16 **Alan Cooper:** The head of ESA's Brussels office explains the co-ordination of ESA and EU space activities

55 **Rodrigo Martins:** The E-MRS's past president discusses the sustainability of materials and smart specialisation

264 **George Monbiot:** The eminent environmentalist discusses the possible return of Eurasian lynx to Scotland

Pan European Networks

Science & Technology

GIFTED CRYSTALS

Liquid metal holds the key to revolutionising energy storage

Front cover inspired by: GroupSadoway EXTREME ELECTROCHEMISTRY

Supported by: cabmm

in association with: European Science Foundation

DR ENRICO CAPOBIANCO AND **DR M GIOVANNA TRIVELLA** ADDRESS THE PROBLEM OF BIG DATA COMPLEXITY IN THE MEDICAL ARENA

Systems medicine road map

Systems medicine is flourishing and is almost naturally stimulating interest in many regards. Fig. 1 shows a map of the main impact areas, those for which a future role of systems medicine is expected. While observing the map as one integrated object, the first thing that comes to mind is that we are facing an ecosystem, i.e. a connected community of living and non-living elements that work together and form a complex space of relations.

The consequence of such an environment is that massive heterogeneous information is generated, but it is hopeless to target analytics towards each source component in isolation from the others. Furthermore, the co-existence of direct (causal) and indirect (feedback) dynamics contributes to shaping an overall complexity for which only approximate solutions may exist.

A second observation can easily be derived from Fig. 1: data from multiple sources calls for linked-data technologies to assimilate them and for cross-matching methods to ensure their most appropriate analysis at all the embedded scales, from the individual to the community (sub-population or population). For instance, together with active data derived from measurements of various types (i.e. electronic health records or EHR, real-time monitoring of behaviour, sociodemographic factors etc.), passive data is also relevant in such a context, being derived from sensors measuring disease signs and/or unexpected disturbances otherwise not detectable (during sleep, physical activity or stress conditions).

The third aspect worth very careful consideration is that in order to guarantee a synergistic integration of the ecosystem-related data to lead to high-impact health solutions, three concepts need to be connected: social, technical and methodological interoperability. In particular, the first of these enables a communication network centred on the healthy and diseased individual, and calls for novel technical and methodological standards to be created.

Fig. 1 Systems Medicine Ecosystem. Circles include the major components, while the indicated actions and solutions (labels joining the circles) involve determinations by both individual choices and motivations, and by organised decisions, activities and evidences

Since pattern discovery, community detection and network interdependence will be goals of such a translational research approach, the shift to data-intensive systems medicine research follows as a necessity for holistically deciphering how the described ecosystem behaves, what its equilibrium/disequilibrium dynamics are and what types of attractors are determining its states.

Opportunities

Among the current needs, a priority is establishing a patient-centric vision. Individualisation, personalisation, precision and other modern attributes adopted in medicine are all converging to that vision, and in a timely manner. Changes in society and technology inevitably alter the established relationships between the systems medicine ecosystem components, and the novelty is determined most significantly by two main drivers: the speed at which changes occur, and the depth and width with which the effects of changes are perceived. Therefore, the ecosystem complexity is a constantly transient entity subject to a redefinition of its components and a remodulation of its dynamics.

If under such variable conditions it is clear that translational approaches are needed, the question becomes: how are they really going to be enabled, and when can we say that they would be successful?.

We believe that cultural cross-fertilisation is required, involving interdisciplinary research teams to redefine health and disease through a better understanding of the mechanisms of homeostasis and their chances of recovery when they are compromised by events, including the 'innovative' roles played by both systematic variables and surprise factors during a person's lifespan.

Challenges

Several difficulties must be faced before systems medicine opportunities can be fully exploited. In particular, we would like to emphasise the following factors:

1) Barriers are present among different disciplines, mainly due to different languages and practices. Examples include the cultural protection of sectorial data importance and hierarchy, the lack of acceptance of the new paradigm of open science, and the vision of teamwork as a form of legacy required for sharing protocols and resources;
2) The necessity of making independent validations central to the quality control of any source of information to be used when inference from data is conducted;
3) Cross-disciplinary problem-solving approaches involving multiple expertise and co-ordinated actions; and
4) Cost-benefit analysis aimed to consider the design of new experiments, protocols, trials etc. *versus* data re-use and investment in analytics to retrieve incremental value.

Fig. 3 Automated Clinical Information Processing. A decision support system scheme for stratifying risk into classes, and these into communities, is represented in its main steps, from input data to outcomes, including significance of scores and evaluation by metrics

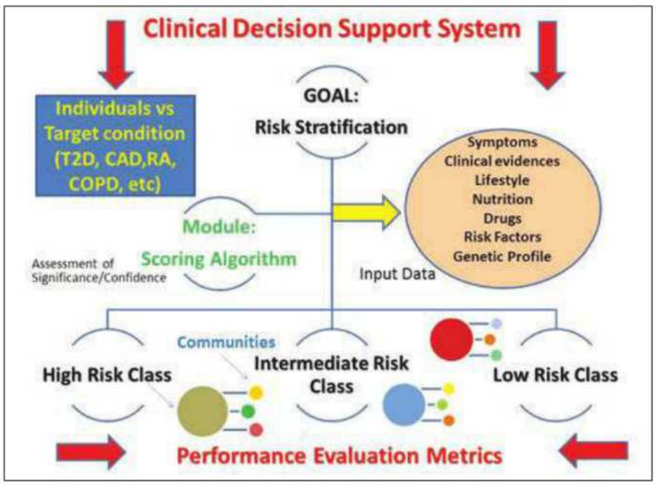
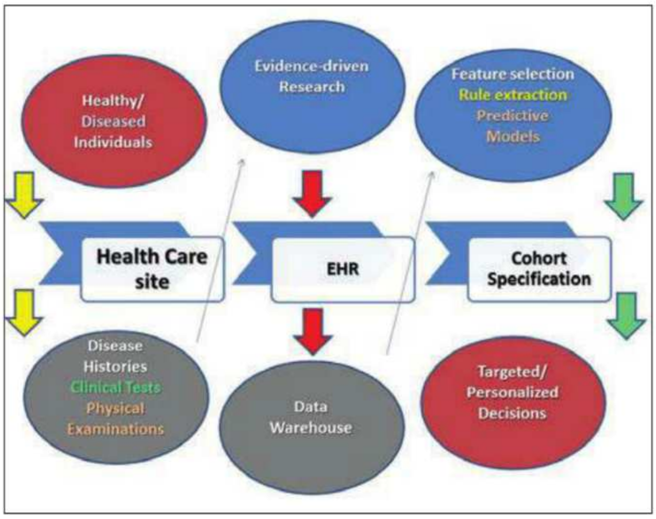


Fig. 2 Personalised Applications. From building evidences to organising the information and warehousing records. Inference applies to data towards individual solutions through selection of features and application of algorithms and/or automatic rules



Looking inside the ecosystem, two main applications are envisioned depending on the target data type, and two internal views are correspondingly offered (see figs. 2 and 3, respectively).

With individual data, personalised applications would rely on evidence of physiological changes and pathophysiological responses to elements such as climatic and environmental conditions (from humidity to pollution, for example), food intake, lifestyle (including physical activity and sleep) and habits (smoking, alcohol consumption), use of drugs (therapeutic and non-therapeutic substances), and work (especially considering health damaging stress conditions and/or unhealthy working place). See Fig. 2.

With population data, the possibility of forming clusters sharing similar characteristics or communities connected by multiple features may allow for the identification of novel risk factors which address disease tendency and health threats. See Fig. 3.

A final point of great relevance in systems medicine is the assessment of the confidence for measurement accuracy and score reliability. It is clear that a certain degree of uncertainty – which is part of the overall observed system's variability – remains unexplained because it is stochastic in its nature and may play a limiting role in the reproducibility of results.

However, the advantage of operating at a systems scale is that a multitude of random fluctuations can be present, also exerting a regularisation effect by averaging and smoothing extreme variation.

Enrico Capobianco
Laboratory of Integrative Systems Medicine
IFC-CNR, Pisa-Italy and Center for Computational Science
University of Miami, Fl-USA

phone **+1-305-689-7031**

email **ecapobianco@med.miami.edu**

M Giovanna Trivella
Experimental Medicine
IFC-CNR, Istituto di Fisiologia Clinica – Consiglio Nazionale delle Ricerche
Pisa-Italy

phone **+39 050 3152730**

email **trivella@ifc.cnr.it**
browse **www.sensorart.eu**