



Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/hlpt



European data-driven economy: A lighthouse initiative on Personalised Medicine[☆]



Denis Horgan^{a,*}, Mário Romão^b, Richard Torbett^c, Angela Brand^d

^aEuropean Alliance Personal Medicines (EAPM), Europe

^bIntel Corporation, SA, Europe

^cChief Economist, European Federation of Pharmaceutical Industries and Associations, Europe

^dInstitute for Public Health Genomics, GROW, Maastricht University, The Netherlands

Available online 23 October 2014

Big data for personalised medicine

“By 2020, the EU should endeavour to achieve widespread benefits for patients and citizens from personalised healthcare by defining in 2015, and subsequently executing a Data Strategy for Personalised Medicine” (EAPM, 2014)

In respect of Personalised Medicine, *Big Data* represents the vast and continuously growing amount of health information (including biomedical and environmental) and its usage to drive innovation in translational research and health outcomes tailored to the individual. Using these data to first understand the cause of disease, the medical profession can then develop new drugs and therapies to find the cure, as well as other health interventions targeting the individual. The personalised, individual approach requires advanced technologies and processes to collect, manage and analyse the information and, even more importantly, to contextualise it, integrate it, interpret it and provide rapid and precise decision support in a clinical and public health context.

[☆]**About EAPM:** The European Alliance for Personalised Medicine brings together Europe's leading healthcare experts and patient advocates to improve patient care by accelerating the development, delivery and uptake of personalised medicine and diagnostics. It is calling for the European Commission, the European Parliament and EU member states to help improve the regulatory environment so that patients can have early access to personalised medicine, and so that research is boosted.

*Corresponding author.

E-mail address: denishorgan@euapm.eu (D. Horgan).

Not only does Big Data offer the potential to revolutionise the effectiveness of health interventions, it may also help ensure the more effective management of resources in what are increasingly cash-strapped public healthcare systems. Over the coming decades, the financial sustainability of health systems will become more and more challenging as the population ages. The number of over 65s in Europe will increase by 75%¹ by 2060. Alongside this ageing population there is likely to be an associated rise in chronic illness which will lead to spending on health and social care reaching unsustainable levels unless we are able both to increase the quality of health outcomes and the efficiency of healthcare resources. Big Data, in theory, offers the potential to do both. It is widely acknowledged that ‘value-based’ approaches to the management of care are an ideal way forward. Big Data will be a key enabler of this. And in future, physicians and health managers should have real-time, real-world evidence on what works and what does not for each patient. Moreover, other trends, such as mHealth, will bring the benefits of Big Data much closer to the citizen. That should lead to more informed individuals, and more rational and less wasteful decision-making.

Getting a *Data Strategy for Personalised Medicine* right in Europe would yield multiple benefits. Not only would it accelerate the development of more effective treatments and potentially help with the management of healthcare resources as described above, it would also act as a foundation for private sector investment and jobs in R&D in Europe. Global developments in approaches to Big Data in healthcare are of major importance to the future of several industries

¹European Commission (2012) “The 2012 Ageing Report”.

including startups and SMEs on ICTs, pharmaceuticals, medical devices and others. A coherent strategy for Big Data would, for example, have a direct effect on the attractiveness of a given health system for the placement of clinical trials. The European Union should see Big Data as a strategic investment that could drive industrial competitiveness.

On the path to personalised care

Nobody likes getting sick. But when we do, it is vital that our doctors have access to the best information and diagnosis techniques available. Thankfully, emerging technologies such as analytics tools for big data can help healthcare professionals improve diagnoses and reshape the way medicine is practiced.

People are also becoming more aware of and receptive to the powerful impact that big data could have on their lives. A recent global study² found that most people are optimistic about technology innovations advancing healthcare and are willing to participate in virtual healthcare visits with their doctor. The survey also found they would be open to using health sensors in their bodies and throughout their daily lives. From mHealth technologies to remote monitoring and sensor systems, new technological innovations are key to the future of a more personalised approach to healthcare.

mHealth and wearables

The European Commission has recently issued a Green Paper on mHealth³ that points to its potential for improving prevention, efficiency, patient empowerment and economic development in healthcare.

Fitness and wellness enthusiasts today are pioneering technology and devices that are worn on or embedded into the body, such as wearable bracelets that combine technological innovation with fitness in new ways. With these types of devices people can better track their heart rate, perspiration and skin temperature to provide data and insights into how their daily routines are affecting their health and wellness.

In the near future, information may be obtained through different wearable or ingestible devices that may help predict changes in the body to prevent emergencies. The information may automatically be shared with healthcare teams for simpler management, as well as anonymously pooled with other people's data to help scientists and researchers find cures and to more rapidly develop effective medications.

Electronic health records

Organisations around the world are starting to realise the importance of the role of data on their efforts to improve the healthcare system. By investing in highly scalable data compute, storage, networking and software capabilities,

²http://newsroom.intel.com/community/intel_newsroom/blog/2013/12/09/the-world-agrees-technology-inspires-optimism-for-healthcare.

³COM(2014) 219 final, GREEN PAPER on mobile Health ("mHealth") <http://ec.europa.eu/digital-agenda/en/news/green-paper-mobile-health-mhealth>.

businesses are equipping themselves with the tools to improve care, discover new insights, reduce costs and meet emerging care models.

The European Union Directive 2011/24/EU on patients' rights in cross-border healthcare,⁴ inter alia, invites Member States to draw up guidelines on effective methods for enabling the use of medical information for public health and research.

The digitisation of health data provides the raw materials as there is no way to meaningfully use paper records. It would not all happen at once, but having electronic data in the long run will help improve patient safety, aid the discovery of new cures and treatments, and give healthcare providers access to a more-complete patient medical history to help with diagnosis and determine the right treatment (Shabo, 2013)⁵.

Computers are now fundamentally changing scientists' ability to track and use the process of trial and error. Today, instead of testing new drugs on thousands of patients to determine whether they are going to work, the pharmaceutical industry can use computer-simulated experiments based on huge amounts of both old and current data to more quickly deliver results. More theoretical tests can be done in less time, with lower cost and lower risk. This means that by the time drugs are tested, they are closer to being broadly usable, and the time required to distribute them to doctors and patients is reduced.

Individualisation and omics

Increasingly, people are embracing a future healthcare system that will allow them to get care beyond hospital walls, anonymously share their information for a healthier community, and enable better patient outcomes or improved personalised care that takes into account an individual's specific genetic makeup.

It is fair to say that the advances leading to the availability of biomedical data, very much driven by digitisation and the decreasing costs of human full genome sequencing (in 2014 a US company announced the \$1000 barrier had been conquered⁶) have outpaced Moore's Law,⁷ heralding a new era for healthcare comparable to that which computers did to transform society over the past decades. Equally important is the development and dissemination of tools and processes able to analyse and interpret the data, thus really creating new knowledge that can benefit patients accurately and directly, rather than at the endpoint of a lengthy process riddled by trial and error treatments and policy bottlenecks.

⁴<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:088:0045:0065:EN:PDF>.

⁵Shabo A. The patient-centric translational health record. *Pharmacogenomics*. 2013;14(4):349-52.

⁶<http://www.businessweek.com/articles/2014-01-14/illumina-dna-supercomputer-ushers-in-the-1-000-human-genome>.

⁷Moore's Law is the observation that the transistor count of integrated circuits, with respect to minimum cost, doubles every 24 months. In other words, it foresees the doubling of 'compute power', for the same cost, every two years. Today this period has actually been reduced to roughly 18 months.

The ability to cost-effectively sequence a whole genome and the value this would bring to a patient's care will soon bring genomic data into routine practice and its integration into electronic health records. Several countries have already embarked on government-sponsored genome sequencing programs, and in these projects genome data will be integrated with a patient's health record.⁸ By ensuring the right regulatory frameworks, researchers would potentially be able to access millions of genetic markers. In turn this would accelerate science towards better understating between diseases and specific patients. Crucially, this data will be more commonly leveraged directly in patient care, rather than research, to guide choice of therapy, prevention and screening programs, increasing overall healthcare efficiency and patient outcomes.

The potential, yet to be fully realised, for knowledge coming from these data to be used to improve medicine in many areas, is significant. Examples include the following:

- Increasing the understanding of the causes of disease by, amongst others, correlating data from vast patient populations to identify DNA variations and other factors such as environment and lifestyle that impact disease and influence treatment outcomes. This is of particular relevance in the area of rare diseases, where many patients (an estimated 50 million worldwide) are still left without a diagnosis - it takes these patients on average seven years to obtain a correct diagnosis.⁹
- Development of new drugs and therapies by speeding up the R&D pipeline using advanced computing biological modelling, clinical trials design and patient recruitment.
- Bringing genomics and their interaction with environmental factors to routine clinical and public health practice, thus enabling the identification of the most clinically and cost-effective treatments targeted to the specific patient, by matching and comparing the patient's DNA against known genomic variations and clinical data.
- Using genomics to track the evolution of disease (already prominent in cancer) during therapy, to develop tailored and adaptive therapy throughout the course of the disease, thus increasing overall survival rates and ultimately allowing more patients to eliminate the disease entirely. Deep sequencing identifies many mutations in each tumor thereby increasing the complexity of analysis and treatment.
- Further integration of exposure and environmental data with health data, to discern the extent to which risk of disease is affected by genomics interacting with environment. Both the US and Europe have now invested substantial amounts in Exposure and Health Surveys, and this is likely to lead to very comprehensive disease risk assessment in the next 10-20 years, allowing a more personalised disease risk calculation.

⁸<http://www.technologyreview.com/news/520571/why-the-uk-wants-a-genomic-national-health-service/>.

⁹<http://shire-hgt.isebox.net/file/%5B37458%5D%20RareDiseaseIm pactReportforWeb.pdf?download=>.

Perhaps the area that is most advanced is that of research geared towards understanding the mechanisms of disease and developing new drugs and treatments. However, translating these advances into daily clinical practice takes time and is still in its infancy. An example in the area of prevention is that which followed the discovery of the BRCA1 and BRCA2 genes mutations, which are indicators of a woman's risk of developing breast and/or ovarian cancer. The discovery allows early detection of women at high risk of breast cancer, prompting them to perform more frequent screening or to opt for mastectomy to eliminate the risk entirely. In the area of therapy molecular markers which emerged from early genomics studies allowed the development of novel therapies, such as, for example, Herceptin, a drug used in the case of breast cancers overexpressing the HER2 gene, which was an early pioneer of many subsequent drugs based on similar "personalised" principles.

Making big data work

Personalised Medicine requires computing environments able to process massive amounts of information for research and diagnostics, and at the same time, patient- and healthcare provider-friendly systems able to interface with patient medical records at the point of care. However, due to the massive amounts of data being created through working with patients' genomes, there are real technological challenges to overcome.

The volume, variety and velocity of healthcare data are increasing:

Volume - the wealth of data may range from clinical data from Electronic Health Records, genetic data, medical imaging and real-time monitoring devices, to environmental, nutritional, and even data from social media. Personal mobile devices are breaking ground at increasing pace in health and wellness areas, creating the potential to generate rapidly a vast increase in data in this area. Worldwide healthcare data in 2012 was estimated at 500 petabytes, which is equal to 50 billion four drawer filling cabinets.¹⁰ Processing large volumes of information, conducting complex analyses and rapidly generating new insights, requires hardware, networking, storage and analytics software that can deliver outstanding performance.

Variety - processing unstructured data (e.g. clinician's notes, images and videos, or scientific publications) alongside traditional structured data becomes necessary given the wide range of data sources. Moreover, globalisation of data sharing and integration is currently increasing, rather than decreasing variety, as the development of international standards and formats is much slower than the pace of innovation in the industry.

Velocity - achieving real time or near- real time analytics has been one of the main objectives in dealing with big data. Short discovery periods may be crucial for decision making in hospitals, whilst decreasing gene sequencing and analysis times can speed up the development of new targeted drugs, or patient genetic information-driven decision making can help doctors at the point of care.

Independently of the source of data, be it from DNA sequences, MRI scans, lifestyle information or scientific

¹⁰Bigger Data for Better Healthcare, IDC, 2013.

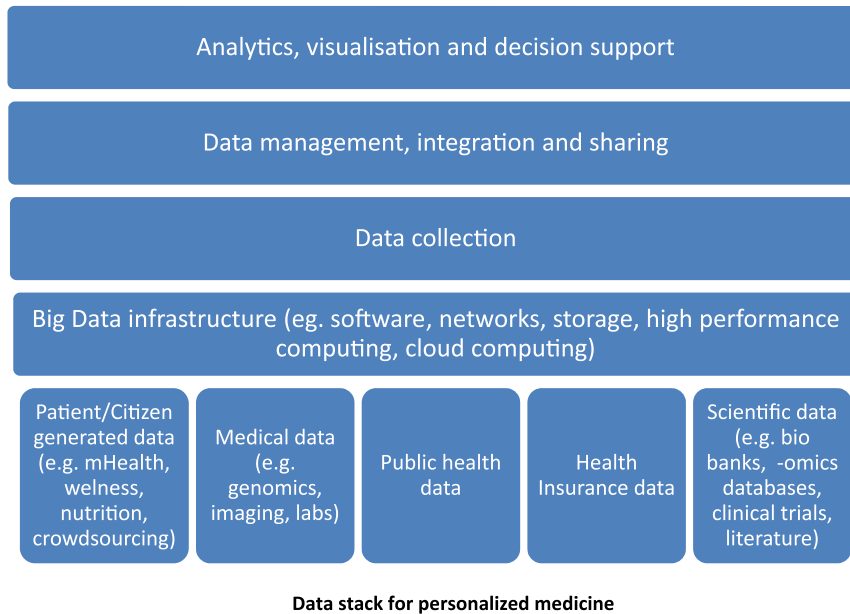


Figure 1 Data stack for personalized medicine.



The three sets of policy actions in respect of big data for personalized medicine

Figure 2 The three sets of policy actions in respect of big data for personalized medicine.

literature, these data have to become ‘information’ to aid new knowledge. This process, from data acquisition to interpretation, can be visualised in the figure below. The *Data stack for Personalised Medicine* is built from several layers supported

by a computing infrastructure that successively collects, organises and manages data, through integration and sharing, ending with an analytics layer where decision making, and visualisation are performed. Ultimately these will be fed back to the patients through better prevention, diagnoses, treatments and drugs; and to the healthcare systems through more efficient use of resources and informed policy decision making [Figs. 1 and 2](#).

In real settings, this stack may involve diverse configurations that may include vast computing resources, dedicated facilities, general purpose computers, “clouds of clouds” or top-of-the-range high performance computing machines. Currently, a single experiment involving genomic analytics may take from hours to several days, depending on the volume of data and computing power available.

The biggest challenges are making sense of the data and being able to access it when and where it is needed, whilst navigating complex regulatory frameworks. To realise these

benefits both new approaches and technologies are necessary. Broadly speaking, collecting, storing, accessing and processing (big) data from diverse locations and multiple sources and formats are more than just grand technical challenges.

As important as the technologies involved and cross-cutting the multiple layers of the Big Data stack are the policies and regulations that frame data collection, sharing and processing. These map legal and ethical obligations that should be implemented across the stack.

*Proteome-based cancer diagnosis*¹¹: Most diseases manifest at the level of protein activity. The entirety of all proteins is called the human proteome. With the right indicators at hand, it is possible to draw conclusions on patients’ current state of health by identifying correlations between changes in the proteome and diseases.

The goal is simple and yet ambitious: developing new, cheap and minimally invasive medical tests based on only a drop of blood.

¹¹See <http://www.sap-innovationcenter.com/2013/09/24/proteome-based-cancer-diagnostics-sap-hana/#>.

Finding these changes - or disease patterns - in the proteome at a very early stage is a big data problem. A patient's blood sample consists of more than 150 Mio data points. Once such a sample data set is processed and analysed, it needs to be compared to hundreds or even thousands of others in order to detect the disease patterns. Searching for complex patterns in protein activity is therefore computationally intensive and demands for the fastest and latest big data technology.

The opportunities are immense but the challenges are no less great. Big data in health consists almost exclusively of personal health data. The knowledge and value gains will come from the integration of diverse data sets (medical, health, fitness, nutritional, environmental) from millions of people, whilst assuring data protection and following subject's consent. There is an innate suspicion of sharing personal data among data experts and medical specialists as it challenges data privacy. The creation of a legal framework for personal data processing in scientific research as well as regulation that exploits technology to strengthen data security allowing reuse and secondary use of data is proposed as a path to explore.¹² But an alternative and radically different approach is to encourage data owners to actively participate by giving certain data for research purposes. In this proposal, to enable citizen/patient participation personal data banks would be required in which citizens can safely store, manage and share their data. The premise is that having the personal data banks organised as citizen-owned cooperatives would allow citizens and society to participate in the economic value of their personal data.¹³

It may well not happen that every healthcare provider will own the capabilities needed to deliver personalised care to the highest degree to each and every patient (e.g. consider the resources required for genomic testing, clinical interpretation and matching patients with optimal therapies). The answer may lie in the pooling of knowledge to make it available anytime and everywhere. But effective data interpretation requires the rapid involvement of experts and organisations globally. The hurdles of healthcare data fragmentation, representation, organisational boundaries and cross-border information sharing, and fragmented data protection regimes will need to be tackled for innovation in the field to succeed. Also, the massive quantities of data involved¹⁴ from hundreds of different sources hint at the potential of networked environments working in tandem with centralised repositories to keep up with knowledge accessibility and management. Moreover, going forward, translating discoveries into clinical advances for the benefit of patients today will require (quasi-) real-time data sharing and analysis. Current examples span from multi-source patient cohort analysis for personalised care to real-time analysis of hospital patient management data.¹⁵

¹²See <http://euapm.eu/wp-content/uploads/2012/07/EAPM-RE-PORT-on-Innovation-and-Patient-Access-to-Personalised-Medicine.pdf>.

¹³See http://www.datenundgesundheits.ch/wp-content/uploads/2014/03/me13-02-0051_Hafen.pdf.

¹⁴The output data from a genomic sequencer amounts roughly to 200GB; sequencing the 2.6 million new cancers each year in the EU would amount to 500PB of data.

In summary, big data for Personalised Medicine brings together technological, regulatory, ethical, organisational and political considerations:

- It is essential to navigate the ethical and policy considerations of information access, ownership, privacy, and intellectual property, balancing the need for the individual to provide consent for the way in which his/her data is utilised with the need for society to gather and integrate large amounts of data to innovate and deliver in Personalised Medicine.
- Foster collaboration by multiple stakeholders through the alignment of interests and sound information governance.
- Organisations need new knowledge management systems and sophisticated analytics solutions plus robust infrastructures, able to handle these data and generate results rapidly.¹⁶
- Data standardisation is much needed to build on knowledge created and maintained globally (-omics, phenotype studies, imaging, functional in vivo studies, etc.).
- Data distribution systems must be established including data archives, as has been done in the academic field by Institutes such as the European Bioinformatics Institute, with the European Genotype Archive and Array Express archive of functional genomics data electronic health records.¹⁷
- End-to-end security must be well understood and applied to ensure uncompromised operation, respect for privacy and overall trust.

These are just some examples of the multitude of areas that must be addressed.

While there isn't a quick recipe for success, EAPM thinks that a **structured approach** by the EU to Big Data for Personalised Medicine is warranted.

Personalised Medicine and a European data-driven economy

By 2020, the EU should endeavour to achieve widespread benefits for citizens and patients from personalised healthcare by defining in 2015, and subsequently executing a **Data Strategy for Personalised Medicine**. This strategy aims to create the conditions for the EU wide exploitation of the entire data stack for Personalised Medicine. The strategy should consider a 360° view of policy enablers to ensure a comprehensive analysis of all the interrelating decisive factors in play for the development and adoption of Big Data for Personalised Medicine in Europe.

¹⁵See SAP cases <http://www.sap-innovationcenter.com/2013/09/19/medical-explorer/#>; <http://www.sap-innovationcenter.com/2014/01/10/sap-patient-management-analytics/#>; <http://www.sap-innovationcenter.com/2013/09/24/virtual-patient-platform/#>.

¹⁶Capitalizing on the Power of Big Data for Healthcare, Intel, SAS, 2013.

¹⁷Use of 'omics' technologies in the development of personalised medicine, European Commission, 2013.

Thus, with the aim to support a European Data-driven Economy, EAPM proposes a major lighthouse initiative for 2015: a Lighthouse Initiative on Personalised Medicine.

A lighthouse initiative on Personalised Medicine

The internet is a vast worldwide decentralised network of billions of different devices that work as a whole because all the nodes of this network speak the same standard protocol (TCP/IP) on top of which email, www, voice and all the other services we are used to are built. Imagine that such a linked network of data resources could be established in support of Personalised Medicine, closing the circle from the citizen to research to clinical care and back efficiently and in timely fashion.

*The European Commission is invited to focus resources to setup a **Lighthouse Initiative on Personalised Medicine** which, through Member States and multi-stakeholder collaboration would drive policy, regulatory, research and innovation activities to establish a **Europe-wide Data ecosystem for Personalised Medicine**.*

The **Lighthouse Initiative on Personalised Medicine** would not start from scratch, but it would bring a much needed holistic and focused approach to what is a multi-dimensional challenge.

By stimulating collaboration and activities addressing computing infrastructures, data collection, storage, analytics, management, governance, security and privacy these would be put to work to establish a Europe-wide Data ecosystem for Personalised Medicine, supporting among other things, domain specific research, medical decision making at the point of care, patient engagement and entrepreneurship through innovative startups and SMEs.

Many nodes of this ecosystem are being developed across Europe and beyond. For instance, ELIXIR is working on an infrastructure that brings together and coordinates many of Europe's leading bioinformatics resources. The EU-funded project EHR4CR¹⁸ under IMI (Innovative Medicines Initiative) is working on tools and services for reusing data from Electronic Health Record systems for Clinical Research. The ICT industry is developing innovative solutions building upon technologies such as machine learning, high performance computing and cloud computing plus advanced analytics and visualisation. The Smart Data Innovation Lab¹⁹ is working on high performance research with big data/smart data, with real data sources from partners, with Personalised Medicine as one of the focus areas.

New developments on citizen-centric solutions based on mHealth and telehealth are delivering care outside hospital walls whilst capturing more data about individuals' health, contributing to prevention and care personalisation, and supplying crucial information with the potential to increase knowledge about response to treatments. This, when cross-referenced with clinical and genomic data, can be used to gain novel insights about the genesis, progression and treatment of diseases.

Big Data for Personalised Medicine for the benefit of citizens and patients can only realise its full potential after considering its dependencies spanning the technological, regulatory, ethical, organisational and political dimensions. Thus, it is vital to identify those policies that unleash the power of Big Data and orchestrate a strategy around those that can mobilise stakeholders across Europe.

EAPM has identified **three foundational** sets of interrelated policy actions that the EU must comprehensively address in 2015.

Liberate the data but do no harm

The best and fastest possible outcomes will be achieved if scientists can work with and test on large datasets, which are currently not available. How to link these results to clinically meaningful and actionable information, and how to create tailored responses to them, will be a further challenge. Broad collaboration frameworks with transparent working principles and clear governance need to be established.²⁰ Crucially, these need to encompass both public and private players. Currently, academics have archives for public data sharing in genomics, but these may not be accessible to industry and vice-versa. For the innovation cycle to deliver, ultimately, new products and stronger models need to be developed for industry to leverage and foster large data acquisition. As Personalised Medicine requires, in most instances, personalised data, it is important to navigate across the complex regulatory landscape of data protection and, in many instances, clarify the boundaries of what is and is not possible.

Actions:

- The European Commission is invited to take stock of the regulatory landscape for the protection of personal health data across the EU, with Member States and the Article 29 WP with a view to interpreting and clarifying its application towards **Personalised Medicine**. Consent approaches should allow data to flow freely, but patients should remain in charge of their data.
- The European Commission is invited to coordinate an approach with the Member States to **streamline** the cross-border sharing of EHR data and genomic data for secondary use.
- Through broad stakeholder engagement, the European Commission is invited to analyse and share good practices, identify roadblocks, propose corrective measures and issue EU-wide guidance on issues such as breaking silos of single-use data; facilitating data processing; notifications of breaches of privacy; conditions for international data transfer for biomedical research and the Commission should also consider frameworks for data governance that can be applied across Europe.

²⁰For example see http://www.biomedbridges.eu/sites/biomedbridges.eu/files/documents/biomedbridges-principles-of-data-management-and-sharing-at-ris_.pdf for principles of data management and sharing at European Research Infrastructures from the EU funded project BioMedBridges.

¹⁸<http://www.ehr4cr.eu/>.

¹⁹www.sdil.de.

Bring it now

More powerful computing environments are being created every day and the bioinformatics community is working hard to link fragmented silos of information. Cloud computing is one technology that allows moving the compute to the data instead of moving data to compute. This seems particularly suitable for Big Data in Personalised Medicine where moving large amounts of data across sites may not be the most efficient way to access it. Keeping the data where it resides and using technologies such as virtualisation, parallelisation and high performance computing, would, in principle, facilitate the access to this knowledge for large and small research and medical institutions.

Ensuring security in networked environments such as the Cloud is very important. As such, implementing privacy and security should include administrative, physical and technical security safeguards (e.g. encryption), with careful attention to comprehensive training for healthcare workers. Security must mitigate risk in a way that preserves an optimum user experience; otherwise, workers may seek alternatives that can circumvent or disable security. If these risks are not properly managed, eventually the healthcare community may decide not to use such Cloud technologies in critical projects or for some datasets.

Data quality, transparency of methods, and reproducibility are all key requirements of 'good science'. For this reason great efforts are currently being made by some of the major industry, research and academic players in this field to create standardisation and interoperability of data and to open up large datasets for researchers to use (see ELIXIR²¹ and GA4GH²² for examples of European and global initiatives). This effort also covers issues related to enhancing data sharing in a safe, secure and ethical manner. Add interoperability of clinical data, and the promise of an impact upon routine healthcare delivery comes closer. Broadly speaking, collecting, storing, accessing and processing data from diverse locations and multiple sources and formats is a key technical challenge.

But there are foundational principles that once agreed upon will help move the equation forward. What to do with these data in the long term, what to keep, how long to keep it for, by whom and to what use (curation)? How can we ensure the quality of the data, know its provenance and compare "apples with apples" (veracity, interoperability)? How can we be reassured that the right security policies are in place to ensure data is protected (security)?

The future is here and we need to invest a renewed sense of urgency to bring the advantages of Big Data to the bedside and public health *today*. The challenge is to integrate the new data into advanced clinical support systems connected with Electronic Health Records with the clinical training required to use the patient data and bring this science into daily clinical operations.

Actions:

- The European Commission and Member States are invited to identify the opportunities brought by the adoption of

data-derived knowledge in daily clinical operations, including mHealth, and issue a short- and medium- term set of implementing actions, addressing inter alia the adoption of genomic information in clinical care, patient generated health data, medication adherence and pharmacovigilance;

- The European Commission is invited to step-up the support for initiatives that promote the expansion of technologically advanced, secure, quality-assured and harmonised data registries, their linkage and outreach across the healthcare delivery system actors.

Prepare the future

In the last decade, Next Generation Sequencing alongside ICT has changed the way biologists and geneticists do science. The data produced with these machines requires sophisticated computational skills to be analysed. This has given rise to the emerging field of bioinformatics.²³ Because many bioinformaticians are biologists who move to the computational end of the spectrum, they might not have formal training in, for example, computer science and perhaps need support to better-use equipment they have available. The training of this user group and the next generation of scientists is key to sustainable growth in this sector. Importantly, while bioinformatics is now a well-developed discipline, a new set of skills is required now to navigate the plethora of data generated, that of "big data scientists", who utilise data mining, statistics, and domain knowledge (in this case biomedical knowledge), to interpret data and derive solutions from large datasets. The vast amounts of data (volume), the pace at which it is created and at which updates and new findings are released (velocity), and the diversity of such data (variety) as presented above, cannot be dealt with without ICT infrastructures and solutions, such as High Performance Computing, cloud computing, machine learning and analytics, for example. More collaboration²⁴ between the ICT and the life science industries is required to create solutions that biologists and scientists can use, rather than expecting them to adapt to the solution.²⁵ Also, as new Big Data technologies evolve, for storing, processing and analysing the increasing number of data sets, this poses new challenges for security and privacy of healthcare data, alongside the 'traditional' considerations, such as information access, ownership, IP, etc. For example, by combining anonymised data sets and public information, it will soon be possible to "de-anonymise" those data. Therefore, the data accesses

²³Bioinformatics is conceptualizing biology in terms of macromolecules (in the sense of physical-chemistry) and then applying "informatics" techniques (derived from disciplines such as applied maths, computer science, and statistics) to understand and organize the information associated with these molecules, on a large-scale' (Luscombe, Greenbaum, and Gerstein 2001: 346).

²⁴See example of training on HPC for next generation sequencing <https://www.crick.ac.uk/contact-us/intel-crick-ngs-workshop/>.

²⁵Biology must develop its own big-data systems, Boyle 2013, <http://www.nature.com/news/biology-must-develop-its-own-big-data-systems-1.13299>.

²¹See for example <http://www.elixir-europe.org/>.

²²See Global Alliance for Genomics and Health <http://genomicsandhealth.org/>.

need to be controlled properly, and the anonymisation techniques may need to be improved, or at least revised. In this Big Data era, Personalised Medicine will deliver its benefits through greater involvement of patients in treatment decision-making and health management. Equally, healthcare professionals cannot be expected to adapt to new ways of approaching patients and coping with new technologies unless they are suitably trained.

Actions:

- The European Commission is invited to focus Horizon 2020 and IMI 2 funds to address some of the domain-specific challenges and advance the state-of-the-art of core technologies at the **crossroads of ICT and Personalised Medicine**, namely:
 - development of hardware, software and workflow algorithms to accelerate cost efficient analytics and visualisation of genetic abnormalities that cause cancer and other complex diseases;
 - research at the convergence of mHealth, Big Data, Cloud Computing, security and anonymisation techniques to meet the requirements of High Performance Computing and data throughput throughout the life sciences and healthcare value chains;
 - research on policy implications and new models of ownership; and
 - centres of Excellence for big data science and life science applications of HPC systems, including algorithmical research, software technology and education.

In support of **education and training**, the European Commission, Member States, patients' and healthcare professionals' organisations are invited to collaborate in order to:

- complement general public health campaigns with specific messaging to broaden the awareness of Personalised Medicine;
- provide education that improves health professionals' ability to involve patients;
- recognise patients' rights to seek information about care options;
- adapt the provision of continuing professional development and training activities;
- adapt curricula for undergraduate, post-graduate and specialist education; and
- develop training systems that provide for interprofessional collaborative practice and produce interdisciplinary professionals.

To nurture **data-enabled entrepreneurship** the European Commission and Member States are invited to create the conditions for **startups and SMEs** to derive value from a data ecosystem for Personalised Medicine.

Funding

None.

Competing interests

None declared.

Ethical approval

Not required.