

Personalised medicine and genetic prediction – are we there yet?

Miles Parkes

ABSTRACT – A working definition of personalised medicine is the delivery of a tailor-made treatment to the right patient at the right time. How close have recent advances in genetics come to realising this in the clinic?

KEY WORDS: Prediction, stratification, genome-wide association scanning (GWAS), biomarker

Introduction

The paradigm of personalised medicine based on genetic insights has been established in oncology for some years. Among the first examples, demonstrating the feasibility of moving from molecular characterisation of a pathogenic defect to a specific therapy that targets this defect, was the Philadelphia chromosome in chronic myeloid leukemia (CML). The translocation between chromosome 9 and 22 leads to the formation of a chimaeric BCR–ABL fusion protein with tyrosine kinase activity. The latter has the unfortunate effect of driving proliferation of the CML clone – but its kinase activity can be inhibited by the small molecule therapy imatinib. The latter and related compounds now form part of the therapeutic armamentarium against CML.

The concept of stratification of a clinically phenotyped patient cohort, which is also a key tenet of personalised medicine, has also been proven in oncology. One of the earliest and best known examples here relates to expression of the HER2 (human epidermal growth factor receptor 2) gene in breast cancer. Approximately 20% of patients with breast cancer have a mutation in HER2 which leads to its over-expression on the surface of malignant cells – rendering these more aggressive and less responsive to hormonal therapies. This can be detected by immunohistochemistry or fluorescence in situ hybridisation (FISH), and patients who are HER2-positive can be effectively treated with antibody therapies that target the extracellular domain of HER2, such as trastuzumab. This treatment is expensive and substantially less effective in HER2-negative patients – bringing forward the concept of stratification of breast cancer patients and targeting of trastuzumab.

A number of other examples have also derived from the field of oncology (*ALK* mutations in non-small cell lung cancer, *BRAF* mutations in melanoma and so on), but the question arises as to how applicable these concepts are to common, non-oncological disease. The rapid progress that has been made particularly by genetics studies in furthering our understanding of the pathogenic basis of common inflammatory, metabolic and cardiovascular

diseases has led many to ask this question. Another question is whether molecular genetics can help in the allied issues that come under the personalised medicine umbrella – namely risk prediction, genetic diagnosis, choice of optimal treatment through pharmacogenetics, and molecular biomarkers of prognosis. The field is still some way off widespread application of genetics at the clinical coalface, but nonetheless progress is being made in each of these domains, as discussed below.

Insights from genome-wide association scanning

Much of the progress made in genetics has been based on the technology of genome-wide association scanning (GWAS). This technique delivers an unbiased survey of the whole genome for regions associated with disease susceptibility. It is achieved by genotyping upwards of 0.5 million markers genome-wide in panels comprising thousands of cases and controls, and using stringent statistical thresholds to identify ‘associated’ single-nucleotide polymorphisms (SNPs) for which there is a significantly different allele frequency between cases and controls. Genes mapping within the linkage disequilibrium block under the association signal are identified, and then (if there is more than one) filtered using a variety of informatics techniques to identify causal genes and refine causal variants. In the ideal world the next steps should include an assessment of the impact of the associated gene variant on protein function in relevant cell type, but for most diseases such studies are complex (even to define which cell type is relevant can be challenging) and on-going.

Among the earliest published studies, the Wellcome Trust Case Control Consortium in 2007¹ set the bar for statistical interpretation of GWAS data. It gave an early steer that Crohn’s disease and other immune-mediated diseases would be particularly tractable to such approaches, and provided evidence of overlap (‘pleiotropy’) of signals across inflammatory phenotypes. Subsequent years have seen exponentially increasing numbers of GWAS-associated variants identified across the spectrum of common, complex disease. In inflammatory bowel disease the total currently stands at 163 independent loci meeting genome-wide significance² – some being specific for Crohn’s disease or ulcerative colitis, but most being shared between these phenotypes. Undoubtedly the results of such studies have provided profound new insights into pathogenic mechanisms and environmental interactors, and potential targets for new drug therapies. But do they facilitate personalised medicine?

Risk profiling

One area where one might expect increasing knowledge of the genetic basis of disease to impact is in terms of risk prediction.

Miles Parkes, consultant gastroenterologist
Addenbrooke’s Hospital and University of Cambridge

Although clinicians may prefer to think about this in relation to their own disease areas of interest, the reality is that risk profiling runs genome-wide and generates risk scores for all phenotypes. This spectrum includes disease risk, pharmacogenetic markers and SNPs associated with non-medical attributes as diverse as hair colour, athletic prowess and ear wax consistency. A number of commercial companies now offer genetic risk prediction, running GWAS arrays and risk prediction algorithms in return for payment of ~\$200 upwards. The accuracy of the method remains low for most phenotypes based on current knowledge and the fact that not all risk variance is attributable to genetic factors. Analysis protocols require both refinement and inclusion of known environmental risk factors, and these and related issues represent areas of active endeavour by the research community. Such topics are among the sources of discussion on blog sites such as Genomes Unzipped (www.genomesunzipped.org). Importantly, mindful of the potential for anxiety and harm of genetic data, the 'direct-to-consumer' user interfaces offer appropriate safeguards against 'accidentally' viewing data relating to relatively highly penetrant variants associated with conditions with significant medical impact. Examples of such sensitive data include carriage of the ApoE4 variant (associated with Alzheimer's disease) or BRCA1 (for breast cancer).

Available evidence from physiological phenotypes such as height³ and disease phenotypes such as rheumatoid arthritis⁴ suggests that use of all data genome-wide (including SNPs which are modestly associated but do not reach traditional genome-wide significance threshold of $P < 5 \times 10^{-8}$) substantially increases the proportion of phenotypic heritability explained – and will therefore contribute to the accuracy of risk prediction. This probably reflects the contribution of several hundred or more variants to many such phenotypes, each having an effect size so small as to be difficult to discern on marker-by-marker association tests. As ever larger datasets are analysed by GWAS so more of these signals are identified, but it remains likely that many will remain forever below the threshold of detection. Accurate risk prediction using genetic factors alone will remain challenging, particularly for relatively uncommon diseases due to the low pre-test probability and hence low positive predictive value of the test. Furthermore, its relevance hinges on availability of effective preventative strategies in 'at risk' individuals.

Published data suggests that early adopters of direct-to-consumer tests are well-educated about genetics and that genetic risk predictions do not increase anxiety. However, nor do they magically change negative behaviours (although they do increase *intention* to change!). Approximately a quarter of those surveyed for one publication had discussed the results of their test with their doctor⁵ – a sign of things to come for many physicians.

Genetic diagnosis

Does molecular genetics offer anything useful with regard to the diagnosis of common disease? The answer is 'no', except perhaps in relation to extreme Mendelian forms. These will most

commonly present to paediatricians, particularly in the context of parental consanguinity. There are increasing numbers of reports in the literature where exome sequencing has proved helpful in characterising the defective pathway and giving a steer regarding therapeutic possibilities. A recent example relating to inflammatory bowel disease led to the identification of mutations in the interleukin 10 receptor gene as causing devastating refractory intestinal inflammation in infants, mandating allogeneic stem cell transplantation, which proved curative.⁶ However, although the sequencing technology itself is increasingly routine, the informatic and analytic challenges such approaches present should not be underestimated. For the most part they remain in the research domain at present.

Optimising treatment through pharmacogenetics

A potentially more generally applicable field for predictive and personalised medicine relates to pharmacogenetics – in particular, whether genetic markers can predict response and intolerance to existing drug therapies. This would be a chapter topic in its own right, and will only be considered briefly. Already in inflammatory bowel disease we routinely use assays of thiopurine methyltransferase (TPMT) to guide therapy with azathioprine and 6 mercaptopurine. There are a number of polymorphisms in TPMT which affect the enzyme's activity and hence risk of bone marrow over-suppression – and which should be assayed biochemically or genetically prior to treatment initiation to reduce the risk of toxicity. Recent evidence relating to an HLA genetic marker of flucloxacillin hepatitis⁷ suggested that other drug toxicities may also be genetically mediated. How generally this (and indeed drug response) is genetically influenced or mediated is unclear. One problem is that detecting more such pharmacogenetic effects will require large-scale new GWAS studies for each drug, using populations defined by drug response, non-response or specific toxicities. Of note, there is no prior expectation that disease susceptibility genes detected by the conventional GWAS studies will correlate with pharmacological outcomes.

Genomics in the prediction of disease course

The final key area of genomics application to personalised medicine that will be considered relates to the identification of prognostic biomarkers. Again the paradigm has been established in oncology, where specific transcriptional profiles have been identified which correlate with disease prognosis. Several have been patented, and the 'Mammaprint' RNA microarray is, for example, marketed as a tool for stratifying patients with node negative breast cancer into high versus low risk groups for distant recurrence – influencing decisions regarding need for post-operative chemotherapy. There are currently no equivalents in inflammatory disease but recent data from Cambridge suggest the potential utility of such approaches. By blood sampling at diagnosis, separating leukocytes and undertaking transcriptional profiling on the CD8+ subset, over-lapping signatures have been

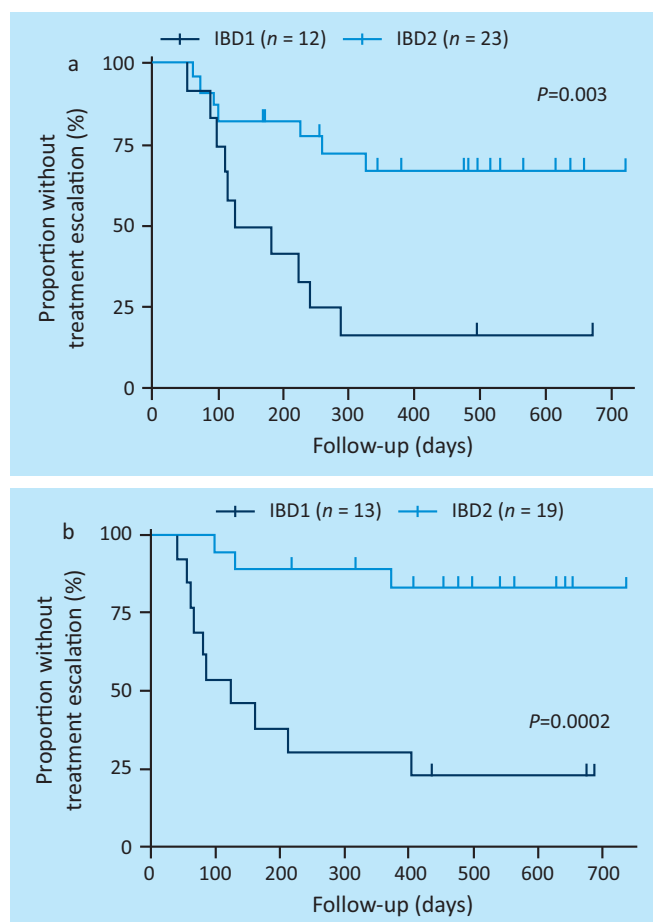


Fig 1. Genetic biomarkers in inflammatory bowel disease prognosis. Kaplan-Meier survival curves demonstrating the proportion of (a) Crohn disease and (b) ulcerative colitis patients who did not require a subsequent treatment escalation (immunomodulator or surgery) after enrollment as stratified by IBD1/2 subgroup. IBD = inflammatory bowel disease. Reproduced with permission from Lee *et al.*⁹

identified for SLE, ANCA positive vasculitis, Crohn's disease and ulcerative colitis that correlate tightly with subsequent disease course (Fig 1).^{8,9} The high risk signature in inflammatory bowel disease (IBD) appears to predict much more frequent disease relapse and significantly greater need for escalation to heavy immunosuppression compared to individuals who carry the low-risk profile. The next goal is to test in a formal CD8+ biomarker stratified drug trial whether 'top-down' treatment with anti-TNF therapy from diagnosis in individuals with the high-risk signature produces better clinical outcomes.

Conclusions

These are early days for genomics at the clinical coalface, and currently the impact of these new tools is limited. However, there

are a number of factors which will drive the implementation of genomics approaches in general and personalised medicine in particular. These include the dramatic decrease in the cost of DNA- and RNA-based technologies – particularly sequencing, relevant not just to detecting germline variation in DNA but also the detection and quantitation of gene expression through 'rna-seq', epigenetic analysis, microbial and microbiota characterization among many other applications; direct-to-consumer marketing of genomics products; and the application of these new technologies in a framework of experimental medicine by a handful of avant-garde clinicians.

In summary, genetics studies through GWAS have provided many new insights re pathogenic mechanisms in complex disease, which was their goal. The application of genomics technologies to personalised medicine is already a reality in oncology and technology will drive other areas of medicine to catch up. It seems likely that even the most reluctant of physicians will be swept along in this tide as it gathers momentum – and it is incumbent on us all to welcome the possibilities that these new tools offer and embrace their development as a means to optimise clinical care of our patients.

References

- Wellcome Trust Case Control Consortium. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:132.
- Jostins L, Ripke S, Weersma RK *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119–24.
- Yang J, Benyamin B, McEvoy BP *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42:565–9.
- Parkes M, Cortez A, van Heel D, Brown M. Genetic insights into common pathways and complex relationships among immune mediated diseases. *Nat Rev Genet* 2013;39:1329–37.
- Bloss CS, Schork NJ, Topol EJ. Effect of direct-to-consumer genome-wide profiling to assess disease risk. *N Engl J Med* 2011;364:524–34.
- Glocker EO, Kotlarz D, Boztug K, *et al.* Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med* 2009;361:2033–45.
- Daly AK, Donaldson PT, Bhatnagar P *et al.* HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet* 2009;41:816–19.
- McKinney EF, Lyons PA, Carr EJ *et al.* A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. *Nat Med* 2010;16:586–91.
- Lee JC, Lyons PA, McKinney EF *et al.* Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn's disease and ulcerative colitis. *J Clin Invest* 2011;121:4170–9.

Address for correspondence: Dr M Parkes, Addenbrooke's Hospital and University of Cambridge, Cambridge CB2 0QQ
 Email: miles.parkes@addenbrookes.nhs.uk