

Health data technologies: the current challenges

Himanshu Tyagi

In the future, when we look back at this time, healthcare informatics will be seen as the biggest healthcare revolution of our generation. Within the field of healthcare informatics, mental health is arguably one of the most information-rich fields. Most mental health problems are chronic in nature, necessitating several encounters between the clinician and the patient over many years. In addition, mental illnesses are poorly understood in terms of their exact neurobiological mechanisms, making it difficult to have standard laboratory tests to quickly establish the diagnosis. In practical terms, this means that unlike other professionals in health care a mental health professional does not have the luxury of quick and reliable methods to make clinical decisions. Therefore, the practice of mental health is marked by multiple lengthy clinical interactions with individual patients, and the information generated by such encounters differs both in quantity and quality from most other fields of medicine. The information-rich practice of mental health professionals makes it vital that they are actively engaged to shape the agenda in the wider field of health informatics. In this article, I am trying to cover some of the current issues with health data and informatics. It is based on the insights I gained from being at a unique vantage point of a psychiatrist and a tech entrepreneur, but the problems and potential solutions I am highlighting are equally applicable to every part of the modern healthcare information technology (IT) systems.

When we think of using IT in healthcare settings, we are mostly focusing on its transformative potential. The broad assumptions about the immense benefits that IT can bring to this sector are undoubtedly valid, but successful implementations of such technologies have proven to be difficult and challenging. Despite being an early adopter of IT, the healthcare sector in the developed world is still under-utilising this technology to the peril of the patients and policy-makers alike. While other major industries are moving on to the next stages of their digital transformations, health care is still negotiating with the transition from paper-based records to digital records. Many factors unique to health care make this transition difficult, but a commonly seen theme is the slow pace of policy innovation in healthcare IT. Let us examine some of the factors that policy-makers might find useful to drive innovation in this area.

Data disconnect

Healthcare data is mainly collected and stored in the following three separate data pools: clinical data; financial data; and research data.

Because the responsibility for collecting, storing and utilising this data rests with different individuals or institutions (for clinical data, the hospitals and clinics; for financial, managers and governments;

for research, universities and pharmaceutical companies), they remain largely isolated from each other with little interconnectivity. The disparate systems on which these datasets are located are typically unsuitable for complex integrative analysis. This disconnect between healthcare data can be detrimental to the early identification of important healthcare trends or adverse events, as highlighted by the withdrawal of a popular pain-relief drug, rofecoxib, in 2004.

Rofecoxib was approved by US Food and Drug Administration in 1999 and gained widespread acceptance among physicians worldwide, who prescribed it to over 80 million people. In 2004, a California-based, integrated, managed-care consortium Kaiser Permanente connected clinical and financial data to compare the risk of adverse cardiovascular events for users of rofecoxib against a similar drug; it found that rofecoxib might have been responsible for more than 27,000 avoidable myocardial infarctions (heart attacks) and sudden cardiac deaths between 1999 and 2003. This study led to a voluntary withdrawal of the drug from the market. Interestingly, between 1999 and 2004, similar conclusions were suggested by a number of small-scale studies; however, none was considered large enough to raise sufficient concerns. The simple act of combining clinical and financial data provided the crucial research dataset that was required to trigger one of the largest medication withdrawals in history.

Although the above example is a powerful indicator of the potential benefits of having an integrated approach on healthcare data, large-scale implementation of such approaches have proven to be challenging and therefore remained largely underutilised. In the UK, implementation of Payments by Results (PbR) in its national health service to integrate clinical outcomes with financial remuneration has produced mixed results.¹ Such approaches usually require a fundamental reorganisation of the industry processes and support by technology appropriate innovations in policy.

One such example is the recently launched government-funded Secure Unified Research Environment (SURE) project in Australia, which aims to overcome such limitations by providing a central data centre where researchers can form connections between data sources and access the necessary computing power required to perform such analysis.² In its short span of active operation, researchers using this integrated database have been able to confirm the intuitive beliefs that the older Australians are more likely to have higher consistency of care, and that lower consistency of care is associated with geographical remoteness. It also led to a counter-intuitive discovery that wealthier and more highly educated Australians have a lower consistency of care. It is important to note that although researchers were able to test intuitive beliefs using more complex and time-consuming methods before the existence

of the SURE database, counter-intuitive discoveries would not have been possible by looking at a single dataset alone.

Data flow

One of the core drivers in the healthcare industry is the information about health and disease. Clinicians are performing the task of eliciting information about illness from the patients, filtering out the noise from signals and producing actionable information, which leads to interventions. They are also burdened with the task of storing (documenting) and disseminating this information to patients and other relevant healthcare professionals. Some estimates put the time spent by a clinician in eliciting, documenting, storing and disseminating this information as up to two-thirds of their clinical time, with less than 20 per cent actually spent in performing interventions (treatment). The role of information is so central in health care that it can be argued that its main business is the business of information. Therefore, any factors that can alter the dynamics of information in our society are bound to affect healthcare business. The last two decades have witnessed many fundamental shifts in these dynamics. If we look beyond the obvious proliferation of technology in the form of round-the-clock access to information from the internet from mobile computing devices or via online social networking, we would find that the way we consume information is undergoing a fundamental change. As a society, we are rapidly moving away from synchronous interpersonal communication (wherein both the originator and recipient of information have to be in same place at same time) to asynchronous communication. This shift is not new, however, as it has been a part of human civilisation ever since we learned to draw in caves in ancient times. Asynchronous communication became significant after the invention of paper and printing, but the internet is giving it the impetus to become the dominant form of communication in human society. This also means that an ever-larger amount of information is accessible by ever-greater numbers of individuals.

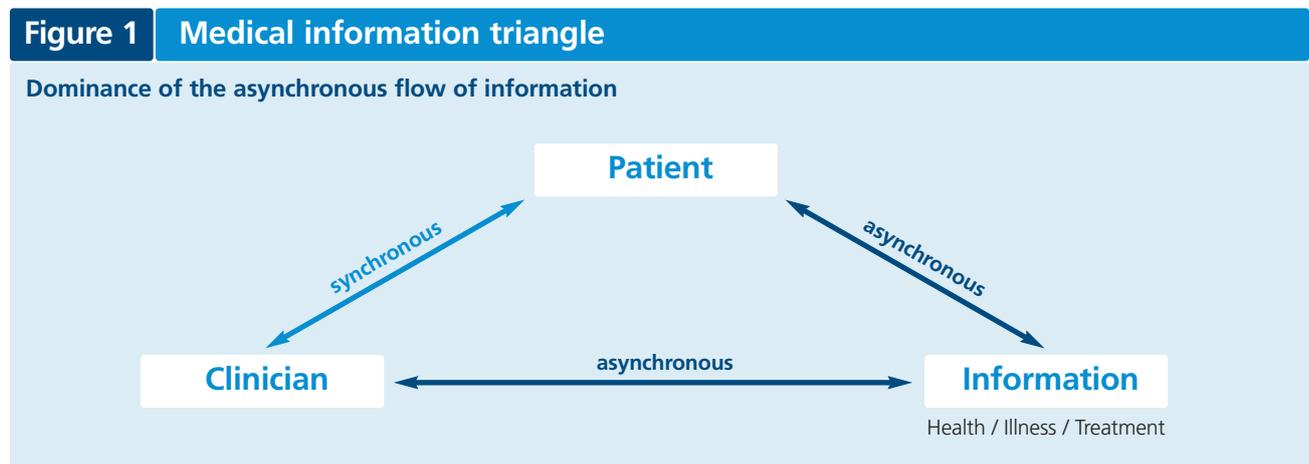
In healthcare settings, this means that the clinician might no longer be the primary information provider to a patient. Today, both doctor and patient have become voracious information consumers, using easily accessible knowledge to inform our decisions in millions of different ways. This shift is also responsible for gradually changing the public perceptions of healthcare professionals, who

are now less valued for their knowledge but more for their skills and experience. With easy access to detailed information, a tech-literate patient could possess more information about his illness than the doctor who had to study several thousand illnesses.

However, in spite of that, our healthcare IT systems are usually designed to suit 20th-century practices – i.e. one-to-one synchronous interactions between the doctor and patient. This approach fails to capitalise on possible advantages offered by asynchronous communication, which has the potential to dramatically increase the reach of individual clinicians. But most importantly, it fails to correct one of the most important drawbacks of the dominance of asynchronous communication – i.e. it is more error-prone than synchronous modes of communication. This is because unlike synchronous communication, which has spontaneous feedback loops (doubts can be immediately clarified by interacting parties in real time), feedback loops have to be created in asynchronous information ecosystems. However, as an industry we have made little progress in this area, and still treat to keep most of the patient’s information as the proprietary information of the organisations. By evolving future healthcare systems in the direction of the shift in the flow of health information, the industry can potentially boost the productivity of individual clinicians. Rather than be restricted by traditional, physical, case-load limits of a few hundred patients, a future clinician can practise on an unprecedented scale when aided by the right technology and policies.

Big data

Medical information almost doubles every few years and the rate of production of new medical information is accelerating. Advances in medical knowledge often make established treatment models obsolete. In a typical year, frontline clinical staff would have 22,000 new peer-reviewed articles, 30 new drugs and 6,000 new combinations of drug compatibilities to consider in addition to their existing knowledge of medicine. The number of drugs has grown by 500 per cent in just the last decade and the technological advances in medical imaging are producing more data than ever before for the same procedures (for example, high-resolution CT Scans). Not only has medical data exploded in recent years, it has also become more accessible to both the patients and providers alike.



Traditionally, meaningful information has been extracted from large sets of data by sampling a representative portion of the dataset. Sampling methods, with all their limitations, were essential as we did not have the tools and resources to comprehend the entire datasets. However, with newer technology this limitation is rapidly being removed. One such approach, encompassing various techniques commonly referred to as 'big data', is helping information-driven industries to analyse entire datasets regardless of their size and scope. At the heart of these new techniques is a simple premise – 'Why analyse a fraction of data when we can analyse everything?' Big data also helps us to move away from the post-hoc statistical analyses that are unable to provide real-time measurements. Although other information-driven industries have been quicker to adopt big data models, the healthcare industry is uniquely placed to benefit the most from it for the following reasons:

- most healthcare data is recorded and validated;
- it spends billions on research and development;
- healthcare data is ever increasing, thereby stretching the resources; and
- it would help to focus on preventive measures.

Big data models have emerged only in the last few years and their growth has been fuelled by internet-based companies like Google, Yahoo! and Facebook who face the challenge to meaningfully analyse the datasets generated by billions of individuals. 'Flu trends' by Google³ is just one example wherein the algorithmic analysis of big data sets (i.e. all search queries) is providing almost real-time estimates of current flu activity throughout the world. This online tool, which has been designed by engineers at Google with little background in health care, is accurate enough to closely match the official government estimates of flu activity and still has the advantage of spotting the emerging trends two to four weeks before healthcare agencies. This is just one example of the transformative power of big data in health care.

Data tools

Big organisations in information-driven industries are gradually moving away from creating software data tools themselves. The task of creating software is being outsourced to anyone with an interest in it by using an approach referred to as Model Driven Development (MDD). Considered a fanciful idea just a few years ago, MDD has moved from a novel concept to a pragmatic business necessity in large corporations, with Apple's App store being just one of the examples of the success of this approach. Implementing MDD in health care can prevent the sector from spending billions on complex IT projects, while delivering comparable or even better software tools. The main benefits of such an approach in health care could be:

- reduction in development cost and time;
- interoperability between different software;
- localised specifications to better address complex variable needs; and
- better opportunities for clinicians and patients to engage in software development.

The MDD approach requires the healthcare authorities to look at healthcare IT as an ecosystem rather than as a project. It hence takes on the responsibility for establishing standards and benchmarking, providing immense room for policy innovations. MDD software approaches have proven to be inherently adaptable to changing circumstances, making them an excellent fit in healthcare settings, where changes in government, technology or medical knowledge require resilient IT solutions ('future proofing').

Conclusion

Healthcare IT can benefit by learning from other information-driven industries. However, in most cases, the project leadership is unable to see the big picture in both IT and the health industry. This is because in health care, the creator of technological and data solutions is not usually the consumer of these tools. To succeed and stay ahead of the curve, healthcare IT would need to foster leadership positions for people uniquely qualified to take on integrative leadership roles, which can facilitate such vision. In the UK, some health care authorities have started recruiting Chief Clinical Information Officers (CCIO)⁴ to their boards, which is an encouraging step in this direction. Otherwise we would be perpetuating this strange dichotomy wherein patients and clinicians are technophiles at home with their smartphones, but technophobes once inside a healthcare setting.

Endnotes

- 1 NHS Payment by Results
<http://www.dh.gov.uk/health/category/policy-areas/nhs/resources-for-managers/payment-by-results/>
- 2 Secure Unified Research Environment
<https://www.sure.org.au/about-sure/project-partners>
- 3 Google flu trends
<http://www.google.org/flutrends/about/how.html>
- 4 CCIO <http://www.ehi.co.uk/campaign/CCIO>

Dr Himanshu Tyagi, MRCPsych, is a psychiatrist and a co-founder of the popular medical education website rxpg.com. His current work includes research in obsessive compulsive disorder and mental health informatics. He received a national award from the Royal College of Psychiatrists, London, in 2012.