



Harnessing big data to drive up quality of care

Pooling data that tell the unique story of each cancer patient reveals patterns that could help us learn about which treatments work best for which patients in everyday clinical practice – and about which clinics stick most closely to clinical guidelines. **Marc Beishon** reports on a US initiative that hopes to do just that.

Where do you turn when you have a patient with a rare cancer for which there are a number of possible treatment options? What if current guidelines do not say which is best, or there are no guidelines that are relevant to patients like yours – similar age, gender, health status, treatment history? The literature could reveal some relevant cases and

guidance, and colleagues can be consulted for opinions. But that could be hit or miss and add little to your own experience. And there is no time to lose.

It's a scenario that CancerLinQ (<https://cancerlinq.org>), an ambitious project in the US, aims to address by pooling the 'real time' experience of treating millions of cancer patients in a 'big data' computer system.

Any oncology practice can log in and search for patients with profiles similar to theirs, and look up how they were treated and what their outcomes were.

The project is billed as a 'rapid learning system' for cancer, with a primary mission to improve the quality of everyday oncology practice across the country. It can also be used to test hypotheses for clinical trials,

generate new clinical guidelines, and bring in results from trials, registry data and patient reported outcomes.

CancerLinQ is one of a series of initiatives by the American Society of Clinical Oncology (ASCO), and, according to ASCO's medical director, Richard Schilsky, the most challenging and potentially transformative. "In recent years we've moved from being mainly reactive to starting projects that we believe will change cancer care," he says. "These include the ASCO Value Framework [which assesses the value of new cancer therapies], our first ever clinical trial – a genomic matching study called TAPUR [Targeted Agent and Profiling Utilization Registry] – and CancerLinQ, which we started in 2012.

"We've taken the view that we can do much more than the usual dissemination of information through journals and meetings, by helping to change practice, and to some extent research, by filling gaps in knowledge that can't be filled by the traditional mechanism of the prospective clinical trial."

As Schilsky adds: "The only way we have learnt anything in oncology is by conducting clinical trials, but we have to recognise that only a small percentage of adult patients participate in them in the US, and there is only a certain number of trials that can be completed with limited resources. Yet every day we continue to treat patients whose information is never available to the wider oncology community, because there is no mechanism to collect, analyse and learn from it."

The goal of CancerLinQ, he says, is to aggregate and analyse data from millions of cancer patients in the US and also potentially from centres abroad, to identify new areas of research, but especially to improve care by feeding back to oncologists information on how well they are serving patients, according to quality guidelines. As Schilsky points out, there is a drive in the US and in other countries to monitor and improve healthcare quality, but existing methods have reached their limits.

Quality control in real time

"For example, for the past decade ASCO has run the Quality Oncology Practice Initiative, QOPI, which has been successful, but it is retrospective, as it is based on manual extraction of data from patient charts done at the practices – it gives a sense to oncologists on what they were doing rather than what they should be doing. CancerLinQ is taking QOPI and converting it into an electronic prospective system with built-in quality measures that can be reported back to doctors nearly in real time." Apart from insights from the 'big data', those quality measures will tell oncologists if they have documented factors such as pain and carried out tests such as for HER2 within certain times, and compare their day-to-day performance with that of their peers.

CancerLinQ is designed to take data from most electronic health or patient record systems that are in use in oncology practices and cancer centres. Although these patient record systems have had a tough gestation in the US, they are now in wide use. Schilsky says about 90% of the several thousand oncology practices in the US have a record system suitable for integration – and apart from some interconnection work, there is no more that needs to be done to upload data to CancerLinQ, as the practice will be collecting it anyway.

The data collected comprises both structured information – such as the pathology and treatment of the cancer, and mandatory reports like standard scales of pain and emotional distress – and also unstructured data, which are mainly the notes that accumulate for each patient. Essentially, it collects the whole patient record.

Initially, the structured data are the easier to analyse, and in this early stage of the project there are several examples that show its potential for what many will no doubt see as the more exciting side of the project, as demonstrated at ASCO's annual meeting in June, which was billed as being about 'collective wisdom'. At this point, says Schilsky,

there were about 750,000 patient records in CancerLinQ, but only 130,000 had been 'cleaned up' sufficiently for analysis.

But even using this limited data base, he says, his team has been able to select a cancer that is uncommon and which most oncologists are unlikely to have much experience with – namely male breast cancer – and pull out 350 records. "That's one of the largest series of cases anyone's seen, and we were then able to ask a straightforward question – what treatments were administered to those men? – and we were plotted a histogram of those treatments against patients who received them. If you're an oncologist who hasn't seen a case for ten years or more, you can see what your colleagues are doing now in a couple of mouse clicks." The system is presented to oncologists as a 'dashboard' interface, he says.

And they also looked at another trial that made the news at the conference. "There was data reported on a prospective randomised controlled trial on metastatic colon cancer, and outcomes depending on whether the tumour was on the left or right of the colon," says Schilsky. "So we pulled out all the colon cases in CancerLinQ, looking at the side and treatments given, and found that, regardless of the location, most patients got bevacizumab [Avastin] in addition to chemotherapy, yet the trial indicates that left-side colon cancers did better with cetuximab [Erbix] and chemotherapy. So now we are in a position to see whether oncologists will shift to cetuximab on the left side. This kind of analysis will inform us about the dissemination of research results into practice."

LinQing up

ASCO also announced that 58 oncology practices have signed up for CancerLinQ. These are mostly smaller, outpatient practices rather than the comprehensive and academic cancer centres – Schilsky says they have fewer patient numbers and

could have more to gain, while the larger centres tend to take a proprietary view of their data and have more bureaucracy to navigate. But all centres should have just as much interest in analysis of current practice, and he is optimistic that most will sign up over time. Currently, there is no charge for signing up to CancerLinQ, with much of the \$40 million or so spent so far coming from ASCO's Conquer Cancer Foundation.

One of the practices is Michiana Haematology Oncology, which has six locations in northern Indiana. Robin Zon, a senior partner and a medical oncologist, says it has been difficult to implement an electronic patient record system, but it became much easier once a shared platform among a network of local institutes was set up. As she explains, like many US practices, Michiana has certain expertise – it mainly carries out medical oncology and radiotherapy, so most surgery and pathology data needs to come from other providers. “We are certified for ASCO's QOPI, which means that we cannot treat without a pathology report for each diagnosis, for example,” she says.

Like most cancer practices, Zon and colleagues run tumour boards, some 120 a year, and use guidelines, principally from the National Comprehensive Cancer Network (NCCN), but as she says, while primary treatments can be straightforward, recurrent and metastatic cases often have many options and little guidance on what to use. “CancerLinQ will give us additional data points on how past patients have behaved when they have certain parameters that take in real life experience, not just those in trials. If there are two patients with similar pathology, stage, age and gender, they might be different in other ways, such as other medication, co-morbidities and certain blood chemistry, and we may then be able to differentiate them by treatment.”

Zon can't emphasise enough that basing treatment decisions on clinical trials from just 3% of the population is like comparing apples with oranges, and applying trial

results to an often older group of patients can soon show that the treatments are not appropriate (or indeed can be a “nightmare” as she puts it).

She adds that oncology centres can use different therapies that are seen as equivalent based on their experience and culture, mentioning chemotherapies given to lymphoma patients about to undergo bone marrow transplants. “We use a different regime to the Mayo Clinic, where one of our patients has recently gone for a transplant,” she says. “We use ‘collective wisdom’ at our tumour boards, but this is mostly not based on documentation. Of course we are using precision medicine where we can, but this is for a minority of patients in some cancers. In lung cancer, for example, the new genomic targets are not found in most, so what do we give them? CancerLinQ will help show whether what we are doing is correct or maybe way off – I anticipate using it at our boards to inform our recommendations.”

The IT challenge

There are a lot of administrative and technical issues that have had to be solved with CancerLinQ, such as anonymising the data and ensuring data are collected for the same person over time, so that comprehensive comparisons can be made about when, how and where they were treated, and how they fared at least for five years. Given that in the US there is no national patient identifier code, this is a big challenge. Schilsky explains that most standard data about a cancer is likely to be in oncologists' systems, given they can't practise without it, but pulling in data from primary care, and other specialists such as cardiologists who provide care during the cancer journey, is more a vision than reality at present.

He mentions that ASCO is now collaborating with the American Society of Radiation Oncology (ASTRO) to include

data from its practices; another collaboration is with the Cancer Informatics for Cancer Centers (CI4CC). “And we also want to build in patient-reported outcomes using a mobile phone app that allows people to link how they are feeling with their patient record. That will greatly enrich the data.” Zon, who leads ASCO's pathways task force, says CancerLinQ can be seen as part of the bigger quality picture of clinical cancer pathways, which should aim to manage care from diagnosis to end of life.

A European LinQ?

Is there anything like CancerLinQ in Europe? Probably not at present, as healthcare systems and cancer centres are developing different tools according to various priorities for quality and research. But ‘big data’ is a common theme. A recent paper, ‘Unlocking the treasure trove of information in cancer registries’, which focused on improving outcomes in prostate cancer, pointed to the trend for population-based cancer registries to merge with clinically-based registries as an important ‘direction of travel’ (*Eur Urol* 2016, 69:1013–14).

Healthcare managers may want to receive up-to-date metrics on factors such as population needs, waiting times and quality of care, while oncologists and researchers will in future also be served with the broad amalgamation of data on incidence and survival from databases, increasingly enriched with patient-level information (see also ‘Explaining Europe's survival gaps’, *Cancer World* May/July 2016).

In Germany, for example, the ‘Klinische Datenintelligenz’ (KDI, clinical data intelligence) project, funded by the Federal Ministry for Economic Affairs and Energy, is developing systems that can provide a single view of all data collected from a patient, not just cancer – although breast cancer is one of the first applications mentioned in a paper (The Clinical Data Intelligence Project,

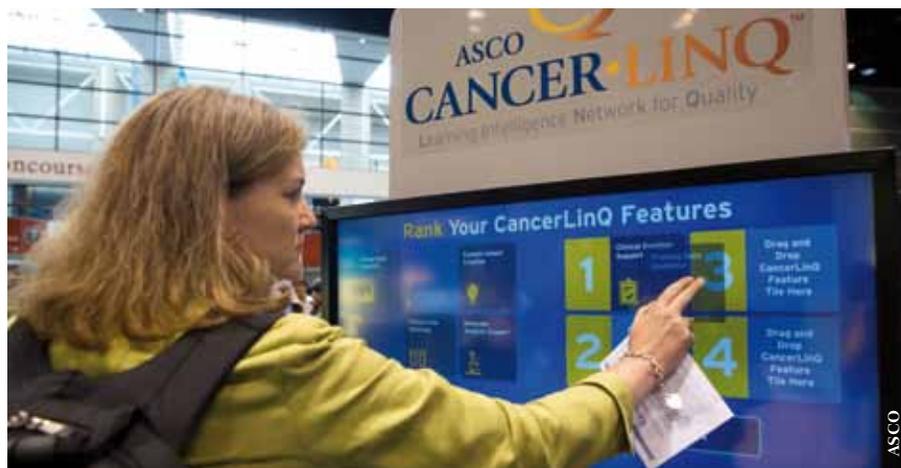
Informatik Spektrum 2016, 39:290–300). It is described as “the first German medical data intelligence initiative where clinical data is tried to be turned into smart data for clinical decision support”, with sources including the Bavarian Breast Cancer Cases and Controls database.

Peter Fasching, a gynaecologist and cancer specialist at Erlangen University Hospital in Bavaria, who is involved with KDI, comments that the US CancerLinQ is an advance, especially for analysing large amounts of data “to find a population similar to the patient sitting in front of you”, given that, increasingly, patients are part of smaller groups as treatment becomes more personalised. He says that the integration of imaging, molecular and biobanking data will be the next step for decision support in personalising treatment, which is what KDI is investigating.

He adds that oncologists don't need to wait for national initiatives – they can start their own databases. “For example, in Germany a group around the country has built one of the largest real-time registries on metastatic breast cancer, called Praegnant [www.praegnant.org] – it is helping us to scrutinise clinical and molecular data to improve patient care right now.”

There are many other projects around Europe. Sweden, for example, has developed a real-time reporting system for its national prostate cancer registry, which oncologists can use to compare data among all of Sweden's 21 counties. “Data include waiting times between referral and first consultation, time between biopsy procedure and cancer information, selection of treatment, surgical outcome (positive margins) and many other pertinent aspects of cancer care” (see *BMJ Qual Saf* 2014, 23:349).

There are certainly a lot of claims being made. The Netherlands Comprehensive Cancer Organisation (IKNL) says the country's national registry (NCR) will include a tumour-specific dataset, and more data will also be gathered about the course of the disease, “thus making the NCR a



A flexible tool for clinicians. An ASCO delegate checks out the different functions CancerLinQ can offer

continuous patient follow-up system – unique in the world”.

Jem Rashbass, of Public Health England's National Disease Registration and Cancer Analysis Service, says the UK is probably the closest to achieving something similar to CancerLinQ – it has the great advantage of integration among all tiers of healthcare, and a much smaller number of oncology centres. In 2013, Public Health England, which had assumed responsibility for all the English cancer registries and the National Cancer Intelligence Network, announced it would develop the world's largest single database of cancer patients.

“We have made great progress since 2013,” says Rashbass, “and now pool nearly all cancer-related-data – referral pathway, screening, pathology, molecular diagnostics, imaging, multidisciplinary discussion, radio- and chemotherapy – at a record level on all patients diagnosed with cancer across England, and from the end of the year Wales. This will be around 32 million records, on over 500,000 tumours, this year.

“We are now about to test feedback to clinicians that will provide them with a view of all the information we know about the patient sitting in front of them. It is an interactive infographic of the whole medical record for that patient. In time we expect to use machine learning algorithms to infer

possible outcomes for individual patients.”

A typical example he gives for how the system in England will work is creating an aggregate view of patients such as older women with breast cancer, to provide details on outcomes, mean time to relapse, and added benefits of adjuvant therapies. A dashboard will also show all the health-related events for a particular patient.

“But there are some datasets that we are working on where I feel we don't have enough information. Our collection of molecular data is limited to about 30 markers at the moment; we need a better assessment of co-morbidity – we are about to link primary care prescription data to do this, and we need to be better at identifying and capturing recurrence and relapse.

“The challenge for all of us is scale and data quality. In comparison with the US potential, our scale in the UK will inevitably be smaller, but we do cover all the 55 million people in England. We are obsessed with data quality, because if personalised medicine is really to deliver, we don't want to spend time tracing spurious data anomalies, so we have around 150 cancer registration staff collating and quality-assuring the data.”

This seems to be a well-balanced big data ‘arms race’ across the Atlantic, which can only be good news for the quality of cancer care.