

- Paper title - Big Data, Big Knowledge: Big Data for Personalized Healthcare
- Abstract - The idea that the purely phenomenological knowledge that we can extract by analyzing large amounts of data can be useful in healthcare seems to contradict the desire of VPH researchers to build detailed mechanistic models for individual patients. But in practice no model is ever entirely phenomenological or entirely mechanistic. We propose in this position paper that big data analytics can be successfully combined with VPH technologies to produce robust and effective in silico medicine solutions. In order to do this, big data technologies must be further developed to cope with some specific requirements that emerge from this application. Such requirements are: working with sensitive data; analytics of complex and heterogeneous data spaces, including nontextual information; distributed data management under security and performance constraints; specialized analytics to integrate bioinformatics and systems biology information with clinical observations at tissue, organ and organisms scales; and specialized analytics to define the “physiological envelope” during the daily life of each patient. These domain-specific requirements suggest a need for targeted funding, in which big data technologies for in silico medicine becomes the research priority.
- Paper authors - Marco Viceconti, Peter Hunter, and Rod Hose
- Link to the paper - <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7047725>
- Summary – In 2009 Google Flu, simply by analyzing Google queries, predicted flulike illness rates accurately. Some analysts started to claim that all problems of modern healthcare could be solved by big data. VPH is a methodological and technological framework that, once established, will enable collaborative investigation of the human body as a single complex system. VPH's vision contains a tremendous challenge, namely, the development of mathematical models capable of accurately predicting what will happen to a biological system. The real challenge is the production of that mechanistic knowledge, quantitative, and defined over space, time and across multiple space-time scales, capable of being predictive with sufficient accuracy. Some examples of VPH applications that reached the clinical assessment stage are described in detail. Suggesting that we revert to a phenomenological approach where a predictive model is supposed to emerge not from mechanistic theories but by only doing high-dimensional big data analysis, may be perceived by some as a step toward that empiricism the VPH was created to overcome.

Both in traditional engineering and in medicine, the research domain is defined in terms of problem solving, not of knowledge discovery. The motto common to both disciplines is “whatever works.” Big data usage in VPH is described in detail using the problem of predicting

the risk of bone fracture in a woman affected by osteoporosis, a pathological reduction of her bone mineralized mass. The most famous phenomenological models capable of predicting the likelihood of fracture, FRAX is discussed. This example of predicting a bone fracture using statistical, population-based knowledge with mechanistic, patient-specific knowledge, whether it can be considered a big data problem, in the light of the “5V” definition (Volume, Variety, Velocity, Veracity, Value) is then described.

In some cases there is need for an explanatory theory, which answers the “how” question, and which may be used in a wider context than that a statistical model normally is. In a complex scenario such as this, are the currently available technologies sufficient to cope with this application context? The brief answer is no. A number of shortcomings that need to be addressed before big data technologies can be effectively and extensively used in computational biomedicine are addressed which are Confidentiality of Data, Big Data: Big Size or Big Complexity?, Integrating Bioinformatics, Systems Biology, and Phenomics Data, Where are the Data? and Physiological Envelope and the Predictive Avatar. Although sometimes overhyped, big data technologies do have great potential in the domain of computational biomedicine, but their development should take place in combination with other modeling strategies, and not in competition and thus big data analytics can be successfully combined with VPH technologies to produce robust and effective in silico medicine solutions