitive grants, favoring an ecosystem that ensures transition from previous pilot experiences [15] to large-scale deployment and adoption of on-going activities showing health value generation (Figure 1).

As a successful example of this transition from pilot experiences showing cost-efficacy to implementation in real world settings assessing health-value generation, we would like to introduce general characteristics of the on-going Nextcare program [16]. Nextcare stands for Innovation in Integrated Care Services for Chronic Patients. It

Figure 1: In the new health scenario, innovative ICT-supported health services will deliver cost-effective care for chronic patients based on systems (precision) medicine approach with a preventive orientation. It will imply enhanced multilevel health risk predictive modeling feeding patient decision support systems (PDSS) and clinical decision support systems (CDSS). Cloud-based computing architectures for data analytics will be a need for dynamic assessments while complying with privacy and regulatory issues. All in all, the novel scenario requires building-up efficient innovation ecosystems involving all the relevant stakeholders contributing to convergence between systems (precision) medicine and Integrated Care as the strategy required to shape the new health paradigm.

is a RIS3 (Research and Innovation Strategies for Smart Specializations) program (2016–2019) that constitutes the umbrella for all the initiatives included in the current research. It has three main objectives: (i) Regional deployment of Integrated Care services for chronic patients with a personalized medicine approach; (ii) Development of a test-bed, willing for international leadership, for the use of ICT in novel services that generate value in the healthcare system of Catalonia; and, (iii) Development and monetization of novel products and services with a high level

of transferability to other healthcare systems, contributing to strengthen Catalan industrial competences.

The project addresses five actions (Figure 2) that encompass the main challenges encountered during the deployment of Integrated Care services. The main objective of *Action 1* is Health Risk Assessment and Service Selection [17] with a special focus on non-pulmonary manifestations and co-morbidities seen in patients with chronic obstructive pulmonary disease were analyzed as a use case [18, 19]. *Action 2* aims at promoting healthy lifestyles with a focus on physical activity [20, 21]. *Action 3* deploys community-based management of complex chronic patients (CCP) [22]. *Action 4* addresses regional deployment of transfer of diagnostic testing to primary care focusing on forced spirometry as a use case [23]. Finally, *Action 5* promotes interoperability between healthcare, informal care and biomedical research shaping the so-called digital health framework (DHF) (Figure 3) [24, 25], as a technological facilitator that supports collaborative adaptive case management among health professionals and with patients.

To this end, coordination between the healthcare system and the informal care environment is a central component of the five actions aiming at achieving an efficient management of multimorbidity. The regional shared electronic health record (HC³) and the personal health folder (La Meva Salut) are the backbone components of the current technological support to the services. NEXT-CARE is transforming the current personal health folder into a citizens' self-management tool and developing technological support to adaptive case management of clinical

Figure 2: Schematic representation of the five actions, A1 to A5, addressed by NEXTCARE aiming at regional deployment of Integrated Care services for chronic patients.

Figure 3: Key dimensions of a Digital Health Framework for enhancing communication among Informal Care (top layer), Health Care (central layer) and Biomedical Research (bottom layer), as first described in ref. [24]. Informal care refers to all factors, traditionally not included in health information systems with an impact on health, as depicted in the figure, including patients' self-tracking data. Moreover, recent studies [19] strongly support the use of population-health information (registry data) to enrich multilevel health risk prediction for clinical purposes, as is extensively explained here.

processes to support collaborative work among the actors. The articulation of the protocols described in the main text provides the frame for a comprehensive evaluation of regional deployment of Integrated Care.

The ultimate aim of Nextcare is to set the basis for large-scale deployment of innovative ICT-supported healthcare services with proven high potential for transferability to different scenarios and geographical areas. It is hypothesized that the outcomes of the program shall contribute to adoption of cost-effective services aiming for high impact on prevention of multi-morbidity, as well as the progression of chronic patients toward the tip of the population-health risk stratification pyramid. Consequently, the implicit general hypothesis of the ongoing actions is that these novel services can effectively facilitate decrease in both the health and societal burden of chronic conditions.

Four challenges

The process of deployment of ICT-enabled services within the framework of Nextcare has facilitated identification of four major areas requiring attention in order to achieve a successful transition toward adoption of systems medicine for chronic patients, within an Integrated Care scenario, as described in the following subheadings.

Enhanced clinical predictive modeling

Health risk assessment and service selection are widely accepted tools facilitating large-scale adoption of

Integrated Care of chronic patients and contributing to enhance healthcare outcomes and patient experience of care while reducing costs and improving the health of populations. The wealth of data available may allow to consider several avenues of computational modeling, including artificial intelligence and deep learning or machine learning tools, based on knowledge and expertise of the data analytics team. However, their application may face major limitations when it comes to accessing and mining health data currently stored in distributed silos of information.

To overcome these major limitations, there is a clear need for application of holistic strategies for subject-specific risk prediction and stratification that incorporate multi-level determinants of health (e.g. socio-economical, lifestyle, behavioral, clinical, physiological, cellular and “omics” information [17] into risk models, which would substantially increase their predictive accuracy and their use in clinical decision-making (Figure 4). To this end, the open data trends of biomedical research should be similarly extended to clinical practice by solving privacy and regulatory constraints. Articulation of these strategies with systems-oriented biomedical research would provide continuous cross-fertilization between research and patient care [2].

An ideal healthcare setting (Figures 1 and 4) should facilitate an optimal support to care decisions and delivery by reducing the complexity of the massive amount of multi-disciplinary data being produced every day and should improve efficiency of health outcomes both in terms of well-being and expenditures. Such a health system relies on the availability of health-related data and analytics tools. Enhanced clinical predictive modeling and personalized diagnostic and treatment tools, such

Figure 4: Scheme for dynamic enhancement of clinical predictive modeling (CPM) feeding clinical DSS (CDSS) and/or patient DSS (PDSS) in cloud-based environments.

as CDSS and PDSS, can contribute to the acceleration of transfer of scientific evidence to practice, helping in the identification of gaps in care and in targeting interventions to the most appropriate sub-populations of patients.

Development of enhanced CPM to feed DSS will require consideration, and eventual integration, of computational modeling of four different dimensions: (i) Underlying biological mechanisms; (ii) Current evidence-based clinical knowledge; (iii) Patients' self-tracked data, including sensors, behavioral, environmental and social information; and, (iv) Population-based health risk assessment data, as already depicted in Figure 3 and described in the text. CDSS/PDSS should be designed to interoperate with existing centralized or distributed hospital information systems, and ultimately with learning health systems (i.e. systems in which science, informatics, incentives and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the delivery process and new knowledge captured as an integral by-product of the delivery experience).

Nextcare is successfully developing a systems medicine approach to better understand patients with chronic conditions with a focus on chronic obstructive pulmonary disease (COPD) as a use case. The overall strategy and specificities of the program, as well as key reported achievements, have been recently analyzed in detail in [18]. Briefly, the novel approach has contributed to uncover shared underlying mechanisms of non-pulmonary phenomena, such as skeletal muscle dysfunction and common comorbid conditions, observed in patients with COPD. The

results pave the way toward enhanced health risk assessment of these patients with relevant implications on clinical management of chronic conditions.

Technology: cloud-based architecture and services

Holistic implementation strategies of cloud computing environments, tackling privacy and regulatory constraints, have been identified as main enablers of this ideal healthcare setting [25, 26]. Currently, the articulation of the main technical building blocks, i.e. multi-level biomedical data integration, tools for clinical predictive modeling in the cloud and high-performance computing, as one integrated system is as yet, an unmet potential. This is strongly highlighted by the exponential growth of health-related data that could enhance current health information systems with functionalities to assist clinical decision-making of health professionals, as well as of patients for self-management. End-to-end exploitation of cloud infrastructures for large-scale data analytics has so far been held back by the lack of a well-integrated set of reliable and flexible services, and user-friendly interfaces. This problem is particularly true for medical research and clinical applications where the additional complexity of handling personal information requires particular care. In this context, a cloud-based data analytics platform shall unlock the full potential of clinical predictive modeling by solving issues around integration, harmonization and

privacy of data coming from different sources in an integrated manner, and by providing a general interface for developing and deploying predictive models.

The added benefit of such a solution includes rapid local prototyping, cost-effective parameter sweeping and validation scale-out to high-performance, large-scale modeling. It should also enable the deployment of the same services across different infrastructures, institutes, laboratories or projects according to local policies, while maintaining interoperability and consistency at the platform layer. Specific implementation of such a platform should identify and deploy cloud-based services for private and public uses with a design that is versatile, scalable, trusted and abstract enough to support a wide range of clinical predictive modeling approaches and application in broad operational contexts for digital health and daily clinical practice.

A high-level description of a proposed cloud-based data analytics platform is displayed in Figure 5 indicating the four types of data sources that are considered for multi-disciplinary computational modeling. Computational modeling in combining these four dimensions should provide the basis for enhanced clinical predictive modeling and elaboration of a cloud-based CDSS/PDSS embedded into clinical processes.

In the implementation process of such a platform, specific requirements include (i) development of a data interoperability engine for standardized data exchange and semantic interoperability among multi-disciplinary data sources; (ii) deployment of technologies to handle

personal data in compliance with agreed legal, policy and standardization requirements; and (iii) defining standard interfaces in compliance with user needs.

Data interoperability engine

A critical aspect of the proposed platform is the development of an efficient data interoperability engine compliant with European regulations and FAIR data principles: (i) Findable: data must be easy to find by both humans and computer systems; (ii) Accessible: data must be put in long-term storage in such a way that either the data itself or its metadata can be accessed easily; (iii) Interoperable: datasets can be combined by humans as well as computer systems, in which the use of shared vocabularies and/or ontologies is of special importance; and, (iv) Re-usable: data can be used for future research and to be processed further by computer programs. To this end, several initiatives are available for consideration, e.g. the ELIXIR (www.elixir-europe.org) interoperability platform or commercial medical data federation tools such as FedEHR (www.fedehr.com).

Data security and data protection

Given the need to handle sensitive data, particular attention is needed to be paid to security and data protection aspects. Future data architectures should assume that

Figure 5: Conceptual view of the proposed cloud-based data analytics platform.

Properly harmonized multi-disciplinary data sources provide the basis for enhanced clinical predictive models that feed real-time decision support systems able to guide health professionals in the clinical decision-making process.

sensitive data cannot be moved around, but rather code and models have to be moved to the data. When considering such architectures three levels of data should be considered: (i) public data; (ii) anonymized data; (iii) pseudo-anonymized data. Public data includes public datasets that can be freely shared and accessed from well-known, community curated data banks and repositories. In this context, creation of public “data lakes” shared across the community and used, for example, to train the machine learning services and build reference models are needed. Anonymized data need to be defined in compliance with existing and future legal definitions and need to be accessed and contributed based on agreed policies of use within specific projects or communities. Such data, although anonymized, is richer than completely public data as it might convey a specific context of use, thus it is of high interest for research purposes. Finally, pseudo-anonymized data should be used to generate models for the CDSS/PDSS that can then be applied to patients in clinically relevant scenarios. In light of the new European General Data Protection Regulation (GDPR) new technologies are also needed to be considered to manage and audit access to data and to move in the direction of implementing the GDPR norms of data ownership and usage. Data transactions and “smart” user consents could be stored and managed using blockchain technologies to enable the implementation of transparent, end-to-end policies for data governance.

User-profiled interfaces

One of the main barriers for translational researchers and clinicians in accessing cloud-based services is currently the lack of intuitive, easy-to-use user interfaces. The main objective of the technical implementation should be to provide an intuitive way for healthcare researchers and practitioners to exploit the capabilities of distributed infrastructures, i.e. accessing a wide range of data sources and services without having to understand any of the complexity of the system. At least two levels of interfaces should be provided in a desired platform, which should be co-designed with the communities of use of reference:

- Data analysis interfaces for translational researchers: this type of interfaces provide an intuitive, scriptable, portable access to the platform services. Technologies like Jupyter Notebooks, CERN SWAN (Service for Web based ANalysis) and other similar approaches provide good examples of such environments. The possibility of supporting different types of scripting engines including familiar systems such as the R analysis

framework is critical to ease the transition from local computers to cloud-based services.

- Decision support interfaces for healthcare professionals and patients: this type of interface should provide clean, web-based access to specialized services and applications, hiding from the end-user (i.e. healthcare professional and patients) the complexity of the underlying infrastructure service layers.

To comply with the vision of the cloud-based data analytics platform, methodologies used in biological and clinical modeling should converge towards standard operations and tools that can be integrated into general pipelines and implemented in analysis platforms. These pipelines shall be ready to analyze data independent of their source or type, i.e. molecular, clinical or wearable measurements, and integrate them in an operational manner. This assumes pre-processed and well-formatted data input, as well as standardized outputs (Figure 6). When considering subject-specific health risk prediction and stratification as the desired output of such systems, the framework of machine learning defines these input and output needs as well as the identified challenges. In this regard, the main factors to consider should be the dimensionality of the data sources, i.e. the number of features that are used for health prediction, their sample size and differences in sample sizes when considering the integration of multiple datasets. Registry data, EHR data and wearable technologies come with the great promise to bring biomedical research to the Big Data era with population/subpopulation size data, whereas molecular data have great potential to gain biological insight into disease mechanisms, however, for these data sources population-wide availability is still awaited. The integration of such data sources should enable mining health related patterns from data with state-of-the art technologies, such as deep learning that show exciting potential for identifying non-linear patterns from large amounts of raw biomedical data [27–29]. The major potential of this technology is that it promises a universal approximator for many learning and prediction tasks that could substitute several processes that are currently done separately in biomedical and machine learning fields. A fascinating way of using deep learning could help to select biologically important features, organize them into higher abstraction level biological assemblies (e.g. pathways, disease modules), highlight their role in the disease and also to predict disease risk using them [27]. A major obstacle, however, is that they are often associated with the need for large longitudinal sample sizes, which is a barrier especially in molecular data sources. Furthermore, more research is needed in this field to abolish its

Figure 6: Proposed interoperability architecture.

The data sourcing layer is responsible for integrating cloud-based services to required structured and unstructured data sources at site-level. The extract-transform-load (ETL) layer is where data is extracted from homogeneous or heterogeneous data sources and transformed for storing in the proper format or structure for the purposes of querying and analysis the target data lake or data warehouse. Finally, the application layer provides access to common platform services for data analytics whereas the pre-processing layer is responsible for integrating DSS with site-specific clinical workstations and patient gateways (e.g. La Meva Salut in Catalonia, Spain).

current stigmatization as a “black box” approach, which is often seen as a barrier for clinical application.

Moreover, integration of data from different sources and with different formats is a major challenge of *in-silico* modeling pipelines. Current practice shows that in different fields different models evolve, such as the disease maps, disease trajectories, mechanistic models, other multiscale hybrid modeling already combining some of the previous approaches; or, data-driven approaches using machine learning. In order not to waste the field specific knowledge encapsulated in these models, integrative approaches are needed. In this context, the patient similarity framework shows immense potential to be applied in the clinical field, mainly due its high abstraction level leading to broad applicability, its patient-centered approach and its transparent methodology, which is especially important for acceptability from the clinical side [30, 31]. Patient similarity enables the separate comparison of patients on different biological organizational levels, e.g. using molecular profiles (transcriptome, genome, epigenome, etc.), clinical traits, comorbidities, and allows retrieving groups of similar patients, or the most successful treatments based on similar cases, as

well as to predict health risk in an unsupervised manner [32–35].

Evaluation and adoption of decision support systems (DSS)

For successful adoption in real world settings, interoperability of the proposed cloud-based data analytics platform with healthcare information systems is indispensable (Figure 6). On the one hand, cloud-based services should be integrated at the site level with the required structured and unstructured data sources.

On the other hand, information systems departments of clinical sites should take into account in-place health information exchange infrastructures, where standard terminology (e.g. SNOMED-CT, SERAM, SEMN, LOINC, etc.), message encoding (e.g. HL7 2.x/3.x, MLHIM, openEHR, ISO 13606, etc.), message routing and security (e.g. IPSec, Audit trail, Node authentication, etc.) are of special importance. Where available, existing controlled vocabularies such as the ICD-10 or the FMA human anatomy, standards for data and metadata format (e.g. ISA-TAB) and content

(e.g. MSI or MIAME) should be used. Where standards are currently not yet broadly accepted, agreements should be generated to deploy site-level interoperability middleware based on an HL7 FHIR standard specification (e.g. HAPI FHIR, hapifhir.io).

Moreover, successful adoption in real world settings will likely require specific evaluation designs to assess effectiveness and health value generation of complex interventions, ranging from standard randomized controlled trials to different implementation science designs [36, 37]. The need for looking for complementarities among different types of study designs has been recently pointed out [38].

Barriers and facilitators for deployment and adoption

Governance of a cloud-based system based on FAIR data principles generates several major challenges in terms of: (i) data/services administration accessibility; (ii) continuous control of quality assurance programs; (iii) compliance with ethical and regulatory issues; as well as, (iv) sustainability of the approach over time. These management and governance challenges are expected to be overcome by adopting block chain technology that allow complete traceability of transactions, and most importantly finely granular enforcement of rules in observance of regulation at data origins, and of consent design. All actors of cloud-based systems should be able to verify that sharing, analysis and other handling and use of the information is in accordance of the individual's intentions and applicable laws, regulations and processes.

Regulatory aspects

Several ethical and regulatory issues currently fall into gray areas in terms of regulation (i.e. computational modeling and DSS assessment for medical use). In this regard, the generation of recommendations on gray areas should be addressed in future work, with the specific objectives to be covered:

- To ensure the protection of privacy and compliance with the GDPR (EU) 2016/679, Directive 95/46/EC and ISO27001 conformance for secure storage of pseudonymized multi-disciplinary data, which will improve trust in health research and therefore will facilitate adoption of innovative digital health services;

- To advise on applicable legislation and regulations to which innovative mathematical models, and components using them, need to comply before they can be deployed and used in healthcare settings;
- To ensure the ethical use of innovative models within patient and professional decision-making.

Service adoption

Adoption and both organizational (e.g. data processing agreements, liability/responsibility aspects, etc.) and financial sustainability (e.g. entrepreneurial actions) of cloud-based services constitutes a major challenge, wherein service models such as those developed in projects like MyHealthMyData (ICT-18-2016-732907) could be considered. The development of a roadmap for large scale deployment and adoption of cloud services at national and EU level should be of main interest for future initiatives pursuing adoption of novel cloud-based services.

Conclusions

The cloud-based data analytics platform has been proposed to successfully address the implicated potentials of health risk assessment and stratification and to facilitate large-scale adoption of Integrated Care of chronic patients [17, 39], contributing to enhance healthcare outcomes and patient experience of care while reducing costs and improving the health of populations. Applying holistic strategies for subject-specific risk prediction and stratification, that consider multilevel covariates influencing patient health, would increase the predictive accuracy and facilitate clinical decision-making based on sound estimates of individual prognosis.

Acknowledgments: Supported, in part, by NEXTCARE (COMRDI15-1-0016), PITES-TIiSS (PI15/00576) and AGAUR research groups (2009SGR911 and 2014SGR661).

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Torkamani A, Andersen KG, Steinhubl SR, Topol EJ. High-definition medicine. *Cell* 2017;170:828–43.
2. Maddox TM, Albert NM, Borden WB, Curtis LH, Ferguson TB, Kao DP, et al. The learning healthcare system and cardiovascular care: a scientific statement from the American heart association. *Circulation* 2017;135:e826–57.
3. Murray CJ, Lopez AD. Measuring the global burden of disease. *N Engl J Med* 2013;369:448–57.
4. Blumenthal D, Chernof B, Fulmer T, Lumpkin J, Selberg J. Caring for high-need, high-cost patients – an urgent priority. *N Engl J Med* 2016;375:909–11.
5. Hidalgo CA, Blumm N, Barabási AL, Christakis NA. A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 2009;5:e1000353.
6. Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, et al. Uncovering disease-disease relationships through the incomplete interactome. *Science* 2015;347:1257601.
7. Jensen AB, Moseley PL, Oprea TI, Ellesoe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nat Commun* 2014;5:4022.
8. Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. The human disease network. *Proc Natl Acad Sci USA* 2007;104:8685–90.
9. Lee D-S, Park J, Kay KA, Christakis NA, Oltvai ZN, Barabási A-L. The implications of human metabolic network topology for disease comorbidity. *Proc Natl Acad Sci USA* 2008;105:9880–5.
10. ACT@Scale project. Advancing Care Coordination and Telehealth at Scale. 2016–2019. <https://www.act-at-scale.eu/>.
11. World Health Organization. Integrated health services: what and why? Switzerland: World Health Organization, 2008.
12. World Health Organization. Practical guidance for scaling up health service innovations. 2009. http://apps.who.int/iris/bitstream/10665/44180/1/9789241598521_eng.pdf.
13. Bousquet J, Farrell J, Crooks G, Hellings P, Bel EH, Bewick M, et al. Scaling up strategies of the chronic respiratory disease programme of the European Innovation Partnership on Active and Healthy Ageing (Action Plan B3: Area 5). *Clin Transl Allergy* 2016;6:29.
14. EIT-Health. European Institute of Innovation and Technology in Health. <https://www.eithealth.eu/>.
15. Hernández C, Alonso A, García-Aymerich J, Grimsø A, Vontetsianos T, García Cuyàs F, et al. Integrated care services: lessons learned from the deployment of the NEXES project. *Int J Integr Care* 2015;15:e006.
16. NEXTCARE program Innovation in Integrated Care Services for Chronic Patients, COMRD15-1-0016 2016. <http://www.next-carecat.cat>.
17. Espieén ID, Vela E, Pauws S, Bescos C, Cano I, Cleries M, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. *Br Med J* 2016;6:e010301.
18. Tenyi A. A Systems Medicine approach to multimorbidity: towards personalised care for patients with COPD. PhD Thesis, 2018. <http://hdl.handle.net/2445/124046>.
19. Vela E, Tényi Á, Cano I, Monterde D, Cleries M, García-Altes A, et al. Population-based analysis of patients with COPD in Catalonia: a cohort study with implications for clinical management. *BMJ* 2018;8:e017283.
20. Barberan-García A, Gimeno-Santos E, Blanco I, Cano I, Martínez-Pallí G, Burgos F, et al. Protocol for regional implementation of collaborative self-management services to promote physical activity. *BMC Health Serv Res* 2018;18:560.
21. Barberan-García A, Gimeno-Santos E, Blanco I, Cano I, Martínez-Pallí G, Burgos F, et al. Personalised prehabilitation in high-risk patients undergoing elective major abdominal surgery: a randomized blinded controlled trial. *Ann Surg* 2018;267:50–6.
22. Cano I, Dueñas-Espín I, Hernandez C, de Batlle J, Benavent J, Contel JC, et al. Protocol for regional implementation of community-based collaborative management of complex chronic patients. *NPJ Prim Care Respir Med* 2017;27:44.
23. Vargas C, Burgos F, Cano I, Blanco I, Caminal P, Escarabill J, et al. Protocol for regional implementation of collaborative lung function testing. *NPJ Prim care Respir Med* 2016;26:16024.
24. Cano I, Lluch-Ariet M, Gomez-Cabrero D, Maier D, Kalko S, Cascante M, et al. Biomedical research in a digital health framework. *J Transl Med* 2014;12:S10.
25. Cano I, Tenyi A, Vela E, Miralles F, Roca J. Perspectives on Big Data applications of health information. *Curr Opin Syst Biol* 2017;3:36–42.
26. Directorate-General for Health and Food Safety (European Commission). Study on big data in public health, telemedicine and healthcare. 2016. https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata_report_en.pdf.
27. Alkawaa FM, Chaudhary K, Garmire LX. Deep learning accurately predicts estrogen receptor status in breast cancer metabolomics data. *J Proteome Res* 2018;17:337–47.
28. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016;6:26094.
29. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018;1:18.
30. Brown S-A. Patient similarity: emerging concepts in systems and precision medicine. *Front Physiol* 2016;7:561.
31. Gallego B, Walter SR, Day RO, Dunn AG, Sivaraman V, Shah N, et al. Bringing cohort studies to the bedside: framework for a ‘green button’ to support clinical decision-making. *J Comp Eff Res* 2015;4:191–7.
32. Lee J, Maslove DM, Dubin JA. Personalized mortality prediction driven by electronic medical data and a patient similarity metric. *PLoS One* 2015;10:e0127428.
33. Ng K, Sun J, Hu J, Wang F. Personalized predictive modeling and risk factor identification using patient similarity. *AMIA Jt Summits Transl Sci* 2015;2015:132–6.
34. Panahiazar M, Taslimitehrani V, Pereira NL, Pathak J. Using EHRs for heart failure therapy recommendation using multidimensional patient similarity analytics. *Stud Health Technol Inform* 2015;210:369–73.
35. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014;11:333–7.
36. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. Standards for Reporting Implementation

- Studies (StaRI): explanation and elaboration document. *BMJ* 2017;7:e013318.
37. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. *Br Med J* 2013;347:2–7.
38. Gershon AS, Jafarzadeh SR, Wilson KC, Walkey AJ. Clinical knowledge from observational studies. Everything you wanted to know but were afraid to ask. *Am J Respir Clin Care Med* 2018;198:859–67.
39. Cano I, Alonso A, Hernandez C, Burgos F, Barberan-Garcia A, Roldan J, et al. An adaptive case management system to support integrated care services: lessons learned from the NEXES project. *J Biomed Inform* 2015;55:11–22.
-
- Article note:** Lecture given by Prof. Josep Roca at the 2nd EFLM Strategic Conference, 18–19 June 2018 in Mannheim (Germany) (<https://elearning.eflm.eu/course/view.php?id=38>).