Make precision medicine work for cancer care

To get targeted treatments to more cancer patients pair genomic data with clinical data, and make the information widely accessible, urges **Mark A. Rubin**.

Ten months ago, the physicians of a feisty 76-year-old sales clerk from New Jersey who had an advanced carcinoma in her urinary tract decided to try an unconventional therapy. A few weeks earlier, they had sent a sample of her tumour to my team at the Institute of Precision Medicine at Weill Cornell Medical College and NewYork-Presbyterian Hospital in New York City. Genetic sequencing had revealed that she had more copies than usual of the *HER2* gene (also known as *ERBB2*)[1,2].

After years of failure with the usual arsenal of surgery, chemotherapy and radiation, the physicians included the drug Herceptin (trastuzumab) in the woman's treatment. Herceptin is more commonly used for breast cancer, but it targets the *HER2* mutation. Since taking the drug, she has been free of disease.

Advances in sequencing have dramatically increased the likelihood of discovering mutations that drive tumour growth in certain people and in certain tumours — even in specific cells within tumours. Yet mountains of genomic data are accumulating that are of little use because they are not tied to clinical information, such as family medical history. What is more, genomic data are generally confined to documents that cannot easily be searched, shared or even understood by most physicians.

To achieve the level of success in precision medicine for cancer care that US President Barack Obama and others are anticipating,

sequence data needs to be linked, in real time, to the patient sitting in front of his or her doctor. Integrated genomic and clinical data will also need to be available, in a searchable way, to a broad community of practitioners and researchers. Prototypes for centralized data banks are showing promise, but serious and sustained investment is needed to scale them up.

COMPLEX RECORDS

Clinicians are used to appraising 20–50 measurements from routine laboratory tests, such as for blood-sugar levels. Such data can be easily entered into patients' electronic health records. Genomic data introduces a whole new level of complexity.

To give an idea of the scale, it would take more than 25 days to transfer from one computer server to another the 2.5 petabytes (a petabyte is 1,000 terabytes) of data generated by The Cancer Genome Atlas — a US project started in 2005 to catalogue the mutations that drive cancer. This is according to my colleague Toby Bloom, deputy director for informatics at the New York Genome Center, a consortium that specializes in large-scale human genome sequencing.

Hugely complicated genomic reports are rarely available in electronic form and are seldom tied to basic information about the patient. Whole-genome sequencing on tumour samples from nearly 14,000 people by the International Cancer Genome Consortium (ICGC), for instance, has revealed nearly 13 million mutations across the genome. But numerous factors aside from the mutations in a person's DNA will affect whether any one patient will respond to a particular treatment. Unfortunately, in the ICGC effort — and many like it — only the most minimal of clinical data, such as type and size of a tumour, are available (see 'Missing metrics').

Since 2013, working with a team of computational biologists from Weill Cornell and the Centre for Integrative Biology at the University of Trento in Italy, my colleagues and I have conducted a pilot programme to determine the feasibility of tying genomic to clinical data in real time. So far, we have created easy-to-read reports for 250 people with cancer.

Each report carries a barcode, allowing patients to be de-identified and re-identified as needed, and is designed to be integrated easily into the electronic health-records system of the NewYork-Presbyterian Weill Cornell Medical Center. The data, which are presented much like pathology results, capture clinical information (family history, medication use and so on), information about mutations for which specific drugs exist, and findings about genetic anomalies with unknown effects.

We have discovered that more than 90% of our patients carry a mutation that may be responsive to a known drug — although less than 10% of the patients may be eligible for a clinical trial either for logistical reasons or because there is insufficient evidence to warrant trying a non-approved drug.

To be useful more broadly, these data need to be sharable across institutions. Take, for instance, current efforts to investigate the efficacy and safety of the drug neratinib in patients whose tumour growth is driven by various mutations in either *HER2* or *EGFR*[3]. Aside from lung cancer (in which *EGFR* mutations are common), the frequency of these mutations is in the range of 1–6%, so achieving the numbers required for a phase II clinical trial has meant recruiting patients from multiple medical centres. Sharing data across institutions could dramatically increase the ease and efficiency of recruitment for such trials — currently a frustratingly slow process that is largely dependent on word of mouth.

Yet the barriers to achieving this type of sharing are formidable. In the United States, incompatible electronic systems make transferring patient records between facilities extremely difficult — often requiring the shipping and scanning of printouts.

DIGITAL DATA

Various initiatives are trying to address the creation of standards for communal digital medical data. One example is the non-profit New York City Clinical Data Research Network (NYC-CDRN). Funded by the Patient-Centered Outcomes Research Institute in Washington DC, this non-governmental organization is bringing together 22 institutions, led by the Weill Cornell Medical College and NewYork-Presbyterian Hospital, to document and manage clinical data[4].

Sixteen months in, the NYC-CDRN has more than 6 million records with hundreds of thousands of data elements, ranging from simple measurements of, say, calcium levels in the blood, to the results of magnetic-resonance-imaging scans. The ultimate goal is to include genomic data in the database and to follow patients longitudinally. Particularly in countries with private health-care systems, centralized 'warehouses' of shared, standardized, searchable patient data may be the most feasible way forward.

The promise of precision medicine for cancer is now clearly evident. For instance, drugs that target *BRAF(V600E)* mutations (seen in around 60% of melanomas) and *IDH1* or *IDH2* mutations (seen in around 80% of brain tumours) have either been approved or are undergoing testing in

clinical trials[5,6] — although, as with most targeted therapies, resistance is a major problem[7]. And in one of the most ambitious precision-medicine trials ever conducted, which is taking place at multiple institutions in France, 141 patients out of the 708 enrolled have already been matched to targeted-therapy trials[8].

MONEY MATTERS

Yet the 'precision' approach raises some hard questions. The more patient-specific information included in centralized databases — crucial to the long-term success of precision medicine — the harder it will be to ensure contributors' anonymity. What rights should people have over their own health data? Should such data be shared internationally? Also unclear is who should manage and sustain such data warehouses, and who should pay for them.

The NYC-CDRN has already cost US$7 million, and annual costs will increase as more information is collated. This adds to the considerable expense of the treatments themselves — annual costs for the targeted therapies in cancer now available generally exceed $100,000, and most extend patients' lives by only months.

Should targeted drugs for patients with mutations found in only 10% of the

MISSING METRICS

For much of the genomic data obtained for nearly 14,000 patients by the International Cancer Genome Consortium, key clinical information is missing.

Legend: Missing / Collected

Metric	Missing %
Proportion of tumour cells in sample	92%
Relapse type	92
Length of disease-free interval	91
Tumour stage at diagnosis	81
Survival time	77
Clinical tumour stage	76
Tumour grade	76
Age when sample was taken	72
Type of tissue	70
Disease status at last follow up	29
Age at last follow up	22
Vital status (alive or dead)	9
Diagnosis	8
Age when diagnosed	6
Sex	2

population be developed and used if they extend survival by just three months, say? Should drugs be supported only if they extend people's lives for at least one year?

To complicate things, the full benefits of many drugs may become apparent only after they have been approved. Herceptin, for instance, was initially approved by the US Food and Drug Administration as a treatment that can extend the survival of people with a certain advanced metastatic breast cancer by months[9]. Increased use of the drug has since revealed that it can improve the chances of long-term survival for people with earlier stages of breast cancer[10].

"Incompatible electronic systems make transferring patient records between facilities extremely difficult"

Some organizations have already given guidance on the rationing of precision treatments. In the United Kingdom, the National Institute for Health and Care Excellence (NICE) examined data on the usefulness of different types of genomic test in the treatment of breast cancer. In September 2013, NICE recommended a test called Oncotype DX for clinical decision making but determined that three other genomic tests currently available (MammaPrint, IHC4 and Mammostrat) be used only in research because of insufficient evidence supporting their usefulness in clinical care.

There are many reasons for hope. But turning the wealth of insights potentially available from genomics into targeted treatments for cancer will require difficult decisions and the costly, laborious task of creating shared and searchable information. ∎

Mark A. Rubin *is professor of oncology in pathology and director of the Institute for Precision Medicine at Weill Cornell Medical College and NewYork-Presbyterian Hospital in New York City, New York, USA.*
e-mail: rubinma@med.cornell.edu

1. King, C. R., Kraus, M. H. & Aaronson, S. A. *Science* **229,** 974–976 (1985).
2. Hudziak, R. M., Schlessinger, J. & Ullrich, A. *Proc. Natl Acad. Sci. USA* **84,** 7159–7163 (1987).
3. Rabindran, S. K. *et al. Cancer Res.* **64,** 3958–3965 (2004).
4. Kaushal, R. *et al. J. Am. Med. Informat. Assoc.* **21,** 587–590 (2014).
5. Flaherty, K. T. *et al. N. Eng. J. Med.* **363,** 809–819 (2010).
6. Wang, F. *et al. Science* **340,** 622–626 (2013).
7. Ramos, P. & Bentires-Alj, M. *Oncogene* http://dx.doi.org/10.1038/onc.2014.314 (2014).
8. Massard, C. *Ann. Oncol.* **26,** (suppl. 2) ii4 (2015).
9. Slamon, D. J. *et al. N. Eng. J. Med.* **344,** 783–792 (2001).
10. Slamon, D. *et al. N. Eng. J. Med.* **365,** 1273–1283 (2011).