Virtual Clinical Trials, an Essential Step in Increasing the Effectiveness of the Drug Development Process

Hans Lehrach

Max Planck Institute for Molecular Genetics (MPIMG), Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS), and Alacris Theranostics GmbH, Berlin, Germany

in both patient treatment and prevention as well as drug development in the future if we do not want to continue to make dangerous and expensive mistakes in real life.

© 2015 S. Karger AG, Basel

Abstract

Every patient is different – his/her genomes, environment, disease history and exposure to drugs. Tumours, in particular, are often heterogeneous in their genetic make-up and their response to drugs, both within and between samples. Classic clinical trials basically ignore this complexity or, as in stratified medicine, attempt to reduce it to an analysis of a small number of still enormously heterogeneous patient groups. Medicine, however, is not the only area in which we are faced with such complex 'n = 1' (every individual case is different) situations. The weather we experience today, characterised by tens of terabytes of measurement data, has never occurred before and will never occur again. Similar to the situation in medicine, we cannot predict the development of today's weather by looking for identical weather conditions in the past, and we cannot, in real life, test drugs for every individual patient in clinical trials with large numbers of biologically identical patient replicas. We can, however, do it with the help of models duplicating the 'n = 1' situation on the computer, an approach which will also have to be used

Introduction

Despite significant progress in the quality of services offered by health care systems over the last decades, leading to drastic improvements in health and well-being of the general public, our health care systems still fail millions of people each year. Throughout Europe, 4,000 people die of cancer every day [1]. A stark statistic that represents the untold misery and suffering of patients and families, as well as a growing financial pressure: in 2009 alone, cancer caused costs of EUR 126 billion in Europe [2]. Given the rapid ageing of populations – by 2025 more than 20% of Europeans will be 65 years or older [3] – our health care systems are set to be challenged even further. The cost of meeting such challenges is spiralling. Every day close to EUR 4 billion are spent by European Union (EU) member states on health care [4], an amount likely to rise with the emerging demographic transition, raising doubts concerning the sustainability of our current approaches to health care.

Prof. Hans Lehrach
Max Planck Institute for Molecular Genetics (MPIMG)
Ihnestrasse 63–73
DE–14195 Berlin (Germany)
E-Mail lehrach@molgen.mpg.de

In general, health care is based upon a 'one-size-fits-all' approach. When it comes to administering drugs, patients are given treatments that have been found statistically to be the best option for a similar group of patients. This approach does not always mean the majority of patients will recover. A fraction will respond positively, while others may actually become sicker or might even die due to unexpected side effects of a chosen therapy. Available estimates suggest between 38 and 75% of patients are irresponsive to drugs selected in this manner [5, 6], which not only results in unnecessary suffering for both patients and families but also represents an enormous economic burden on health care systems, amounting to EUR 100 billion or more per year within Europe alone. Drugs (often expensive) that are not effective for the patient receiving them not only delay recovery but can potentially trigger detrimental effects – requiring additional treatments, possibly long-term care – and potentially cause a reduction in productive life span.

Underlying these differential responses to drug treatments is the simple fact that we are all very different from one another. Our genetic legacy provides us with different (forms of) genes, we all have different disease histories (and futures), and we lead different lifestyles and are exposed to different environments. It should not be too much of a surprise that we might respond in a different way to the drugs we take. This is exemplified in cancer, where the scope for heterogeneity is manifold. Tumours are not only different from each other because our genomes are different; they are formed on the basis of random processes in their own genome and epigenome, which make each tumour absolutely unique. This vast molecular diversity of tumours means that each one has never existed before or will ever exist again. However, it is not just individual tumours which are dissimilar, even subpopulations of cells from the same tumour can react differently to the treatment a patient receives. In particular, we do not know the full extent to which various components interact, such as the genetic make-up of individuals, environmental and lifestyle factors and microbiomes (i.e. the microbial residents in the gut and other bodily surfaces). Current paradigms (and knowledge) are still not fully embracing the irreducible complexity of the 'biological system' that each human being is.

What Does This Mean for Clinical Trials?

Current clinical trial designs essentially ignore this complexity, with the heterogeneity of the patients in a trial translating into a similar heterogeneity of responses to the drug after approval, forcing doctors to gamble with their patients' health. Due to the differences in their biology, some will improve, some might get worse, and some might even die due to a drug badly fitted to their biological make-up. Patient stratification by biomarkers has, in certain cases, aided in improving the situation, helping to subdivide patients into groups that are more or less likely to respond to a therapy. To date, however, clinically useful biomarkers have been difficult to identify. Despite intense research, fewer than 100 out of 150,000 publications in this field have identified biomarkers that have been subsequently endorsed for routine clinical practice [7]. In addition, predictions through biomarkers are often still statistical (not all of the predicted responders respond, and not all of the predicted non-responders do not), leaving doctors and patients still in a lottery, albeit one biased in their favour. Within the context of clinical trials, an essentially unsolvable problem is also posed by trying to predict the likely response of a patient from a specific combination of biomarker results. It quickly becomes technically impossible to test these results robustly in clinical trials.

I am convinced we now can do (a lot) better, based on the basic approach used in many other situations where we have to handle difficult, unique problems with dangerous and/or expensive consequences: the use of 'mechanistic' computer models, which essentially duplicate part of the real world on the computer in sufficient detail to make them respond to challenges in a way similar to their real-life counterparts. Mechanistic models can be used to investigate the dynamic behaviour of complex systems, based on fundamental knowledge of the interactions occurring between the components of a biological system, to generate in silico models with large-scale predictive capabilities [e.g. see 8–11].

We use such models to predict the weather and to find out whether building designs will translate into structures that can withstand storms or even earthquakes, or if a new car design has any fatal flaws before it is put into production. Pilots are trained using computer models which allow them to make mistakes and crash 'virtual planes' until they are proficient enough to fly the real thing. This general 'risk avoidance' strategy of using computer models for defining the optimal response in complex situations allows us to make unavoidable mistakes on the computer rather than in reality, ultimately improving designs, accelerating developments, reducing risks and saving lives. Computer models provide the only way of using the large amounts of information necessary to define a complex situation in sufficient detail to pre-

dict how one system will evolve in contrast to a superficially similar system.

Application of such a strategy to the extremely complex, highly interactive biological networks acting in us to keep us alive and healthy or make us sick may also reap comparable benefits in the context of our health and well-being. So why do we not use a similar strategy to make unavoidable mistakes first on a computer model rather than in a real patient or in a real clinical trial? The answer is quite simple. Although for quite some time we have gained increasing amounts of relevant information on the basic biological networks in human cells (probably not all of it correct) as well as rapidly increasing computer power [12], in principle enabling us to handle models of increasing complexity, we have not had one key component that is available in all other situations handled successfully by such computer models, the equivalent of the tens of terabytes of information on every patient, which allows weather forecast models to accurately define the situation at the start of the modelling run (e.g. 'a forecast model is only as good as the data put into the model' [13]). The decisive difference between human patients (or clinical trials) and weather forecasts, virtual crash tests and many other areas where we use mechanistic models to be able to predict what will work and what will end in disaster has therefore not primarily been the difference between physics and biology, the types of differential equation used or the overall complexity of the model, but simply the detailed knowledge available about the situation at the start of the modelling run.

Fortunately the situation is changing rapidly, driven to a large extent by huge technological leaps forward in next-generation sequencing and other analytical technologies. The first human genome took more than 10 years and billions of dollars to sequence; today multiple genomes can be sequenced in days by one machine at a cost of approximately USD 1,000 per genome. In a relatively short period of time, costs have decreased by roughly 6 orders of magnitude, a trend that is set to continue. Third-generation instruments are now under development, which have the potential to generate longer reads in a shorter time at lower costs. Sequencing is, however, not just used to characterise the genetic component of the processes in the body. It can also be used to characterise the epigenetic processes controlling whether genes are read out. Sequencing of the transcriptome gives us information on which genes are read out and when, as well as detection of patterns of allelic expression and alternative splicing. DNA sequencing can also be used to characterise the immune system of individual patients [14], opening the way

to make predictive models of the state of the immune system of every individual patient, which is invaluable for diseases with immune components (e.g. type 1 diabetes, rheumatoid arthritis, etc.) or the optimisation of immunotherapies for cancer patients. Similarly, through a number of technologies, we have made huge progress in our capability to characterise proteomes and metabolomes, to characterise exomes, epigenomes and transcriptomes of single cells, and to even carry out some of these analyses in a spatially resolved fashion [15–18]; this opens the route towards modelling the tumour of every patient, including the complex interactions occurring between genetically or epigenetically different subgroups of cells of the same tumour with each other, with the stroma and with cells of the immune system localised within the tumour.

In a sense, we can, at reasonable cost, learn more about every single patient than we knew about the whole of human biology just a few years ago. We can use this information to model these patients accurately on the computer, and treat these models with any drug or drug combination to find the optimal treatment. The optimal treatment is 'experimentally' determined in a safe, quick and cheap manner in a computer model instead of in the real patient – as current practice dictates, with potentially lethal consequences. A route towards provision of patients with drugs they actually respond to would not only substantially reduce current expenditure on drugs but also avoid many downstream costs, such as extended care, the need for additional therapy to counteract the side effects of the original treatment as well as illness-related absences from work – the latter currently costing European nations approximately 2.5% of their gross GDP per year [19].

Although such models are still far from perfect, they are in many cases (particularly in oncology) likely to perform better than current clinical practice. As information on biological networks improves, and disease mechanisms and parameters become increasingly well defined, based on a systematic comparison of predicted and actual therapy responses, overall accuracy will improve asymptotically. This continuous 'reverse engineering' of biological mechanisms will, in addition, provide valuable input for hypothesis-driven basic research, ultimately providing as much (or more) information on disease mechanisms in humans as other sources of information.

The same technology also promises to solve the basic dilemma we are facing in clinical trials. Real clinical trials can be highly deleterious to the health of participating patients through drug side effects, especially for those patients who have no clinical benefits. They are extremely expensive and slow, and in addition (at least with new

Lehrach

drugs) mostly fail, for a large number of reasons. Only 1 in 10–20 drugs tested in clinical trials will actually gain approval, with each new drug reaching the market having spent 10 to 12 years in development and cost over USD 2.5 billion [20]; even if the drug is deemed to be successful, this often results in the approval of drugs that are still ineffective for most patients, resulting, on average, in relatively small clinical benefits at very high costs. Within a virtual clinical trial scenario, multiple individual models are used to make mistakes on the computer, quickly, cheaply, safely and at a low risk to the developer, prior to the initiation of a real-life clinical trial. In itself, this will reduce the number of failed clinical trials and, by allowing the focus to be placed on the best drug candidates, increase the number of drugs reaching the market.

By modelling patients in virtual clinical trials using publicly available data, e.g. the colon cancer patients characterised in the International Cancer Genome Consortium (ICGC) [21] and similar programmes, there is also a chance to identify candidates for biomarkers that can, for instance, identify a significant fraction of patients likely to respond to a drug. The identification of such biomarkers potentially leads to rapid, cost-effective and low-risk real clinical trials of a well-defined responder population (maybe for multiple tumours in parallel, maintaining the overall market size, but vastly increasing the response rate).

The inclusion of such virtual trials in the drug development process could have a number of additional advantages over the exclusive reliance on real clinical trials. Virtual clinical trials do not endanger patients. In contrast to real trials, which can only be performed after an extensive, lengthy and already quite costly development process, virtual clinical trials (and even virtual health technology assessment exercises) could be performed at very early developmental stages (i.e. as soon as the likely binding strength to the different molecular targets can be defined, possibly even before synthesis of the compound, if docking programmes give sufficiently accurate predictions), thereby accelerating development and reducing the very large fraction of drugs failing at some stage during preclinical or clinical development.

In contrast to real clinical trials, virtual clinical trials would incur limited costs per patient. Also, compared with real clinical trials, which can rapidly become unaffordable for even the largest pharmaceutical companies, virtual clinical trials can be routinely carried out on all previously characterised relevant patients, allowing clinical trials to encompass tens of thousands, and ultimately even millions, of subjects (e.g. cancer patients) characterised in detail during their therapy.

Sample Applications of Virtual Clinical Trials

Virtual clinical trials could have myriad applications within personalised medicine as well as the drug development and discovery pipeline, including:
1 early drug development, supporting the pharmaceutical industry and biotech companies in early developmental stages (responder groups can be predicted from public or industry-supplied data and efficiently guide experimental validation in animal or cell line experiments);
2 support of the pharmaceutical industry in clinical trials of newly developed drugs (responder and non-responder groups are identified, enabling a highly cost-efficient, several-fold increase in approval of new drugs);
3 drug repositioning, with approval for different disease applications;
4 drug rescue ('fallen angels'), typically for drugs that have failed in phase II or III clinical trials due to low response rates (typically below 20%) rather than due to non-tolerable side effects (the identification of responder groups will provide the means of gaining approval for use of these drugs commercially for the benefit of defined patient groups);
5 selection of the most efficient drug/drug combination for patients;
6 prediction of additive and synergistic effects of drug combinations.

Virtual trials have the potential to dramatically reduce the requirement for animal testing in preclinical drug developmental stages and ensure only patients who are most likely to respond positively to a drug will be enrolled in real-life clinical trials. Aligning with EU initiatives that are calling for improvements in the competitiveness and quality of public services and people's lives [6], such a 'prescreening' stage using virtual clinical trial technology is likely to become a future prerequisite for any clinical trial, helping to protect patient welfare and increase cost-efficiency.

The Next Stage: The Patient-Specific Clinical Trial?

Although a virtualisation of clinical trials could already generate major benefits within the current stratification- and biomarker-based regulatory framework, we would consider this just a step towards a truly personalised medicine and drug approval framework. Once patients are routinely characterised and modelled to predict

the optimal treatment (or prevention) for any individual, the ultimate test of whether a drug should or should not be used to treat the individual patient would, in a sense, be dependent on his/her personal characteristics rather than, by necessity, the outcome of heterogeneous real clinical trials. It might therefore be worth considering modelling results as the equivalent of a current phase III trial carried out on each individual patient [22]. Such trials could still explore multiple possibilities, e.g. hundreds of subtly different models of the individual patient reflecting remaining uncertainties in the biology of the patient, the mechanisms of action of the drug or the biological networks underlying the modelling. Critical differences in outcome of the individual variants could then be resolved by re-analysis of, for instance, the raw reads of the primary analysis of the patient, or, if sufficiently critical, by additional diagnostic tests initially not carried out due to high costs or invasiveness.

Such models would be self-learning, with information on differences between prediction and actual treatment results from an increasing number of patients as well as many experimental systems (e.g. from preclinical analyses) flowing back to progressively define the parameter space and improve the structure of the models, increasingly ensuring both optimal treatment and optimal predictions in virtual clinical trials.

The Legislative Landscape

Changing paradigms also requires the adaptation of existing and provision of new legislation. There are currently many legal and regulatory hurdles to be faced that hamper virtual trial scenarios, especially on a large scale. The information flow required for such virtual trial scenarios is currently, to some extent, hindered by prevailing data protection rules. To allow large-scale virtual clinical trials, regulations such as the General Data Protection Regulation [23] need to be defined to permit access to data for the benefit of current and future patients while ensuring a robust governance framework which can provide appropriate technical and ethical safeguards to protect people's personal data.

Conclusions

At the congruence of technological and computational developments, and the generation of 'big data', mechanistic models with predictive capacity can be deployed for the simulation of clinical trials with associated benefits for patient welfare and the economy. A testament to the effectiveness of such an approach can be gleamed from the many other areas that have already switched to data- and model-driven development cycles (e.g. the automobile and aviation industries). In many cases, an increased reliance on modelling techniques would improve quality, decrease costs, accelerate development, reduce risks and ultimately make many more (and better) drugs available for patients. An increased virtualisation of the drug development process – with virtual clinical trials as one of the key components as well as more personalised therapy and prevention strategies based on patient modelling – might, in our ageing societies, very well be the only alternative to increased rationing of health care provision [24].

Acknowledgements

Disclosure Statement

Hans Lehrach is founder of and scientific advisor to Alacris Theranostics GmbH, a company which aims to develop 'virtual patient' models for use in therapy choice and drug development.

References

1 Stewart BW, Wild CP: World Cancer Report 2014. Geneva, World Health Organization, 2004.
2 Luengo-Fernandez R, Leal J, Gray A, Sullivan R: Economic burden of cancer across the European Union: a population-based cost analysis. Lancet Oncol 2013;14:1165–1174.
3 European Commission: Population groups: elderly. http://ec.europa.eu/health/population_groups/elderly/index_en.htm.
4 Busse R, Blümel M, Scheller-Kreinsen D, Zentner A: Tackling chronic diseases in Europe: strategies, interventions and challenges. Observatory Studies Series No 20. Geneva, World Health Organization, 2010.

5 Spear ABB, Heath-Chiozzi M, Huff J: Clinical application of pharmacogenetics. Trends Mol Med 2001;7:201–204.

6 Einsidel EF: Emerging Technologies: From Hindsight to Foresight. Toronto, UBC, 2009, p 360.

7 Poste G: Bring on the biomarkers. Nature 2011;469:156–157.

8 Wierling C, Herwig R, Lehrach H: Resources, standards and tools for systems biology. Brief Funct Genomic Proteomic 2007;6:240–251.

9 Wierling C, Kühn A, Hache H, Daskalaki A, Maschke-Dutz E, Peycheva S, Li J, Herwig R, Lehrach H: Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment. Mutat Res 2012;746:163–170.

10 Röhr C, Kerick M, Fischer A, Kühn A, Kashofer K, Timmermann B, Daskalaki A, Meinel T, Drichel D, Börno ST, Nowka A, Krobitsch S, McHardy AC, Kratsch C, Becker T, Wunderlich A, Barmeyer C, Viertler C, Zatloukal K, Wierling C, Lehrach H, Schweiger MR: High-throughput miRNA and mRNA sequencing of paired colorectal normal, tumor and metastasis tissues and bioinformatic modeling of miRNA-1 therapeutic applications. PLoS One 2013;8:e67461.

11 Henderson D, Ogilvie LA, Hoyle N, Keilholz U, Lange B, Lehrach H: Personalized medicine approaches for colon cancer driven by genomics and systems biology: OncoTrack. Biotechnol J 2014;9:1104–1114.

12 Prometeus GmbH: TOP500. http://www.top500.org/statistics/perfdevel/.

13 Haby J: Forecast model data input. http://www.theweatherprediction.com/habyhints/60/.

14 Glökler J, Warnatz HJ: CDR sequence-specific enrichment of live immune cells. European Patent Application 2012 No 12165009.7.

15 Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, Terry R, Jeanty SS, Li C, Amamoto R, Peters DT, Turczyk BM, Marblestone AH, Inverso SA, Bernard A, Mali P, Rios X, Aach J, Church GM: Highly multiplexed subcellular RNA sequencing in situ. Science 2014;343:1360–1363.

16 Lee J, Daugharthy ER, Scheiman J, Kalhor R, Ferrente TC, Terry R, Turczyk B, Yang JL, Lee HS, Aach J, Zhang K, Church GM: Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat Protoc 2015;10:442–458.

17 Gu GJ, Friedman M, Jost C, Johnsson K, Kamali-Moghaddam M, Plückthun A, Landegren U, Söderberg O: Protein tag-mediated conjugation of oligonucleotides to recombinant affinity binders for proximity ligation. N Biotechnol 2013;30:144–152.

18 Weibrecht I, Gavrilovic M, Lindbom L, Landegren U, Wählby C, Söderberg O: Visualising individual sequence-specific protein-DNA interactions in situ. N Biotechnol 2012; 29:589–598.

19 Eurofound: Absence from work. 2010. http://www.eurofound.europa.eu/observatories/eurwork/comparative-information/absence-from-work.

20 Tufts Center for the Study of Drug Development: Cost to develop and win marketing approval for a new drug is $2.6 billion. 2014. http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study.

21 International Cancer Genome Consortium: International network of cancer genome projects. Nature 2010;464:993–998.

22 Hood L, Balling R, Auffray C: Revolutionizing medicine in the 21st century through systems approaches. Biotechnol J 2012;7:992–1001.

23 European Commission: Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

24 OECD: Fiscal Sustainability of Health Care Systems: Bridging Health and Finance Perspectives. Paris, OECD, 2015.