

Genomic markers to tailor treatments: waiting or initiating?

Parvin Tajik · Patrick M. Bossuyt

Received: 3 March 2011 / Accepted: 5 April 2011 / Published online: 13 April 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract The decade since the publication of the Human Genome Project draft has ended with the discovery of hundreds of genomic markers related to diseases and phenotypes. However, the project has not yet delivered on its promise to tailor treatments for individuals. The number of genomic markers in clinical practice is very small. The number of markers to guide treatment decisions is even smaller. In order to speed up discovery and validation of genomic treatment selection markers, we call for considering the brilliant potential of randomized clinical trials. If biomedical research community can collaborate in organizing large-scale consortium of clinical trials associated with well-designed biobanks, these studies would soon act as huge laboratories for investigating genomic medicine; a big step forward towards personalizing medicine.

Introduction

A common experience of practicing physicians is that only a subgroup of patients with a specific disease can derive substantial benefit from a particular therapy. It is rare for a drug or treatment option to be safe or effective for everyone. The inherent variability in treatment response among individuals has a significant effect on the quality and cost of health care. Spear and colleagues (2001) have analyzed the efficacy of major drugs for several major diseases based on

published data. They found that the highest percentage of responding patients is between 80 for Cox-2 inhibitors and 25% for cancer chemotherapy. Many of the drugs fall within the range of 50 to 75% response. The safety of treatment options also varies between drugs and diseases.

The conspicuous variability in response, combined with the costs and safety concerns for drugs, make it attractive, or even necessary, to look for so-called treatment selection markers. These are biomarkers that can prospectively identify individuals who are likely to benefit from a specific treatment, separating them from whom in which the more limited health gains do not outweigh the safety and side effects of treatment.

About a decade ago, when the first draft sequence of human genome was published, hopes were raised that genomic information would finally explain the heretofore inexplicable heterogeneity of treatment responses and lead to a revolution in medicine. Francis Collins (2010), one of the leaders of the project at the time, forecasted that in about 10 years there would be genetic markers available indicating a person's risk of heart disease, cancer, and other common conditions, soon to be followed by individually tailored preventive measures and therapies.

This article comments on the present and the future of genomic discoveries for tailoring medicine. We briefly summarize the current state of genomic treatment selection markers. We then discuss the potentials and partly overlooked role of clinical trials in the discovery and validation of genomic markers.

P. Tajik (✉) · P. M. Bossuyt
Department of Clinical Epidemiology,
Biostatistics and Bioinformatics, Academic Medical Center,
University of Amsterdam, Room J1b-214, PO Box 22700,
1100 DE Amsterdam, The Netherlands
e-mail: P.Tajik@amc.uva.nl

Genome-based treatment selection: Where are we now?

Reviewing the list of valid genomic biomarkers in the labels of FDA approved drug products reveals that by now

only 22 genomic markers have entered routine clinical practice. These are some gene-targeted cancer treatments, some predictors of drug toxicities, a few markers of drug metabolism that have proved useful for dose adjustment, and a gene-specific antiviral drug.

This number is quite small, compared with the myriad of diseases, the plethora of existing drugs, and the global health burden. In addition, evaluations show that most of these genomic biomarkers were discovered well before the sequencing of the reference human genome (McDermott et al. 2011; Varmus 2010). In short, it seems that now, after a decade of research, the Human Genome Project has not yet delivered its initial promises. Stratified medicine, let alone personalized medicine, has not yet entered in everyday clinical practice. Unpredictable variability in treatment outcome remains a given in clinical decisions.

A more balanced and less sceptical assessment of the fruits of the Human Genome Project for clinical medicine would come to the conclusion that the discovery of more than 1,100 genetic loci within a few years can be considered as an excellent start. In this short time, genomics have been insightful about, for example, cancers, the molecular basis of inherited diseases, and the role of structural variation in disease (Green and Guyer 2011). By providing a comprehensive scaffold, the human genome sequence has made it possible for scientists to assemble often fragmentary information into landscapes of biological structure and function (Lander 2011).

In order to translate these findings into improved diagnostics and better treatments we now need, over the next decades, intensive functional studies to characterize the genes and pathways underlying diseases. If this materializes, we expect more of the initially promised improvements in the effectiveness of healthcare. Most likely, the initial promise was overly optimistic in terms of timing. As Green and Guyer (2011) wrote in their updated vision of genomic medicine “significant change rarely comes quickly”.

Association of variants at the 9p21 locus with myocardial infarction can be seen as a very good example for showing that genomic research in medicine still has to mature. This association was one of the first discoveries of Genome Wide Association Studies (GWAS) in 2007 and one of the most consistent and robust variant-disease associations in the GWAS era (Schunkert et al. 2008). However, scientists neither have any hypotheses for the function of the locus nor could they explain why it was associated with increased risk of cardiovascular diseases. It was only known that the locus was located in a gene desert, with the nearest protein-coding region several thousand kilobases away. Continued investigations have recently shown an important regulatory function of the 9p21 gene desert in response to inflammatory signalling in vascular cell types,

an important pathway in understanding the disease etiology and its treatment (Harismendy et al. 2011).

From cohort studies to clinical trials

As forecasted in the National Human Genome Research Institute vision for the future of genomics research, the progress during the second decade will be mainly focused on understanding the biology of genomes and diseases. Improving the effectiveness of clinical medicine can only be expected to happen beyond 2020 (Green and Guyer 2011).

Converting scientific discoveries into improved medical tests and more effective treatments for clinical practice have always been slow and difficult. To harmonize and accelerate the efforts for translating genomic discoveries into health care, Khoury and colleagues (2007) proposed a four-phase framework. Phase 1 moves a basic genomic discovery into a candidate health application (e.g., genetic test or intervention); Phase 2 assesses the value of the genomic application for health practice leading to the development of evidence-based guidelines. Phases 3 and 4 introduce evidence-based guidelines into health practice and evaluate the outcomes in the real world.

At present, the driving engine of genomic discoveries in Phase 1 is formed by large-scale collaborative cohorts (Psaty and Hofman 2010) or case-control studies, which have the genotype information of their participants. The output of these studies consists of genomic markers that have robust associations with the studied diseases or traits. There have been similar efforts to identify putative treatment selection markers through studying cohorts of treated patients and finding genomic markers that are associated with the response to treatment. For example, O’Doherty and colleagues (2009) analyzed 61 SNPs in 34 candidate genes as possible determinants of IFN-beta response in 255 relapsing-remitting multiple sclerosis patients. The authors compared genotype and allele frequencies between responders and nonresponders. Yet, the mere identification of allelic combinations that differ significantly between responders and nonresponders do not qualify these as multi-allele treatment selection markers. For that purpose, one has to compare the benefit from treatment for marker-positive patients, identified as such before treatment starts, and marker-negative patients (Janes et al. 2011).

The best way to identify treatment selection markers would be in randomized comparative clinical trials. This assures that the patients who were treated with the agent for whom the marker is purported to be predictive are comparable to those who were not. In a nonrandomized design, it is impossible to isolate any causal effect of the marker on therapeutic efficacy from the multitude of other factors that

may influence the decision to treat or not to treat a patient (Mandrekar and Sargent 2009).

Looking for markers of treatment benefit in randomized trials is an excellent strategy for accelerating genomic marker discoveries, but it is also helpful in validating them for clinical applications. MINDACT, for example, is a trial designed for validating a genomic signature for breast cancer. The trial is expected to enroll 6,000 breast cancer patients to evaluate whether risk assessment by a 70-gene profile can improve risk assessment avoiding non-necessary chemotherapy for the low risk patients (Bogaerts et al. 2006). The study was launched in 2007 and the estimated primary completion date is 2019.

A simple search in Pubmed will show that only about 2% of present research publications in the field of human genomics have been classified by the National Library of Medicine as clinical trials reports, most of which are not randomized trials. A landmark example of a genomic marker that has been successfully discovered and validated using RCTs is KRAS. This genomic marker is the predictor of the efficacy of panitumumab and cetuximab in advanced colorectal cancer. In an analysis of data from a randomized phase 3 trial of panitumumab versus the best supportive care, KRAS status was assessed on archived tumor specimens of about 400 enrolled patients (Amado et al. 2008). This analysis demonstrated a statistically significant advantage in the overall survival for patients with wild-type KRAS on panitumumab, whereas patients with mutant-type KRAS enjoyed no survival benefit. In addition, multiple phase 2 trials consistently demonstrated similar results (Mandrekar and Sargent 2009). The American Society of Clinical Oncology estimates that routinely testing people with colon cancer for mutations in the KRAS oncogene would save at least \$600 million a year, demonstrating how timely translation of genomic markers could increase health care efficiency (Poste 2011).

From individual trials to prospective meta-analysis with biobanking

One important obstacle in using clinical trials for discovery and validation of genomic markers is the large sample sizes needed for such analyses. Few studies have enough participants to provide adequate statistical power, particularly in view of the multiple-testing inherent in genetic analyses. Cohorts and well-conducted case–control studies have successfully adapted to the genomics era by forming large-scale consortia. Some consortia comprise more than a million participants, which helps in overcoming the limitations associated with the large volume of genomic data and in allowing for continued genomic discoveries. The consortia provide the opportunity to prospectively plan meta-analyses and find independent studies for replication of the findings

(Psaty and Hofman 2010). Now it is time for clinical trials to take their role in genomic era and move seriously toward forming large-scale consortia.

A very interesting development is the increasing use of prospective meta-analysis of randomized trials, a form of meta-analysis in which trials are identified, evaluated and determined to be eligible before the results of any of those studies became known (Pfeffer et al. 2003; The PPP Project Investigators 1995). Such collaborative projects allow a priori coordination of data collection efforts, entry selection criteria and end-points, prospective harmonizing of collected data, and preplanned data pooling. Prospective meta-analysis of individual patient data in large RCTs may result in the sample size required to test hypotheses about genomic markers for treatment selection. When the pooling process was planned as an initial step in clinical trials, such as in the Cholesterol Treatments Trialists' (CTT) Collaboration, with about 170,000 recruitments (Baigent et al. 2010), it has succeeded. Major companies have also realized the importance of such cooperation. Merck, Eli Lilly, and Pfizer have planned a joint Asian Cancer Research Group that will help speed up research on new drugs to treat gastric and lung cancers (Lancet Editorial Board 2011).

The presence of RCT registries is an exceptional opportunity for prospective meta-analysis, one that does not exist for other study types. Although trial registration is in place to increase transparency with respect to performance and reporting of clinical trials (ICMJE 2004), it can now also perform as an excellent platform for the formation of the core of large-scale RCT consortia. Registries can serve as search engines for finding suitable studies for the replication and validation of the genomic markers.

Another key necessity for the discovery of genomic treatment selection markers is generating biobanks that accompany clinical trials. Preserving biological specimens from trial patients can be expensive to initiate and maintain, and there may be practical and ethical considerations related to biobanking, but these issues should not discourage investigators from taking on the challenge. Biobanks allow precise future genomic investigations, which may be beyond imagination or technical possibilities at the time of collection. If standardized procedures are used for specimen collection, handling and storage, the anticipated scientific gain from biobanks combined with thorough trial databases would be enormous, deserving the investment (Gustafsson et al. 2010; Poste 2011).

The presence of large RCT consortia associated with well-designed biobanks would have the capacity to perform as a laboratory, allowing for continued discoveries and validation of genomic markers. This approach can potentially save a lot of time and expenses required for conducting trials for validating each newly discovered marker. So, a more feasible and timely option to speed up the translation

process of discovered genomic markers would be using data from previously well-conducted randomized controlled trials. Retrospective validation—like KRAS validation—can aid, when conducted appropriately, in bringing forward effective treatments for marker-defined patient sub-groups in a timely manner that might otherwise be impossible because of ethical and logistical considerations. In particular, if such a retrospective validation can be demonstrated in data from two independent RCTs, this provides a strong evidence for a robust predictive effect (Mandrekar and Sargent 2009). Regarding the expected sudden growth of the number of discovered genomic markers in near future, retrospective use of RCTs seems an intelligent shortcut for keeping the translational research in a proportional pace as the discoveries occur.

Conclusion

The high hopes that genomic information would explain the heterogeneity of treatment responses and revolutionize health care, leading it into an era of stratified or even personalized medicine, have not been met so far. The number of validated genetic treatment selection markers is small in number and applies to a fraction of health care. Yet, there are reasons to remain optimistic. In the years to come, investigators designing and conducting clinical trials have an ethical and financial responsibility to maximize the scientific knowledge gained from the participation of the subjects. Designing biobanks and collaborating with other investigators in forming large data sets would enormously increase the generation of evidence needed for the discovery and validation of treatment selection markers. Even if only a fraction of more than four million subjects enrolled worldwide in interventional trials were to be captured, the potential for progress would be enormous.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Amado RG, Wolf M, Peeters M, Van CE, Siena S, Freeman DJ, Juan T, Sikorski R, Suggs S, Radinsky R, Patterson SD, Chang DD (2008) Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* 26:1626–1634
- Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhalra N, Peto R, Barnes EH, Keech A, Simes J, Collins R (2010) Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* 376:1670–1681
- Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, Bines J, Mook S, Decker N, Ravdin P, Therasse P, Rutgers E, van 't Veer LJ, Piccart M (2006) Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. *Nat Clin Pract Oncol* 3:540–551
- Collins F (2010) Has the revolution arrived? *Nature* 464:674–675
- Green ED, Guyer MS (2011) Charting a course for genomic medicine from base pairs to bedside. *Nature* 470:204–213
- Gustafsson F, Atar D, Pitt B, Zannad F, Pfeffer MA (2010) Maximizing scientific knowledge from randomized clinical trials. *Am Heart J* 159:937–943
- Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA (2011) 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* 470:264–268
- ICMJE (2004) Clinical trial registration: a statement from the International Committee of Medical Journal Editors. http://www.icmje.org/clin_trial.pdf. Accessed 22 February 2011
- Janes H, Pepe MS, Bossuyt PM, Barlow WE (2011) Measuring the performance of markers for guiding treatment decisions. *Ann Intern Med* 154:253–259
- Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L (2007) The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med* 9:665–674
- Lancet Editorial Board (2011) Where will new drugs come from? *Lancet* 377:97
- Lander ES (2011) Initial impact of the sequencing of the human genome. *Nature* 470:187–197
- Mandrekar SJ, Sargent DJ (2009) Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges. *J Clin Oncol* 27:4027–4034
- McDermott U, Downing JR, Stratton MR (2011) Genomics and the continuum of cancer care. *N Engl J Med* 364:340–350
- O'Doherty C, Favorov A, Heggarty S, Graham C, Favorova O, Ochs M, Hawkins S, Hutchinson M, O'Rourke K, Vandembroek K (2009) Genetic polymorphisms, their allele combinations and IFN-beta treatment response in Irish multiple sclerosis patients. *Pharmacogenomics* 10:1177–1186
- Pfeffer MA, Swedberg K, Granger CB, Held P, McMurray JJ, Michelson EL, Olofsson B, Ostergren J, Yusuf S, Pocock S (2003) Effects of candesartan on mortality and morbidity in patients with chronic heart failure: the CHARM-overall programme. *Lancet* 362:759–766
- Poste G (2011) Bring on the biomarkers. *Nature* 469:156–157
- Psaty BM, Hofman A (2010) Genome-wide association studies and large-scale collaborations in epidemiology. *Eur J Epidemiol* 25:525–529
- Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghorri MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, König IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ (2008) Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 117:1675–1684
- Spear BB, Heath-Chiozzi M, Huff J (2001) Clinical application of pharmacogenetics. *Trends Mol Med* 7:201–204
- The PPP Project Investigators (1995) Design, rationale, and baseline characteristics of the Prospective Pravastatin Pooling (PPP) project—a combined analysis of three large-scale randomized trials: Long-term Intervention with Pravastatin in Ischemic Disease (LIPID), Cholesterol and Recurrent Events (CARE), and West of Scotland Coronary Prevention Study (WOSCOPS). *Am J Cardiol* 76:899–905
- Varmus H (2010) Ten years on—the human genome and medicine. *N Engl J Med* 362:2028–2029