European Alliance for Personalised Medicine

A Europe–wide data ecosystem for personalised medicine
A proposal for a Lighthouse Initiative

Background

The European Commission in its Communication “Towards a thriving data-driven economy”[1] calls for the EU to support "lighthouse" data initiatives capable of improving competitiveness, quality of public services and citizens’ lives.

Such initiatives maximise the impact of EU-funding within strategically important economic sectors. Furthermore the Communication presented personalised medicine as one of the possible target areas.

Generally speaking, economically, Europe needs to close a widening productivity gap between itself and the US and this can be greatly assisted by the use of more information technology across 28 Member States mostly in the use of this data rather than its actual collection.

For example, it was recently estimated that Big Data could save the public sector €100 billion in ‘operational efficiency improvements’.

Current users of Big Data include Royal Dutch Shell, which has formed partnerships with the likes of IBM, HP and Dreamworks to use data, sensors and more to explore thousands of oil wells while, in healthcare, data could be a valuable tool to upgrade healthcare quality and lower costs.

For example, regarding personalised medicine, Pfizer has used advanced data to develop a drug for a type of lung cancer in connection with a specific gene mutation. This has been given conditional approval by the European Medicines Agency.

Increases in computing power and new technologies have shot down the price of using data in healthcare.

These days ‘wearables’ and more are already modernising healthcare as tracking activity, remote monitoring and reminders to encourage medicine adherence help to bring about better healthcare outcomes. This means greater independence for the patient, a better quality of life and much-lower hospital bills. It is estimated that reducing non-adherence in medicines could save €125 billion annually and reduce premature deaths by 200,000 each year. That is in Europe alone.

Moreover, using Big Data effectively will bring about improvements in efficiency, improvements in decision making, and smarter investment.

1 https://ec.europa.eu/digital-agenda/en/news/communication-data-driven-economy

EAPM Office: Avenue de l'Armee/ Legerlaan 10, 1040 Brussels,
Board: H. Brand, M. Baker, D. Byrne, G McVie, H. Galbraith, A. Paradiso, D. Jacqmin
Denis Horgan, EAPM Director, Tel: +32 (0)472 5 3 51 04, E-Mail: denis.horgan@euapm.eu,

The creation of the Big Data Value Association (BDVA)[2] and the European contractual Public Private Partnership[3] (cPPP) on Big Data Value brought forward two powerful vehicles to support the development and adoption of data-centric initiatives that can contribute positively to the socio-economic and technological progress of Europe.

Why Personalised Medicine matters

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.

A focus on personalised medicine would be extremely beneficial to Europe. The area of personalised medicine has grown rapidly over the last decade due to improvements in areas such as genomic sequencing, diagnostic testing etc. This has allowed us to gain a deep mechanistic understanding of a disease, meaning that patients can now be sorted into groups who are likely to benefit from a specific treatment.

Personalised medicine therefore has massive potential to make treatments more effective (with fewer side effects) and more cost effective. Creating a Data Ecosystem for Personalised Medicine 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, 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 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.

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 $1,000 barrier had been conquered[4]) have outpaced Moore's Law[5], 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.

---

2 http://www.bigdatavalue.eu/

3 http://ec.europa.eu/digital-agenda/en/data-public-private-partnership

4 http://www.businessweek.com/articles/2014-01-14/illuminas-dna-supercomputer-ushers-in-the-1-000-human-genome

5 *Moore'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 programmes, and in these projects genome data will be integrated with a patient's health record[6]. 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 programmes, 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:

- 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[7].

- 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.

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.

6* http://www.technologyreview.com/news/520571/why-the-uk-wants-a-genomic-national-health-service/

7* http://shire-hgt.isebox.net/file/%5B37458%5D%20RareDiseaseImpactReportforWeb.pdf?download

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, 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 over-expressing the HER2 gene, which was an early pioneer of many subsequent drugs based on similar "personalised" principles.

The BDVA, in its Strategic Research and Innovation Agenda[8], refers to a Lighthouse project as running data-driven large scale demonstrations whose main objectives will be to create high-level impact, and being the major, high impact mechanism for Europe to demonstrate Big Data Value ecosystems and sustainable data marketplaces that lead to increased competitiveness of established sectors, as well as the creation of new sectors in Europe.

In the next section we propose a large scale data-driven demonstrator for Personalised Medicine on cancer-related data aggregation and analysis.

A Data Place for the advancement of cancer research and treatment

The Lighthouse initiative will setup and operate a Europe-wide ecosystem for personalised medicine, able to aggregate, share and analyse millions of data points. It would make this data and analytic tools available to hundreds of researchers and healthcare professionals and it would develop a business model that incentivises data donation and citizen participation.

Personalised medicine requires computing environments able to process massive amounts of information for research and diagnostics and, at the same time, the ability to integrate within the clinical environment 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.

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

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 lies in the federation of knowledge to make it available anytime and everywhere.

8" http://www.bigdatavalue.eu/images/SRIA/EuropeanBigDataValuePartnership_SRIA__v1%200_final.pdf

But effective data interpretation requires the rapid involvement of experts and organisations globally. The hurdles of healthcare data fragmentation, representation and organisational boundaries will need to be tackled for innovation in the field to succeed.

Also, the massive quantities of data involved[9] from hundreds of different sources hint at the potential of networked environments working in tandem with 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[10].

In summary, commoditisation of genome sequencing brings a wealth of genomic data, yet it is being submitted to multiple repositories in different formats. This fact limits the ability of researchers to collaborate across data silos and makes it harder to run analytics across federated pools of data. Also, the computing intensity required to draw insights from the data is not within reach of many institutions.

The Lighthouse proposal would provide a virtual Data Place for the advancement of Personalised Medicine for cancer and other diseases which, albeit decentralised, would be able to operate securely and with respect for patient confidentiality as a cloud of centres providing comparable data with analytics and visualisation capabilities.

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, validation, 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 start-ups and SMEs.

Many nodes of this ecosystem are being developed across Europe and beyond. For instance, ELIXIR[11] is working on an infrastructure that brings together and coordinates many of Europe's leading bioinformatics resources and the ICGC[12] represents an international effort to create a comprehensive description of genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes.

The EU-funded project EHR4CR[13] 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

9* The 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

10* See 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/#

11* See for example http://www.elixir-europe.org/

12* International Cancer Genome Consortium https://icgc.org/

13* http://www.ehr4cr.eu/

such as machine learning, high-performance computing and cloud computing plus advanced analytics and visualisation. The Smart Data Innovation Lab[14] 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.

The Lighthouse on PM could:

- provide a Data place through a virtual platform - intensive discovery;
- make available cancer genomes, ancillary data and put in place mechanisms to incentivise data contributors;
- provide advanced bio-informatics tools while optimising genomic pipelines for performance and accuracy, and provide an environment for the clinical deployment of genomics;
- provide a platform for public health genomics and link biobank infrastructures, while meeting the diversity of systems, data formats and interoperability;
- ensure governance, including quality assurance, security and patient confidentiality;
- encourage cooperation between healthcare systems across the European Union;
- help to improve investment in research while driving innovation across the EU;
- provide a platform for saving millions of euro by, for example, using data and the associated technologies to 'treat' and monitor patients at home rather than in expensive hospital beds
- contribute to improving (and, in many cases, saving) the lives of a potential 500 million patients across 28 Member States

14 www.sdil.de