

Current challenges in handling genetic data

Felix Gutzwiller, Patricia R. Blank

Institute for Social and Preventive Medicine, University of Zurich, Switzerland

Summary

In no other field of biomedicine has such revolutionary change taken place in recent decades as it has in molecular genetics. The accumulated knowledge in this field will not only enable clinicians to make new treatment decisions in future, but will also help to save on healthcare costs. A positive test result will be the prerequisite for carrying out targeted drug treatment (companion diagnostics). Specific molecular diagnostics provide doctors with additional information that was not previously available, enabling them to optimise treatment accordingly. At the same time, prognostic tests mean that targeted preventive measures can be taken. Highly informative non-invasive tests will enable early detection and prevention to play a greater role. Technological breakthroughs, such as high-throughput sequencing, will lead to a flood of data in the future. The challenge lies in the quality of interpretation, which means extracting useful information for doctor and patient.

Unlike data collection, interpretation is complex and expensive: it requires a high degree of expertise and a lot of resources. At the same time, experts stress that – as well as improvements in the accuracy and speed of data analysis – defined quality criteria must be generated for reliable interpretation of results. These challenges need to be tackled so that the population can benefit to the utmost from the opportunities offered by these developments: rapidly available and informative tests for targeted therapies based on high-quality data.

Key words: *molecular genetics; data; biomarker; test*

Introduction

In recent years the opportunities for collecting individual genetic data have expanded dramatically. Technological advances in high-throughput sequencing of genetic material (including next generation sequencing [NGS]), improvements in microarrays and the steadily falling costs of performing such tests are among the factors that have contributed to this expansion. Today this opens up a whole range of genetic testing possibilities and provides a far more detailed view of the human genome through to its complete sequencing. The “omics era” (genomics, proteomics, transcriptomics, etc.) has definitely arrived.

These technological advances give rise to new opportunities, but also pose certain challenges. Examples of these opportunities include better and faster diagnoses, informative tests that are non-invasive (such as novel prenatal tests on maternal blood) or a reduction of side effects caused by drug treatments (companion diagnostics). Major challenges lie in the interpretation of data, a partial lack of counselling and education of patients, in quality assurance or data security [1]. These factors are particularly evident in so-called direct-to-consumer (DTC) tests, which mean genetic tests that are accessible to all on the internet (in Switzerland DTC tests with medical implications are at present banned).

Challenges posed by direct-to-consumer tests

The challenge posed by DTC tests lies in the fact that they can be lifestyle products, for instance when they claim they can find someone’s genetically “right” partner, calculate a person’s IQ or when they involve ancestry research. However, they can also deliver relevant medical data, for example information about monogenetic diseases such as cystic fibrosis for which there is complete penetrance. Hence a broad range of tests of variable significance are available.

It is essential to respect the individual’s right to autonomy and knowledge, while guaranteeing protection of privacy where necessary. This is because the products can theoretically be used for disguised paternity or for prenatal tests. In these cases the highest priority is the protection of people who lack legal capacity, such as children. Quality control and data security also pose a certain problem: as the analysis of DTC tests is done abroad, it is scarcely possible to check the quality of the results and the security of the data. Certification according to European or internationally valid standards offers a certain degree of security. However, a general prohibition of DTC tests in Switzerland is also rendered unnecessary – simply because it cannot be enforced internationally. In view of these considerations and after thorough consultation, the Expert Commission on Human Genetic Testing (ECHGT) commissioned by the Federal Office of Public Health (FOPH) opted for a cautious opening up to DTC tests (see ECHGT recommendations on revision of the Federal Law on Human Genetic

Testing, March 2013). The Commission does, however, recommend banning such tests if they involve children, for prenatal testing and without a person's consent. It also recommends making the person ordering the test liable.

Challenge of a mountain of data

Another major challenge is the huge amount of data that is essentially generated by these technological advances, especially by NGS. Thanks to NGS, large areas of the genome or exome can now be analysed simultaneously, then interpreted. The advantage is that tests do not have to be conducted consecutively, which is costly and time-consuming. On the other hand, there is a greater probability of incidental findings; hence information that was not even being sought and the patient probably would not wish to know at all. Should the doctor tell the patient about this information? Only in the case of very serious diagnoses? Or not even then? The person having a test done must be aware beforehand what information might be delivered by such a test and know what he/she does and does not want to be told [2].

Large quantities of data in themselves do not signify a real gain in knowledge. It is only the quality of interpretation that creates added medical value. Achieving this requires a high level of expertise and – in contrast to data collection – it remains complex and expensive. Experts agree that the interpretation of genetic data cannot yet keep pace with data collection. Meanwhile the ever-increasing flood of data does not help to simplify data interpretation. At the same time, experts stress that – as well as improvements in the accuracy and speed of data analysis – defined quality criteria must be worked out for reliable interpretation of results. High-throughput sequencing of lots of genes is expected to detect thousands of new variants in the patients being tested. To minimise uncertainties and fears about unclear clinical relevance, approaches need to be validated which will enable benign polymorphisms to be distinguished from potentially pathogenic sequence changes. Differences currently exist between provider laboratories in this respect. Everyone agrees that, as well as solid experience, databases containing very high numbers of cases are required to guarantee reliable results.

A good example of the challenge posed by data interpretation is evident from the requirement to make databases of biomedical companies public. Many of the companies in Switzerland and abroad have extensive databases built up over a number of years. For instance, a company working in this field over the past twenty years has developed one of the largest databases for the analysis of the BRCA1 and BRCA2 genes, recording all the sequence changes to the relevant genes that have been diagnosed in patients. This allows reliable statements to be made about the clinical relevance of most of the variants in these genes [3]. The requirement to make these databases public or at least share them only makes sense superficially. This is because the data merely form the foundation for what is actually the decisive element in genetic testing, namely the reliable interpretation of the data based on very large numbers of cases. Data security is also a key issue. Companies' genetic databases contain sensitive data because they hold highly per-

sonal patient details, such as the data found in cancer registers (cancer diagnoses). If data were to be made public, any residual risk of inferring the identity of the person concerned would have to be entirely ruled out [4, 5].

Good data quality is expensive

Furthermore it is a matter of investment: building up a database is a costly business. Constant maintenance is required to guarantee superior data quality in the long term. On a daily basis, new data sets are generated and identified changes may have to be re-assessed. Doctors and those affected may also be informed of any reclassification via supplementary reports even years after the testing took place. This contrasts with public databases (e.g. BIC, LOVID, ENIGMA) which are not always updated regularly. Users of such databases run the risk of relying on obsolete data. Over the years, the cost of maintenance and updating a large database can add up to substantial amounts of money. Companies make these investments because good data are the basis for high quality and hence for business growth [6].

These challenges need to be tackled so that the population can benefit to the utmost from the opportunities offered by these developments: rapidly available and informative tests for targeted therapies based on high quality data.

Funding / potential conflicts of interest: This article is a summary of the expert forum on the subject "Personalised medicine – hope for the future" which was held in Zurich on 14 October 2013. This workshop was financially supported by Myriad Genetics GmbH (Schweiz). The authors declare that there are no conflicts of interest in connection with this article.

Members of the expert forum: Dr. J. Achermann, I. Bachmann-Mettler, N. Braun, Dr. med. B. Bolliger, N. Chantziaras, Prof. Dr. S. Gallati, Dr. J. Th. Gasser, F. Kalaitzidis, Dr. M. Gattiker, Prof. Dr. med. D. Hochstrasser, Prof. Dr. med. A. Huber, M. Kessler, G. King, Dr. R. Marti, T. Maurer, Dr. med. R. Morant, Dr. C. Noppen, Prof. Dr. med. B. Pestalozzi, Prof. Dr. med. F. Recker, Prof. Dr. S. Rohmann, A. Schaller, Dr. med. L. Schmid, H. Stocker, Dr. E. Vayena, S. Wehnelt, Dr. med. J. Zollkofer

Correspondence: Patricia R. Blank, PhD MPH (cand.), Universität Zürich, Institut für Sozial- und Präventivmedizin, Hirschengraben 84, CH-8001 Zürich, Switzerland, [patricia.blank\[at\]fspm.uzh.ch](mailto:patricia.blank[at]fspm.uzh.ch)

References

- 1 Rehm, et al. ACMG clinical laboratory standards for next generation sequencing. *Genet Med.* 2013;15(9):733–47.
- 2 Green, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med.* 2013;15(7):565–74.
- 3 Eggington, et al. A comprehensive laboratory-based program for classification of variants of uncertain significance in hereditary cancer genes. *Clin Genet.* 2013;8.
- 4 Greenbaum, et al. Genomics and privacy: implications of the new reality of closed data for the field. *PLoS Comput Biol.* 2011;7(12).
- 5 Gymrek, et al. Identifying Personal Genomes by Surname Inference. *Science.* 2013;339(6117):321–4.
- 6 New NIH-funded resource focuses on use of genomic variants in medical care, Press release 2013, <http://www.nih.gov/news/health/sep2013/nhgr-25.htm>