Personalized drug repositioning using gene expression

Koudijs, K.K.M.

CHAPTER 1

General introduction

Need for new and affordable cancer treatments

Cancer is the leading cause of death in 57 countries, including the U.S., China and most European countries and is expected to surpass cardio-vascular disease as the leading cause of death worldwide over the course of this century.[1] Conventional cancer treatment, by chemotherapy, typically consists of combinations of highly toxic antiproliferative compounds with many side-effects. Fortunately, over the past decades many new more specifically acting effective anti-cancer drugs have been introduced with more favorable side-effect profiles, the so-called targeted therapies. These novel drug therapies have greatly improved survival for some previously poor prognosis cancers such as metastatic breast cancer and lung cancer. Nevertheless, currently only 8% of patients qualify for targeted anti-cancer therapies and an even smaller percentage actually benefit from it.[2] An additional problem complicating both conventional chemotherapy and targeted therapy is the development of drug resistance: any cancer cells that are insensitive to the drug treatment could reproduce and form a resistant clone by evolutionary selection.[3] Unfortunately, development of a completely new drug (i.e. a new chemical entity) takes many years of development: for anti-cancer drugs in the period of 2009–2016, the average time from the earliest filing of the patent paperwork to availability to NHS patients was 14 years.[4] However, perhaps the biggest problem is the high cost of new drug development: a 2016 analysis based on the analysis of 106 new drugs reported an average cost of $1.4 billion per drug in 2013 dollars.[5] If the cost of capital and post approval R&D are included, this number more than doubles to $2.9 billion per drug. As cancer consists of more than 200 different neoplastic diseases, most of which can likely be categorized into even smaller subsegments with different initial and delayed responses to drug treatment, it may not be economically feasible to develop new drugs for all possible cancer types and their subsets. Any method which can speed up the availability and/or reduce the cost of new drug therapies would therefore be highly welcomed.

Drug repositioning

A promising alternative method to novel drug development is called drug repositioning. Drug repositioning (also frequently referred to as drug repurposing) is defined as the use of a drug in a new indication, i.e. other than what it was originally marketed for.[6] Because much is already known about the drug's safety and toxicity profile in humans, a repositioned drug can be developed at a reduced cost and time to patient.[7] However, the main disadvantage of drug repositioning lies in the relatively weak intellectual property protection, discouraging financial investment in drug repositioning, especially by commercial parties.[8] The responsibility of investigating drug repositioning of generic and off-patent drugs therefore rests mainly on academic institutions and non-profit initiatives, and has become a

hot-topic of academic research: the number of academic papers available in the biomedical literature search engine PubMed with either "drug repositioning" or "drug repurposing" in the title or abstract has increased over thirty-fold over the last 10 years, from 54 papers in 2011 to 1,685 papers in 2021. The same query combined with the words 'cancer', 'tumor' or 'tumour' shows an even stronger fifty-fold increase, from 8 papers in 2011 to 400 papers in 2021.

Researchers have explored a variety of methods and concepts to discover which drugs can be repositioned for use in cancer treatment.

Drug repositioning using laboratory experiments

Testing a drug in cell cultures (*in vitro*) or in laboratory animals (*in vivo*) gives some information about its efficacy in human cancers. For example, gemcitabine was originally developed as an anti-viral drug but the intended indication changed after pre-clinical testing showed that it killed leukemia cells *in vitro*.[9] The sedative and anti-emetic thalidomide was taken off the market for causing congenital defects, but reappeared on the market with the new indication multiple myeloma after it was shown in a rabbit cornea micropocket assay that thalidomide inhibits the formation of blood vessels *in vivo*.[10] In most cases it is not even necessary anymore to perform these experiments in a wet-lab experiment to discover the anti-cancer properties of existing drugs: many datasets have become available with the results from high-throughput drug screens, in which many cell lines are systematically screened for cell death after administration of multiple different drugs, such as the Genomics of Drug Sensitivity in Cancer (GDSC), Cancer Therapeutics Response Portal (CTRP) and Cancer Cell Line Encyclopedia (CCLE).[11-13] However, comparison of the results from these different *in vitro* high-throughput screens have shown that the results from different datasets tend to weakly correlate, raising serious questions about the predictive validity for the results *in vivo*.[14] Experiments to validate the results of *in vitro* screens to the results *in vivo* and subsequent standardization of *in vitro* methods should be performed to increase the reliability of the *in vitro* results. To start animal testing or a clinical trial based only on the results from a single *in vitro* screen, high-throughput or not, absent a prior hypothesis about the mechanism of action of a particular drug against a specific cancer, would likely result in poor results. For this reason, combining the *in vitro* results with those of another method described below to increase the reliability of the prediction is recommended.

Drug repositioning using clinical observation or retrospective observational studies

The observation that bone marrow and lymph nodes of people got severely depleted after exposure to mustard gas in WW1 and after an accidental spill of sulfur mustards in WWII led to the investigation of these compounds and its derivatives as therapeutics against lymphoma.[15] However, such serendipitous clinical observation can only lead to further development if the effect is consistent, large and occurs quickly after exposure. In addition, it depends on careful observation of clinical researchers. A more sensitive approach is to perform retrospective observational studies, i.e., of cancer patients taking drugs not prescribed as cancer treatment. Systematic retrospective observational studies are able to find more subtle effects and with higher reliability and have fueled enthusiasm to start clinical trials using commonly prescribed drugs such as statins, metformin and aspirin.[16-18] Initiatives such as the PHARMO Database Network in the Netherlands enables follow-up, including prescribed medication, of more than 10 million residents of the Netherlands for on average 12 years, greatly increasing the potential scope of new observational studies.[19] However, the downsides of this approach are that it only tends to generate leads for commonly prescribed drugs (due to sample size requirements needed to detect a statistically significant effect), there is typically no biomarker available to predict which patients may benefit most from the drug and lastly, the inherent potential biases from observational studies. For example, patients which live longer or which have a better prognosis are more likely to receive medications such as statins, and may therefore only seem to experience a survival benefit if this is not corrected for in a time dependent analysis.[20] In addition, many other types of potential bias can exist in observation studies which can compromise its validity.[21] Nevertheless, while most of the time observational studies are not sufficient on its own, retrospective human evidence certainly is a valuable tool.[22] It should be noted however that observational studies are primarily hypothesis generating and not conclusive on its own.

Drug repositioning using prospective studies

A popular drug repositioning approach in oncology is to use the newly developed targeted anti-cancer agents for new cancer types outside of their original limited indication. This method is pioneered by initiatives such as in the The Drug Rediscovery Protocol (DRUP) trial: an ongoing, prospective, multi-drug and pan-cancer trial.[23] In these trials, specific patient cohorts are defined based on existing clinical and biological knowledge (e.g., patients with tumors that have an EGFR mutation regardless of tissue of origin) and are given the same targeted drug (in this case e.g., the EGFR inhibitor erlotinib, currently only registered to treat non-small cell lung cancer and pancreatic carcinoma). By systematically including

eligible patients and recording the outcomes, it eventually becomes possible to observe trends in which 'drug-tumor type-molecular profile' (agnostic) combinations seem to work, providing the rationale for a larger and controlled randomized clinical trial. Preliminary results from the DRUP trial indicate that of the over 600 cases which were submitted for central review, less than half (N = 294) started treatment and of the 215 patients with sufficient follow-up, clinical benefit was observed in 74 patients (34%). While certainly impressive and hopeful, extrapolating these numbers shows that at most 1/6 of patients can be served with the current repertoire of targeted anti-cancer agents.

Drug repositioning using computational methods

Aside from computational methods which rely on data from *in vitro* cell killing experiments or patient records, there are computational methods which rely on completely different data and can thus be used to complement and corroborate the results from these other approaches, increasing the plausibility that a drug can be repositioned against a particular tumor type. For example, it is possible to virtually screen based on the molecular structures of drugs whether any existing drugs can interact with a specific protein of interest, e.g. one which is overexpressed or mutated in a specific cancer (sub)type of interest.[24] The downside of this method is that it relies on existing knowledge: e.g., the target protein and its structure should be determined or inferred (e.g., using DeepMind's AlphaFold), and additional experiments are required to confirm if the affinity of the drug is high enough to affect the protein at a concentration achievable inside the human body and it achieves the intended effect on the targeted cell type(s).

Another computational drug repositioning method relies on data from gene expression perturbation experiments. In these experiments, different cell lines are incubated with various drugs at different concentrations and durations to find out which genes are transcribed more intensively (upregulated) or less intensively (downregulated) compared to control cell lines. This produces a complete readout of all genes affected by the drug at various concentrations and timepoints.[25,26] One way this information has been used is to discover new drug repositioning candidates is through similarities in the mechanism of action: if drug A and drug B share similarities in mechanism of action at the gene expression level, and drug A is already used against a specific cancer (sub)type, then drug B might also be useful against the same cancer (sub)type. Another way this gene expression perturbation data has been used is to discover drugs which may act against a cancer (sub)type using a mechanism of action different from any existing drugs. We call this method Transcriptome Signature Reversion (TSR) and is illustrated in Figure 1.1.

The TSR method was first pioneered in 2011 by Sirota et al. to find drug repositioning candidates against lung adenocarcinoma,[27] but has since been applied to many other tumor types.[28-34]

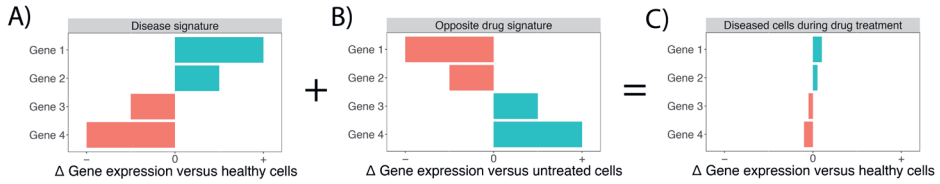

Figure 1.1: Illustration of the Transcriptome Signature Reversion (TSR) method.
A) First the gene expression difference between diseased cells and control cells is compared, to find out which genes are upregulated and downregulated in the disease ('disease transcriptome'). B) This disease signature is then used to find drugs with the opposite gene expression signature, i.e., downregulates genes upregulated in the disease, and upregulates genes downregulated by the disease. C) The hypothesis is that if a drug reverses the disease phenotype at the gene expression level, it might be therapeutically active in treating this disease.

Aim and outline of this thesis

In this thesis, we aim to explore the use of TSR for drug repositioning. In addition, since individual tumor gene expression may vary within a tumor type, we investigated personalized drug repositioning by using individual tumor gene expression signatures rather than expression profiles characteristic of a tumor type.

In **chapter 2**, a review and critical appraisal of the existing evidence on the use of TSR as a method to reposition drugs against cancer is presented. In addition, the challenges of making personalized drug repositioning recommendations are discussed.

In **chapter 3**, the development and application of TSR to find personalized repositioned drugs is presented.

In **chapter 4**, the role of the fraction of tumor cells in a tumor sample in the drug repositioning approach is studied.

In **chapter 5**, the use of TSR as a drug repositioning method against cancer types based on gene expression data from 18 different solid tumor types is validated. In addition, it was investigated whether correcting the expression of genes associated with decreased cell viability after drug exposure improves the predictive power of TSR.

In **chapter 6**, a new method to normalize the gene expression of RNA-seq tumor samples using a reference panel of low-variability genes was developed.

In **chapter 7**, the use of convolutional neural networks (CNNs) on gene expression data is explored. We test different methods of pre-structuring the data (hierarchical clustering and a dense layer) and compare the results to a neural network without a convolutional component.

In **chapter 8**, a general discussion is presented that discusses the opportunities and limitations of TSR for drug repositioning and future perspectives are given.

References

1. Bray F, Laversanne M, Weiderpass E, Soerjomataram I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. Cancer. 2021;127(16):3029-30.
2. Smith CEP, Prasad V. Targeted Cancer Therapies. Am Fam Physician. 2021;103(3):155-63.
3. Vasan N, Baselga J, Hyman DM. A view on drug resistance in cancer. Nature. 2019;575(7782):299-309.
4. Sharpe E, Hoey R, Yap C, Workman P. From patent to patient: analysing access to innovative cancer drugs. Drug Discov Today. 2020;25(9):1561-8.
5. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. J Health Econ. 2016;47:20-33.
6. Jourdan JP, Bureau R, Rochais C, Dallemagne P. Drug repositioning: a brief overview. J Pharm Pharmacol. 2020;72(9):1145-51.
7. Nosengo N. Can you teach old drugs new tricks? Nature. 2016;534(7607):314-6.
8. Rastegar-Mojarad M, Ye Z, Kolesar JM, Hebbring SJ, Lin SM. Opportunities for drug repositioning from phenome-wide association studies. Nat Biotechnol. 2015;33(4):342-5.
9. Hertel LW, Boder GB, Kroin JS, Rinzel SM, Poore GA, Todd GC, et al. Evaluation of the antitumor activity of gemcitabine (2',2'-difluoro-2'-deoxycytidine). Cancer Res. 1990;50(14):4417-22.
10. D'Amato RJ, Loughnan MS, Flynn E, Folkman J. Thalidomide is an inhibitor of angiogenesis. Proc Natl Acad Sci U S A. 1994;91(9):4082-5.
11. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955-61.
12. Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol. 2016;12(2):109-16.
13. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7.
14. Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJ, et al. Inconsistency in large pharmacogenomic studies. Nature. 2013;504(7480):389-93.
15. DeVita VT, Chu E. A history of cancer chemotherapy. Cancer Res. 2008;68(21):8643-53.
16. Graaf MR, Beiderbeck AB, Egberts AC, Richel DJ, Guchelaar HJ. The risk of cancer in users of statins. J Clin Oncol. 2004;22(12):2388-94.
17. Coyle C, Cafferty FH, Vale C, Langley RE. Metformin as an adjuvant treatment for cancer: a systematic review and meta-analysis. Ann Oncol. 2016;27(12):2184-95.
18. Elwood PC, Morgan G, Delon C, Protty M, Galante J, Pickering J, et al. Aspirin and cancer survival: a systematic review and meta-analyses of 118 observational studies of aspirin and 18 cancers. Ecancermedicalscience. 2021;15:1258.
19. Kuiper JG, Bakker M, Penning-van Beest FJA, Herings RMC. Existing Data Sources for Clinical Epidemiology: The PHARMO Database Network. Clin Epidemiol. 2020;12:415-22.

20. Tran AA, Prasad V. Drug repurposing for cancer treatments: a well-intentioned, but misguided strategy. Lancet Oncol. 2020;21(9):1134-6.
21. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002;359(9302):248-52.
22. Cavalla D. Using human experience to identify drug repurposing opportunities: theory and practice. Br J Clin Pharmacol. 2019;85(4):680-9.
23. van der Velden DL, Hoes LR, van der Wijngaart H, van Berge Henegouwen JM, van Werkhoven E, Roepman P, et al. The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. Nature. 2019;574(7776):127-31.
24. Jarada TN, Rokne JG, Alhajj R. A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. J Cheminform. 2020;12(1):46.
25. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313(5795):1929-35.
26. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171(6):1437-52.e17.
27. Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med. 2011;3(96):96ra77.
28. Bhat-Nakshatri P, Goswami CP, Badve S, Sledge GW, Nakshatri H. Identification of FDA-approved drugs targeting breast cancer stem cells along with biomarkers of sensitivity. Sci Rep. 2013;3:2530.
29. van Noort V, Schölch S, Iskar M, Zeller G, Ostertag K, Schweitzer C, et al. Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene-expression profiling. Cancer Res. 2014;74(20):5690-9.
30. Chen HR, Sherr DH, Hu Z, DeLisi C. A network based approach to drug repositioning identifies plausible candidates for breast cancer and prostate cancer. BMC Med Genomics. 2016;9(1):51.
31. Chen B, Wei W, Ma L, Yang B, Gill RM, Chua MS, et al. Computational Discovery of Niclosamide Ethanolamine, a Repurposed Drug Candidate That Reduces Growth of Hepatocellular Carcinoma Cells In Vitro and in Mice by Inhibiting Cell Division Cycle 37 Signaling. Gastroenterology. 2017;152(8):2022-36.
32. Chen B, Ma L, Paik H, Sirota M, Wei W, Chua MS, et al. Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets. Nat Commun. 2017;8:16022.
33. Laganà A, Beno I, Melnekoff D, Leshchenko V, Madduri D, Ramdas D, et al. Precision Medicine for Relapsed Multiple Myeloma on the Basis of an Integrative Multiomics Approach. JCO Precis Oncol. 2018;2018.
34. Vásquez-Bochm LX, Velázquez-Paniagua M, Castro-Vázquez SS, Guerrero-Rodríguez SL, Mondragon-Peralta A, De La Fuente-Granada M, et al. Transcriptome-based identification of lovastatin as a breast cancer stem cell-targeting drug. Pharmacol Rep. 2019;71(3):535-44.