On the Record

Precision medicine is fast becoming recognised as the way forward for the treatment of diseases. Having the right patient medical information is key, and IT technologies are crucial in being able to analyse it. Of all the data sources, electronic medical records seem to hold the most potential

Tigran Arzumanov and Le Vin Chin at Clinerion

Undoubtedly, precision medicine is gaining momentum. Recent reports put the market on a growth track to reach $65-75 billion by 2021, with an estimated average growth rate of 10-12% (1-3). A flurry of announcements regarding the main focus therapy area, oncology, shows that the field is attracting a lot of research attention as well as funding.

Cases in point: the $215 million committed to the Precision Medicine Initiative database in 2015 and the subsequent $1.45 billion allocated to precision medicine as part of the 21st Century Cures Act by the Obama Administration (4). Meanwhile, China launched a five-year, $3 billion precision medicine initiative in 2016, and the UK has announced €300 million funding for a four-year project to map 100,000 of people's genomes this year (3,5). The Swiss pharmaceutical hub of Basel – home to powerhouses Novartis and Roche – has set up a cross-industry innovation platform for precision medicine, and Qatar has established a world-class genome project and biobank (6,7).

This surge of activity has been spurred by innovations in medical technologies, such as drug discovery, companion diagnostics, next-generation sequencing, bioinformatics and big data analytics. These allow the diagnosis of certain diseases based on the analysis of genetic information – via genomics and specific biomarker techniques – and the ability to treat particular diseases with new genetic manipulation methods.

Personalised Medicine, Precision Medicine

These developments enable greater drug differentiation. At the extreme, drugs might be tailored to each individual's genetic makeup to offer a 'personalised' medicine. While economics of scale revolt against going to this extreme in every case, this is nonetheless a radical change for the pharma industry, which has been built on high-selling, large volume 'blockbuster' drugs. Precision medicine is the happy medium found by combining standardisation with individualisation via the accurate prediction of the treatment and prevention strategies for a particular disease and specific groups of people.

Benefits include enabling a better understanding of disease mechanisms, assessment of disease risks and prediction of

optimal therapy, taking into account individual variability in genes, environment and lifestyle for each person. This provides individualised care in the form of disease treatment and prevention.

The clearest case for more personalised medicine is the treatment of cancers, where the therapies focusing on tumour genotypes are the most effective. Oncology therefore takes the largest share of the precision medicine market today (30%), but its advantages can also be seen in the treatment of rare diseases, where next-generation sequencing techniques can play a pivotal role (4). There are an estimated 7,000 rare diseases afflicting as many as 350 million people around the world – 80% of which involve a genetic component (8).

Big Data Challenge

It might become clear that data are key in getting precision medicine to work. This kind of medicine relies heavily on IT, whether in regard to the data analytics that power drug discovery and bioinformatics or the mechanics and logistics of patient data management. To improve success rates, pharma is increasingly attempting to integrate real world evidence into early phases of drug development. This requires the use of big data analytic techniques to collect, make sense of, put into context and draw useful conclusions from large swathes of electronic health records.

The gathering, storage and use of patients' personal medical information – and the rights and responsibilities thereof – are gradually developing and codifying conventions for the protection of personal health data, such as the EU's General Data Protection Regulation (Regulation [EU] 2016/679), which will come into full force in May 2018 (10).

Ultimately, the question precision medicine asks is: can we identify a target patient population for whom we can develop a specialised treatment? This is a core challenge that that involves two stages. The first is to cluster patients intelligently to find the groupings where a targeted treatment can impact a disease more effectively than a general cure for the wider population. The second is to identify these patients so that they can be offered the new treatment. Big data analytic techniques are ideally suited to find the needle in the haystack.

Patient Data Sources

With information as the foundation, where can we get data to profile patients? The first, most obvious source is medical records, which are already stored in hospitals and other healthcare facilities around the world. If these records are digitised, they can be electronically queried to find individuals with the right profile. Another resource might be from historical clinical trial data. Further options include patient registries, especially those which automate patient data collection; new and original data can be collected from research at biobanks; patients suffering from specific diseases may be found via patient advocacy groups; and patient support groups, which can also be a rich resource, whether in the real world or on social media.

It is, of course, important that data drawn from multiple sources are of a good quality, although never 100% complete or consistent. Oracle's Jonathan Sheldon agrees; in his article 'Three key ingredients to clinically actionable precision medicine', he cites "assess data quality and establish governance" as the very first step (9).

The diverse data sources have varying degrees of problems. Social media and other forms of contacting patients via self-identification in the public realm – such as through direct marketing – offer a high degree of false positives. Patients identified in this way may not be connected to hospitals with the right facilities to run further tests or to offer new treatments. Patient advocacy groups will not have complete patient datasets and are not set up to deliver them quickly. Registries also experience delays; there will be a duplication of effort, as they are created for each indication.

Hospital electronic medical records (EMRs) are the best bet. These systems are rich in information and provide data unavailable elsewhere. Data are complete and up-to-date, allowing quick and early detection of indications and patients. Users also gain a direct link to the doctor and not just the patient.

However, these systems often do not have the tools to search on a large scale within each institution. Due to the way they have evolved, the data are stored in silos, and searches can therefore not be carried out easily across institutions and countries. Patient data standards will also differ among providers and countries, making a cross-system search more difficult.

As we have seen in the earlier discussion about the protection of personal health data, the protection of patient privacy is also an issue, which needs to be clarified for this source to be effective and trusted. Overlapping with this concern is the requirement for all techniques for gathering, storing, managing and using the data to be legal according to local laws.

Making it Work

An EMR has the most potential as a data resource for precision medicine due to its richness, quality and speed. There are a few ways to overcome its disadvantages.

Patient Data Privacy
Privacy by design: protection of personal data needs to be 'baked in' to all processes and systems. Operate on the principle of 'privacy by design', which means that patient privacy is considered throughout the design and development process of new tools and data flows.

Anonymisation: regardless of the locale, data protection regulations permit the sharing of de-identified data. For instance, the EU's Data Protection Directive 95/46/EC – which strictly prohibits secondary uses of person-specific data without individual consent – provides an exception to

"Ultimately, the question precision medicine asks is: can we identify a target patient population for whom we can develop a specialised treatment?"

the ruling in Recital 26, which states that: "The principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable." This means that any data, including health information in EMRs, can be reused for research purposes once they are de-identified.

Pseudonymisation: the Handbook on European Data Protection Law, from the European Council (in April 2014) (§2.1.3), states: "…pseudonymisation of data is one of the most important means of achieving data protection on a large scale, where it is not possible to entirely refrain from using personal data… This is particularly useful where data controllers need to ensure that they are dealing with the same data subjects but do not require, or ought not to have, the data subjects' real identities. This is the case, for example, where a researcher studies the course of a disease with patients, whose identity is known only to the hospital where they are treated and from which the researcher obtains the pseudonymised case histories" (11).

In this model, the only individuals with the ability to translate the pseudonym back to a patient's identity are those who hold the keys to the pseudonymisation process. This should clearly be the hospital itself.

Legality

Data residency/local hosting: a clear solution to the challenges of removing patient data from hospitals is to ensure that they remain there the entire time. Another solution would require analytics processes to either deal with the data at the storage location or to only extract anonymised/de-identified data to work on – or both.

Data Silos

Federated network: issues of data residing in different hospital infrastructures, based on distinct systems in different geographical locations, can be solved by the use of federated networks and integrating data at the point of query. For example, the query goes out to each federated node at the same time and the returned results are later aggregated in the cloud.

Data Standards

Semantic interoperability: different systems, hospital groups and countries imply vastly different standards being applied at each hospital. This can be solved through the use of semantic technologies to render the queries and the returned results semantically interoperable.

Conclusion

These methods are worth the effort. EMRs offer many advantages, with speed as the ultimate prize. If done with the approach proposed above (of semantically interoperable query of local database from the cloud), then the results returned are in real time, based on live data. For precision medicine, this offers live data for real world evidence, recruitment for acute indications and other time-dependent actions. Patient data can be clustered for analysis using the full scope of available information, and patients can be matched with new drugs immediately on entry into hospital databases. Then we can unlock the true potential of precision medicine.

References

1. Precision medicine market size by technology, by application – Industry Analysis Report: Regional outlook, application potential, price trends, competitive market share & forecast, 2016-2023, Global Market Insights: July 2016
2. Global precision medicine market – Estimate and forecast through (2015-2022), *Wise Guy Reports*: September 2016
3. Global precision medicine market growth – Trends and forecasts (2017-2022), *Orbis Research*: January 2017
4. Geetter JS *et al*, 21st century cures: Modernizing public and private research, *The National Law Review*: December 2016
5. Murphy P, As US R&D languishes, China pushes precision medicine envelope, Bloomberg: January 2017
6. Nikulski N, The Precision Medicine Group Basel Area is launched, Switzerland Innovation Park Basel Area: February 2017
7. Qatar Biobank paves way for precision treatment, *Gulf Times*: February 2017
8. Visit: www.pharmexec.com/what-precision-medicine-means-rare-diseases
9. Garvin E, 3 keys to achieving clinically actionable precision medicine for providers, *HIT Consultant*: January 2017
10. Visit: www.eugdpr.org
11. Visit: www.fra.europa.eu/en/publication/2014/handbook-european-data-protection-law

About the authors

Clinerion's **Tigran Arzumanov,** PhD, MBA, has been working in the area of Clinical Research Sales and Marketing since 2004 and has extensive experience in commercial aspects of both pharma research services and eClinical solutions.

Email: tigran.arzumanov@clinerion.com

Le Vin Chin is Head of Marketing and Communications at Clinerion and has been working in communications and marketing for 20 years in a wide variety of industries, including software and services.

Email: levin.chin@clinerion.com