



Article

DDCM: A Computational Strategy for Drug Repositioning Based on Support-Vector Regression Algorithm

Manyi Xu ^{1,†}, Wan Li ^{1,†} , Jiaheng He ¹, Yahui Wang ¹, Junjie Lv ¹, Weiming He ², Lina Chen ^{1,*} and Hui Zhi ^{1,*}

¹ College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150000, China; myxu.china@gmail.com (M.X.); liwan@hrbmu.edu.cn (W.L.); j1ah3ng@gmail.com (J.H.); wanyahui@hrbmu.edu.cn (Y.W.); lvjunjie525@126.com (J.L.)

² Institute of Opto-Electronics, Harbin Institute of Technology, Harbin 150000, China; hewm@hit.edu.cn

* Correspondence: chenlina@ems.hrbmu.edu.cn (L.C.); zhihui@ems.hrbmu.edu.cn (H.Z.); Tel.: +86-0451-86674768 (L.C.)

[†] These authors contributed equally to this work.

Abstract: Computational drug-repositioning technology is an effective tool for speeding up drug development. As biological data resources continue to grow, it becomes more important to find effective methods to identify potential therapeutic drugs for diseases. The effective use of valuable data has become a more rational and efficient approach to drug repositioning. The disease–drug correlation method (DDCM) proposed in this study is a novel approach that integrates data from multiple sources and different levels to predict potential treatments for diseases, utilizing support-vector regression (SVR). The DDCM approach resulted in potential therapeutic drugs for neoplasms and cardiovascular diseases by constructing a correlation hybrid matrix containing the respective similarities of drugs and diseases, implementing the SVR algorithm to predict the correlation scores, and undergoing a randomized perturbation and stepwise screening pipeline. Some potential therapeutic drugs were predicted by this approach. The potential therapeutic ability of these drugs has been well-validated in terms of the literature, function, drug target, and survival-essential genes. The method’s feasibility was confirmed by comparing the predicted results with the classical method and conducting a co-drug analysis of the sub-branch. Our method challenges the conventional approach to studying disease–drug correlations and presents a fresh perspective for understanding the pathogenesis of diseases.

Keywords: drug repositioning; support-vector regression; hybrid matrix; potential therapeutic drugs



Citation: Xu, M.; Li, W.; He, J.; Wang, Y.; Lv, J.; He, W.; Chen, L.; Zhi, H. DDCM: A Computational Strategy for Drug Repositioning Based on Support-Vector Regression Algorithm. *Int. J. Mol. Sci.* **2024**, *25*, 5267. <https://doi.org/10.3390/ijms25105267>

Academic Editor: Bruno Rizzuti

Received: 29 February 2024

Revised: 25 April 2024

Accepted: 6 May 2024

Published: 12 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Drug repositioning is an effective and promising strategy for identifying new therapeutic indications for existing drugs. This approach has the potential to significantly enhance the efficiency and productivity of traditional drug discovery and development, ultimately leading to better health outcomes for patients [1]. Drug repositioning has proven to be a valuable solution for the treatment of certain diseases. While the initial instances of successful drug repositioning were largely serendipitous, there are now several examples of effective repositioning, such as acetylsalicylic acid [2,3], thalidomide [4,5], sildenafil [6,7] and dimethyl fumarate [8–10]. As the amount of available biomedical data continues to expand, the benefits of using computational methods to identify potential drug candidates for diseases of interest have become increasingly apparent [11,12]. In the realm of computational drug repositioning, the primary methods include text mining [13–17], network analysis [18–22], and machine learning [23–28]. The advantages of computational drug repositioning have been demonstrated by numerous researchers through the development of algorithms and experimental validation.

Text-mining advancements have facilitated the creation of association networks for biomedical entities, like genes, drugs, and diseases [13]. For its capability of integrating

multiple data sources, network analysis has been a research hot spot and has been utilized in the field of drug repositioning [18]. Machine learning is ideal for computational drug repositioning, as it can process large amounts of data and help identify associations between biological data [23]. By improving the robustness and accuracy of models, machine learning can lead to more effective drug-repositioning strategies. Machine-learning methods currently rely on drug- and disease-based searching to identify new drug–disease interactions. This strategy has emerged as the main approach for discovering potential drug candidates for diseases [29]. The construction of training and test sets is typically required for machine-learning methods. This involves using known disease–drug relationships to establish benchmarks for evaluating new drug candidates [30]. Both machine-learning and network-analysis methods rely on known disease–drug relationships to construct similarity matrices or networks. These are used to identify and evaluate potential drug candidates [31]. Using existing prior information on known disease–drug correlations when constructing similarity matrices or networks can impact the performance of machine learning and similarity network analysis [30]. In binary classification models, disease–drug relationships that are not present in existing data resources are included in the negative training set. The absence of documented associations does not necessarily mean that the association does not exist in biology. In fact, these yet-to-be-observed associations can be vital pieces of information for discovering new drugs. Furthermore, disease-pathway and drug-mechanism data are essential to understanding the reasons underlying the occurrence and progression of diseases, as well as the mechanisms through which drugs can treat such diseases [32].

As a supervised learning approach, machine learning can typically produce superior results for drug repositioning. DDCM is a machine-learning strategy based on the support-vector regression algorithm. Compared with other methods, DDCM considered more levels of data types, which has certain advantages for model construction. DDCM represented the similarity of related biological characteristics of drugs and diseases in the form of a matrix in the construction process, which ensured biological significance in the training process. The SVR-based DDCM approach covered the advantages of regression models. Unlike traditional classifiers, DDCM could describe the correlation between diseases and drugs in the form of continuous scores. This allowed DDCM to quantitatively characterize the association between disease and drug, overcoming the limitation of the traditional binary classification problem. For the diseases where no therapeutic relationship had been observed, the appropriate correlation score could be given for drug pairs, which solved the problem of negative set matching. This also allowed the DDCM method to maximize the inclusion of drugs that differ significantly from the known therapeutic drugs for the disease within the prediction. The prediction of potential therapeutic drugs for a disease could also be further expanded by the introduction of slack variables in the support-vector regression algorithm. Through a stepwise screening pipeline, the DDCM method could effectively reduce the influence of random error on prediction, and this method has been proven to be stable and flexible in multiple aspects.

2. Results

2.1. Potential Therapeutic Drugs for 59 Neoplasms

For these diseases in C04.588 branch of Medical Subject Headings (MeSH), the DDCM approach was applied to predict the potential therapeutic drugs (Figure 1A,B).

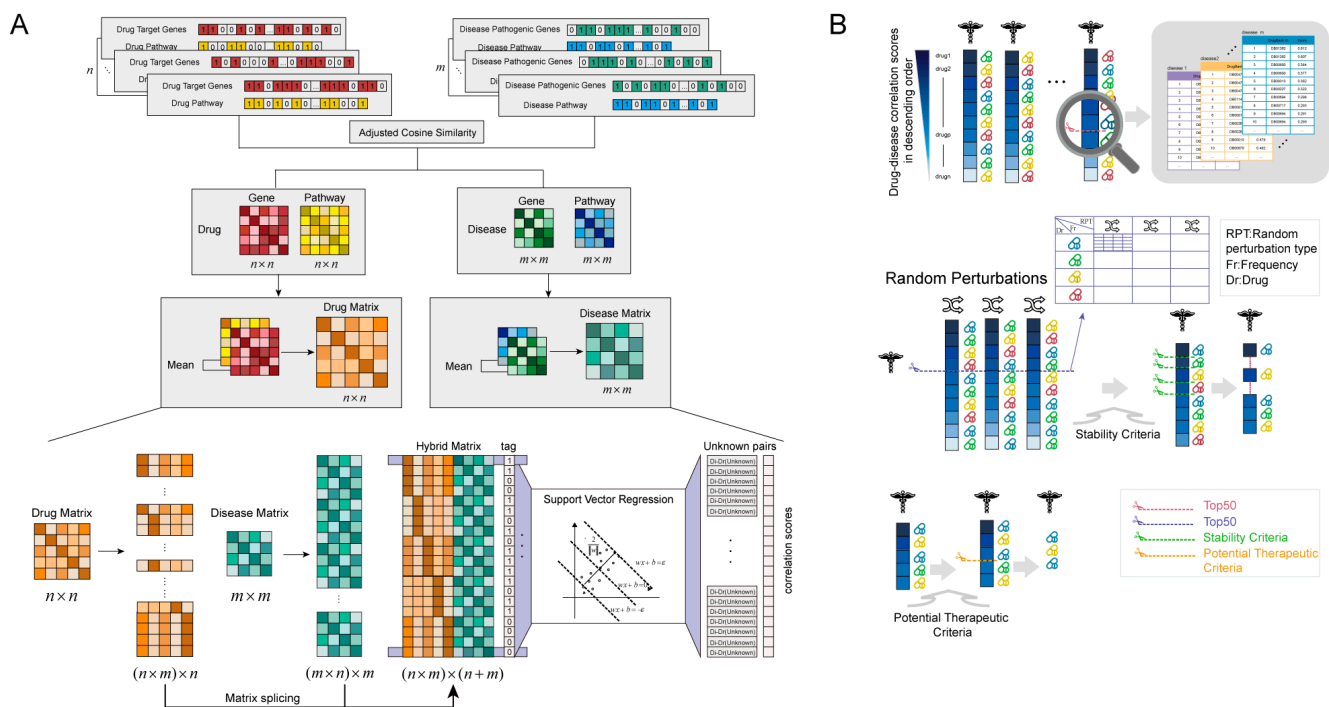


Figure 1. Workflow of the DDCM. (A) The process of forming a hybrid matrix and predicting disease–drug correlation scores in the DDCM approach. (B) The process of screening potential therapeutic drugs for a disease by means of a disease–drug correlation score in the DDCM approach.

The distributions of disease–drug correlation scores were relatively mainly distributed around zero; there were slight differences between them. The higher the positive score, the greater the degree of correlation between the drug and the disease, and the more likely the drug was to be a potential therapeutic drug for the disease (Figure 2A). The disease–drug correlation scores changed with the number and combination of diseases involved in the calculation (File S4 1_c04.588_CorScore). Therefore, a stepwise screening pipeline was adopted to obtain stable drug candidates and potential therapeutic drugs. The distribution of drug–disease correlation scores was derived for 59 neoplasms under three randomized perturbation types and for 59 neoplasms before randomization. The distribution of disease–drug correlation scores showed different degrees of fluctuation in different types of randomizations (Figure 2B). And, the more diseases were removed from the randomized perturbation experiments, the more fluctuations in the final disease–drug correlation scores were obtained than those when no diseases were removed. The stability score for each drug fluctuated across different random perturbation types. Thus, the intersection of the drugs obtained in each of the three random perturbation types was considered here (File S4 2_stability_scores_boxplot). Non-small-cell lung cancer (NSCLC) was an example. The stable drug candidates for non-small-cell lung cancer were screened by a stepwise screening pipeline. The stability scores of the 50 drugs to be screened in different states of random perturbation are shown (Figure 2C). The closer the stability score is to one, the more stable the drug is under such conditions. The stability scores in several random perturbation conditions were close to one, indicating that the volatility of the drug’s stability was low and more likely to be a potential therapeutic drug. The stepwise screening pipeline screened drugs with high stability and a low degree of stability fluctuation in different random states, resulting in 25 potential therapeutic drugs for NSCLC. There were potential therapeutic drugs for neoplasms other than gallbladder cancer and inflammatory breast carcinoma. These diseases were each predicted to range from 11–30 potential therapeutic drugs (Figure 2D). (File S2 Table S5).

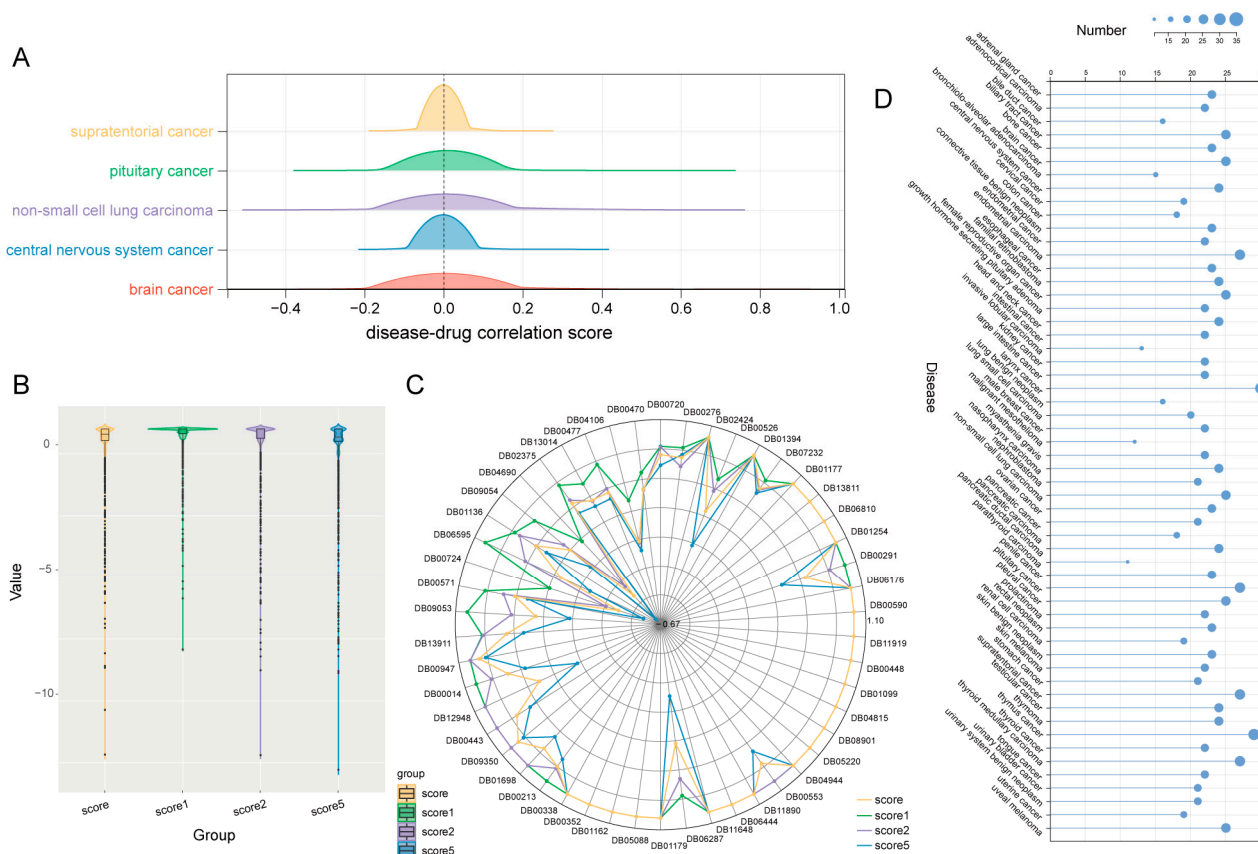


Figure 2. Potential therapeutic drugs for 59 neoplasms. **(A)** The distribution of disease–drug correlation scores for different diseases (The figure shows the distribution of disease–drug correlation scores for five diseases, supratentorial cancer, pituitary cancer, non-small-cell lung carcinoma, central nervous system cancer, and brain cancer). **(B)** The distribution of disease–drug correlation scores for the overall diseases in the C04.588 branch of MeSH before and after random perturbation. **(C)** Radar plot of stable drug candidates for NSCLC and the potential therapeutic drug screening criteria, where drugs with stability scores close to 1 multiple times were considered relatively stable. After the potential therapeutic screening criteria, a total of 25 potential therapeutic drugs for the final NSCLC were identified. **(D)** The number of potential therapeutic drugs for all diseases in the C04.588 branch.

2.2. Potential Therapeutic Capacity Analysis and Validation of Potential Therapeutic Drugs

The effectiveness of potential therapeutic drugs was evaluated in the literature aspect, the functional aspect, the drug-target aspect, and the survival analysis aspect.

2.2.1. The Literature Aspect

From the literature aspect, text-mining techniques were leveraged by mining keywords, whether a potential therapeutic drug for a disease has been investigated for the treatment of the disease, whether a potential therapeutic drug acts in combination with other drugs for the treatment of a disease, or whether the drug target of a drug has been shown to be a potential therapeutic target for the disease. For NSCLC, the literature validation rate of the screened potential therapeutic drugs reached 56.0% (14/25) (Supplementary Table S2).

2.2.2. The Functional Aspect

From the functional aspect, since the functional properties of drugs were such that they exert their effects through target genes, we focused on the effects of drugs through their target genes. In turn, known disease pathogenic genes can reveal pathogenesis. Thus, functional enrichment was performed for drug-target genes of potential therapeutic drugs and drug-target genes of known drugs for each disease.

We focused on genes associated with non-small-cell lung cancer, including drug targets for potential therapeutic drugs, drug targets for known drugs, and pathogenic genes. In an attempt to reveal the therapeutic mechanism of potential therapeutic drugs, these genes were mainly examined. This part of the gene was categorized into four groups (Group I consisted of five genes: genes in the overlapping part of drug targets of potential therapeutic drugs and drug targets of known drugs and pathogenic genes. Group II consisted of seven genes: genes in the overlapping part of drug targets of potential therapeutic drugs and drug targets of known drugs. Group III consisted of ten genes: genes in the overlapping part of drug targets of potential therapeutic drugs and pathogenic genes. Group IV consisted of thirty-six genes: drug targets of potential therapeutic drugs only) (Figure 3A). An enrichment analysis was performed separately for potential therapeutic drugs and drug targets of known therapeutic drugs for NSCLC (Figure 3B,C).

Functional enrichment was performed for each set of genes. Partial results of some major genes in each group are shown (Figure 3D–G). (File S2 Table S7).

In Group I, where the mechanism of pathogenesis and the potential therapeutic mechanism were similar to the known therapeutic mechanism, the literature validation rate of potential therapeutic drugs was 50.0% (3/6). In Group II, where the mechanism of treatment was similar to that of known therapeutic drugs, the literature validation rate of potential therapeutic drugs was 66.7% (4/6). In Group III, where the mechanism of pathogenesis and potential therapeutic mechanism were similar, the literature validation rate of potential therapeutic drugs was 75.0% (3/4). In Group IV, where the potential therapeutic mechanism was specific, the literature validation rate of potential therapeutic drugs was 70.0% (7/10).

The results of the enrichment analysis showed that, for the drug targets of potential therapeutic drugs for NSCLC, all groups of genes were enriched in peptidyl-tyrosine phosphorylation (GO:0018108), peptidyl-tyrosine modification (GO:0018212), positive regulation of kinase activity (GO:0033674), and other such GO terms. The enrichment of drug-target genes for known drugs in NSCLC can be seen in the enrichment of some genes to these GO terms as well. It was known that Groups III and IV do not overlap with the drug-target genes of known drugs. Genes in Group IV do not overlap with the pathogenic genes, which were similarly enriched in a common functional class with the drug targets of known drugs, suggesting that these potential therapeutic drugs, although targeting different genes than known drugs, may perform the same or similar functions through different biological pathways.

Klarisa Rikova et al. characterize tyrosine kinase signaling across 41 non-small-cell lung cancer (NSCLC) cell lines and over 150 NSCLC tumors. Profiles of phosphotyrosine signaling are generated and analyzed to identify known oncogenic kinases, such as EGFR and c-Met, as well as novel ALK and ROS fusion proteins. By focusing on activated cell circuitry, the approach provides insight into cancer biology not available at the chromosomal and transcriptional levels and can be applied broadly across all human cancers [33]. The lymph node metastasis of NSCLC exclusive somatic mutation (LME-SMs) genes was enriched in peptidyl-tyrosine phosphorylation, peptidyl-tyrosine modification, protein tyrosine kinase activity, and transmembrane receptor protein kinase [34]. Santos GdC et al. have proven that tyrosine kinase inhibitors (TKIs) are especially effective in patients whose tumors harbor activating mutations in the tyrosine kinase domain of the *EGFR* gene. More recent trials have suggested that, for advanced NSCLC patients with *EGFR* mutant tumors, initial therapy with a TKI instead of chemotherapy may be the best choice of treatment [35]. And since the drug-target genes of these groups were enriched with the drug-target genes of known drugs on these GO terms, it can be inferred that the potential therapeutic drugs can be therapeutic drugs for NSCLC from the functional aspect. As a whole, the drug targets predicted for potential therapeutic drugs for NSCLC can play a corresponding role in the functional aspect. In terms of each drug individually, the corresponding drugs were also useful for the treatment of NSCLC from different perspectives.

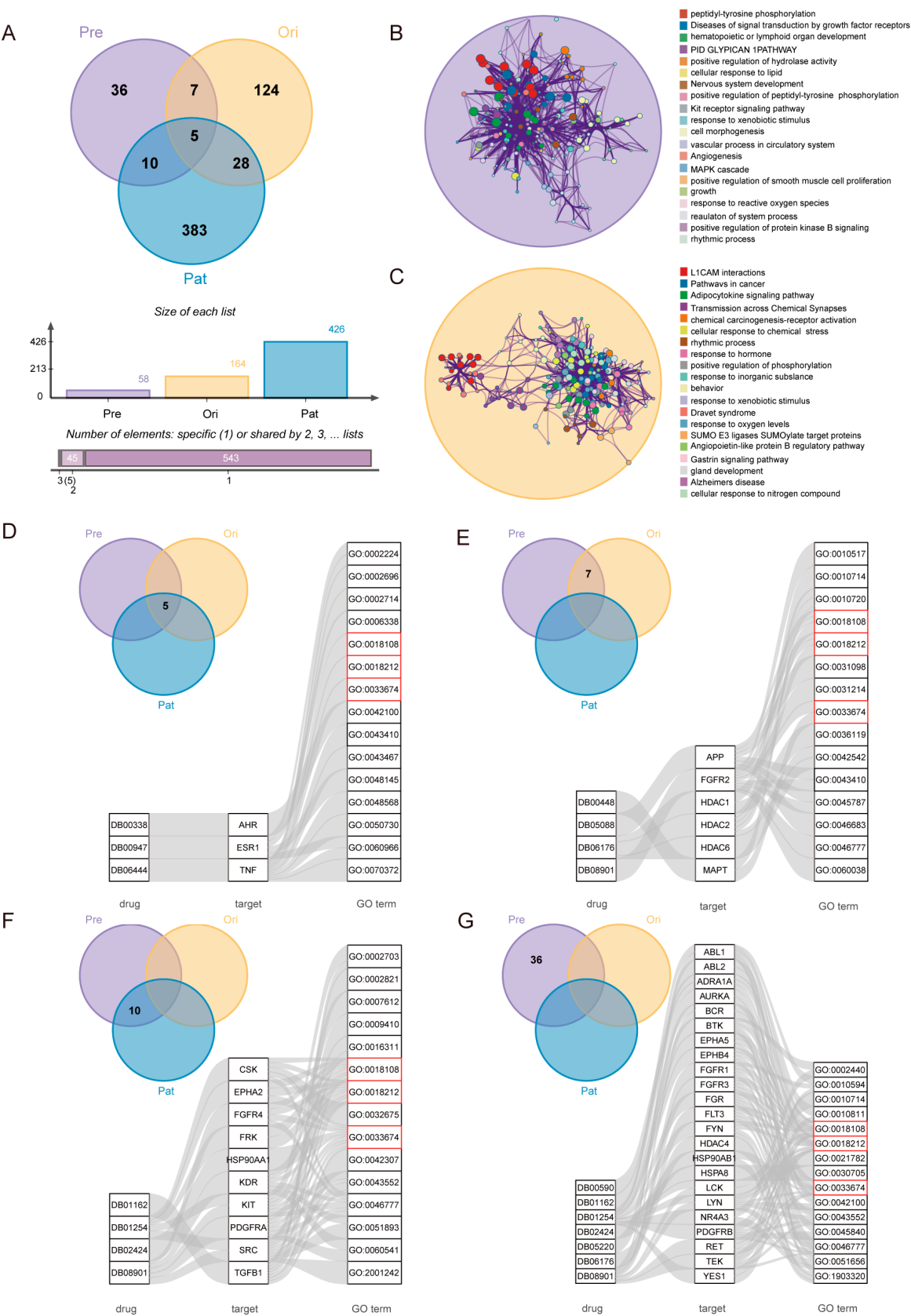


Figure 3. The functional aspect. (A) Venn diagram of drug targets of potential therapeutic drugs, drug targets of known therapeutic drugs, and pathogenic genes for NSCLC (Pre: potential therapeutic drug

targets, Ori: known drug targets, and Pat: pathogenic genes of NSCLC). (B) Drug-target enrichment analysis of potential therapeutic drugs for NSCLC. (C) Drug-target enrichment analysis of known therapeutic drugs for NSCLC. (D–G) Partial Gene Ontology (GO) enrichment of drug-target subgroups for potential therapeutic drugs in NSCLC. (Some genes were not enriched on GO terms) (D) Group I. (E) Group II. (F) Group III. (G) Group IV. (GO terms enriched for all four groups of drug target genes were marked with red boxes.)

In the results of drug-target enrichment analysis for Groups II, III, and IV, we can find that these drug targets were enriched to the function of protein autophosphorylation (GO:0046777) in addition to the common three functional classes, indicating that potential therapeutic drugs and known therapeutic drugs also play a role in the treatment of disease through the pathogenic mechanism of the disease. Jing-Qiang Huang et al. suggested that the SRPK1/GSK3 β axis promotes gefitinib resistance by activating the Wnt pathway in the form of autophosphorylation and may serve as a potential therapeutic target for overcoming gefitinib resistance in NSCLC [36].

For the additional enrichment of Groups II and IV to the common function positive regulation of collagen metabolic process (GO:0010714), potential therapeutic drugs and known therapeutic drugs have a similar therapeutic pattern. It is possible that the therapeutic effect was exerted by affecting this biological pathway. Hao Chang et al. suggested that the overexpression of collagen XVIII was associated with NSCLC progression and a poor outcome. Thus, collagen XVIII expression may serve as a useful prognostic marker in patients with NSCLC [37].

For Groups III and IV jointly enriched beyond the three GO terms, we identified positive regulation of phosphatidylinositol 3-kinase activity (GO:0043552), potential therapeutic drugs that exert a therapeutic effect on the pathogenic mechanism of the disease, and this therapeutic effect is unique relative to known therapeutic drugs. Panpan Zhang et al. proved when phosphatidylinositol 3-kinase/protein kinase B signaling was inhibited by corresponding inhibitors, PD-L1 expression was downregulated, and apoptosis was upregulated in the cisplatin-treated cancer cells. These results suggest that the upregulation of PD-L1 promotes a resistance response in lung cancer cells that might be through activation of the phosphatidylinositol 3-kinase/protein kinase B pathway and suppression of tumor-infiltrating lymphocytes [38].

In addition, for Group IV genes (*BTK/LYN*), the enriched GO term toll-like receptor signaling pathway (GO:0002224) was shown to be associated with NSCLC. Toll-like receptors (TLR) recognize pathogen molecules and danger-associated signals that stimulate inflammatory processes. TLRs have been studied mainly in antigen-presenting cells, where they exert important immune regulatory functions, but they were also expressed by epithelial tumor cells, where they have been implicated in tumor progression. Saradiya Chatterjee et al. proved that, in patients with non-small-cell lung cancer, expression analyses revealed that high TLR7 expression was strongly associated with resistance to neoadjuvant chemotherapy and poor clinical outcomes [39]. *BTK* was targeted by DB01254, and *LYN* was targeted by DB01254 and DB08901, suggesting that both drugs may exert the pharmacological effects of drugs from the perspective of inhibiting the high expression of toll-like receptors.

2.2.3. The Drug-Target Aspect

For the known drug-target genes and drug targets of potential therapeutic drugs for non-small-cell lung cancer, according to KEGG Mapper, most of their drug targets are involved in cancer-related pathways, both for known therapeutic drugs and potential therapeutic drugs. And, the drug targets of potential therapeutic drugs (including four groups of genes) are mainly involved in the PI3K-Akt signaling pathway and the Jak-STAT signaling pathway of the cancer pathway (Figure 4).

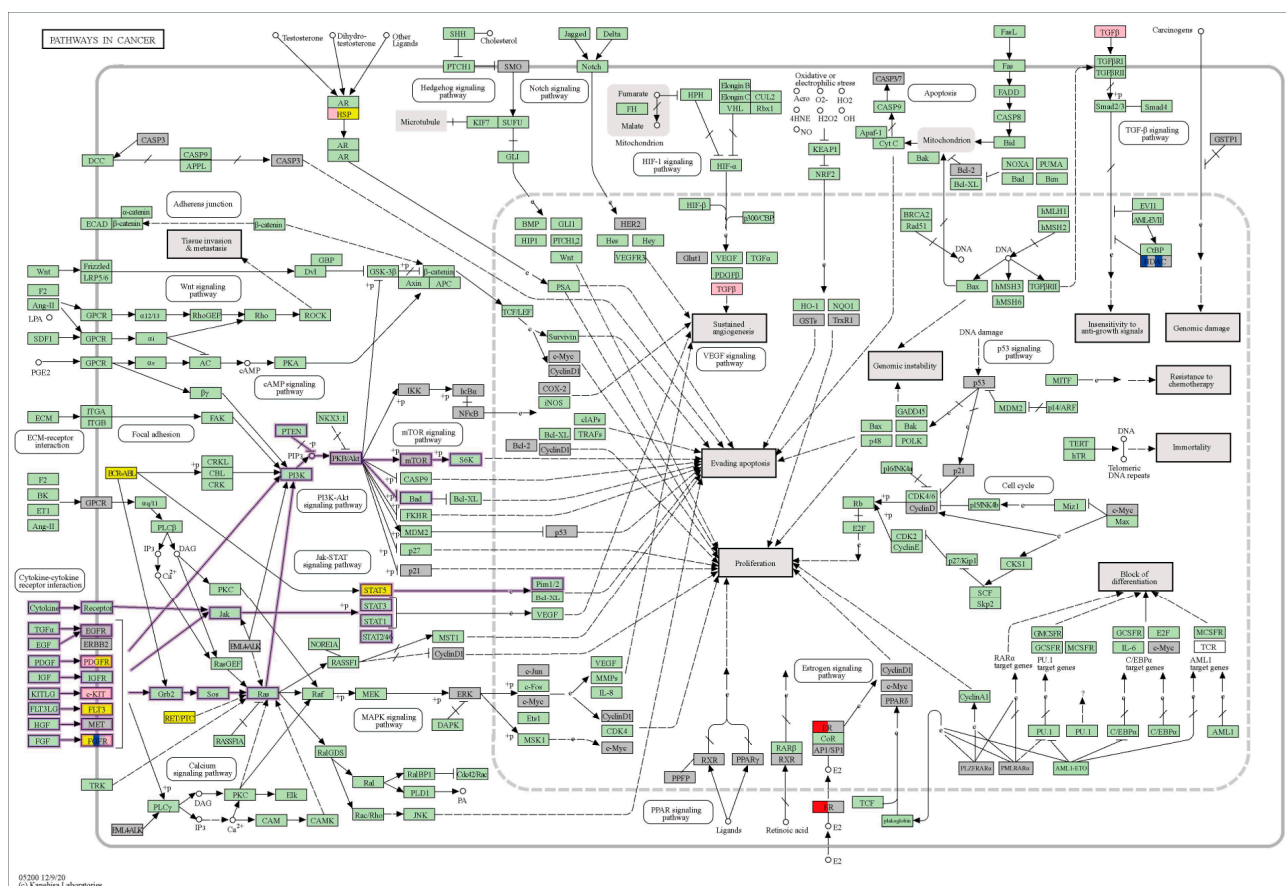


Figure 4. The drug-target aspect. Involvement of various groups of drug-target genes in cancer pathways as shown by KEGG Mapper. (Red: Group I; blue: Group II; pink: Group III; yellow: Group IV; grey: target genes of known drugs; pathways labeled with purple outer box lines are those in which the drug-target genes of potential therapeutic drugs are predominantly involved).

Except for Group I, which was available for all aspects of genes, when a drug targets genes in Groups II, III, and IV, the drug should have a similar therapeutic mechanism to known drugs and be associated with the pathogenic mechanism of the disease. And in some aspects, it has its own unique therapeutic mechanism. For example, ponatinib (DB08901) is a kinase inhibitor used to treat patients with various types of chronic myeloid leukemia (Drugbank: ponatinib). These signaling pathways play an important role in the development of cancer. Ren M et al. demonstrated that pharmacological inhibition of FGFR1 kinase activity with ponatinib may be effective for the treatment of lung cancer patients whose tumors overexpress *FGFR1* [40]. Ponatinib targets *FGFR2* and exerts the same therapeutic effect as known drugs for NSCLC. *FGFR2* and other FGFR kinase family gene alterations have been found in both lung squamous cell carcinoma and lung adenocarcinoma [41]. Also, ponatinib targets genes in parts that overlap with pathogenic genes such as *KDR*. The clinical outcomes of patients with advanced NSCLC receiving first-line bevacizumab plus chemotherapy regimens might be impacted by polymorphism V297 L through mediating the mRNA expression of *KDR* [42]. In addition, ponatinib targets potential therapeutic drug-specific targets such as *ABL1*. Mutations in *ABL1* identified in primary NSCLC tumors and a lung cancer cell line increase downstream pathway activation compared to wild-type *ABL1* [43].

When a drug targets genes in Groups III and IV, the drug should have its own unique therapeutic mechanism for the pathogenic mechanism of the disease. For example, dasatinib (DB01254) is an oral dual BCR/ABL and SRC family tyrosine kinase inhibitor approved for use in patients with chronic myelogenous leukemia (DrugBank: dasatinib). It

inhibited migration and invasion, induced cell-cycle arrest, and partial apoptosis in NSCLC cell lines. Dasatinib is a drug with multiple targets (these target genes are mainly located in Groups III and IV), which works by inhibiting the metastasis of epithelial cells and affecting related signaling pathways to treat non-small-cell lung cancer. It was demonstrated that this drug will be a unique therapeutic idea for potential therapeutic drugs in NSCLC. Molecular alterations of *YES1*, a member of the SRC, can be found in a significant subset of patients with lung cancer. Moreover, high *YES1* protein expression was an independent predictor for poor prognosis in patients with lung cancer [44]. Phase I/II studies have shown that dasatinib in combination with erlotinib is safe and feasible for the treatment of NSCLC [45,46]. Some other targets targeted by dasatinib, such as *PDGFRB* [47] and *MAPK*, have been validated in the literature accordingly.

Romidepsin (DB06176) is a histone deacetylase (HDAC) inhibitor used to treat cutaneous T-cell lymphoma (DrugBank: romidepsin). The drug targets of the drug were mainly in Group IV and Group II. Theoretically, the therapeutic mechanism of the drug has similarities to known drugs as well as its own unique aspects. The drug has some similarities with the known drugs for NSCLC in terms of the mechanism of cure. There were also some individual therapeutic mechanisms (e.g., branching morphogenesis of an epithelial tube, cell–substrate adhesion, etc.). Karthik S et al. proved that romidepsin and bortezomib cooperatively inhibit A549 NSCLC cell proliferation by altering the histone acetylation status and the expression of cell-cycle regulators and MMPs. Romidepsin along with bortezomib might be an effective treatment approach for A549 NSCLC cells [48]. This drug also has an impact on improving the chemosensitivity of NSCLC by targeting the *ABCC1* gene [49,50].

In addition to this, if a drug targets only Group IV, the potential therapeutic drug may have an impact on the treatment of the disease from a unique therapeutic mechanism. Alisertib (DB05220) is a novel aurora A kinase inhibitor under investigation for the treatment of various forms of cancer (DrugBank: alisertib). Alisertib is an oral aurora kinase inhibitor that has been shown to induce cell-cycle arrest and apoptosis in preclinical studies. It is currently under investigation for a wide variety of malignancies, including hematologic and solid tumors [51]. Wang et al. found that an AURKA inhibitor, alisertib, treatment restored the susceptibility of resistant cells to EGFR-TKIs and partially reversed the EMT process, thereby reducing migration and invasion in EGFR-TKI-resistant cells. This study provides evidence that targeting the AURKA signaling pathway by alisertib may be a novel approach for overcoming EGFR-TKI resistance and for the treatment of metastatic EGFR-TKIs in NSCLC patients [52].

Among the potential therapeutic drugs for NSCLC, there are some drugs that are temporarily unproven in the literature.

Flucytosine (DB01099) is an antifungal indicated only to treat severe infections throughout the body caused by susceptible strains of *Candida* or *Cryptococcus* (DrugBank: flucytosine). Flucytosine targets the *DNMT1* gene in Group I. *DNMT1* is mainly enriched to the GO:0033002 functional class, namely muscle-cell proliferation, and Jia et al. demonstrated that NSCLC-derived exosomes promote cell proliferation and inhibit cell apoptosis in both normal lung fibroblasts and NSCLC cells by delivering ASMA. High expression of *DNMT1* protein in serum may increase the pathogenesis of non-small-cell lung cancer and may play an important role in the early development of lung cancer. Genes in Group I were mainly enriched to some cancer-related signaling pathways, immune responses, and cell-differentiation-related functions. Some of the literature has confirmed the association of these functional classes with non-small-cell lung cancer [53].

Terazosin (DB01162) is an alpha-1 adrenergic antagonist used in the treatment of symptomatic benign prostatic hyperplasia and the management of hypertension (DrugBank: terazosin). Terazosin targets *TGFB1* in Group III and *ADRA1A*, *ADRA1B*, *ADRA1D*, *KCNH2*, *KCNH6*, and *KCNH7* in Group IV. For *TGFB1*, Ki-Eun Hwang confirmed that TGF- β 1 induces EMT, which leads to lung cancer cell migration and invasion. It is possible to treat non-small-cell lung cancer by inhibiting this aspect of the *TGFB1* gene [54].

Colchicine (DB01394) is an alkaloid used in the symptomatic relief of pain in attacks of gout and to treat the inflammatory symptoms of Familial Mediterranean Fever (DrugBank: colchicine). Colchicine targets the *TUBB* gene in Group I. Knockdown of *TUBB* sensitizes cells to MTAs, while overexpression confers resistance. A high expression of *TUBB* is correlated with worse survival of lung adenocarcinoma [55]. The *TUBB* gene is primarily enriched for immune-related neutrophil activity. Proinflammatory cytokines are centrally involved in tumor progression and survival in non-small-cell lung cancer, and both the presence of infiltrating neutrophils and bacterial infection in the lung may indicate a poor prognosis. Interaction between bacterial pathogens, neutrophils, and tumor cells might amplify the release of proinflammatory cytokines, which may promote tumor growth in vivo [56].

2.2.4. The Survival-Essential Gene Aspect

Furthermore, survival analysis is widely used in clinical and epidemiological research. Its use in the contemporary medical literature is widespread [57]. The Cancer Dependency Map (DepMap; <https://depmap.org/portal/>; 21Q4 Public+Score, Chronos, released on 3 November 2021) utilizes RNA interference (RNAi) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9) technologies to screen for the necessity of various genes for tumors. DepMap analyzes hundreds of cancer cell-line models to obtain information on the genome of individual cell lines and their sensitivity to genetic and small molecule perturbations. Multiple analyses and studies of tumor cells were attempted to identify the genetic and pharmacological dependence of tumors as well as predict their biomarkers. Based on the potential therapeutic drugs, we performed the identification of survival-essential genes for each group of genes for these neoplasms. For NSCLC, several subgroups of drug targets for potential therapeutic drugs can be achieved by a CRISPR knockout screen, which represents the impact of knocked-out genes on the survival of cancer cell lines [58]. Typically, the cut-off value will be set to -0.5 for gene-effect scores, indicating a significant depletion of cell lines [59].

The gene effects of 94 NSCLC cell lines derived from the genome-wide CRISPR knockout screens were stored in the DepMap database. Each group of genes was screened for survival-essential genes by this method. A total of 58 genes were involved in the identification, of which 17 were essential for survival (Figure 5) (File S3). Targeting these genes, which were essential for survival, could have the effect of treating neoplasms. Further, we screened for drugs corresponding to these survival-essential genes (Table 1). We determined that these drugs that target survival-essential genes were more likely to have an impact on the survival of cancer cell lines and were more likely to be therapeutic drugs for NSCLC.

Among the potential therapeutic drugs corresponding to these survival-essential genes, the literature validation rate was 70.0% (7/10). Among them, the literature validation rate of potential therapeutic drugs in Group I was 50.0% (2/4), in Group II was 50.0% (2/4), in Group III was 100.0% (3/3), and in Group IV was 100.0% (4/4).

Table 1. The drugs corresponding to the survival-essential genes in each group.

Group	Survival-Essential Genes	Drugs
I	<i>TOP2A, DNMT1, AHR</i>	DB01177/DB01179/DB01099/DB00338
II	<i>TUBB, FGFR2, HDAC2</i>	DB01179/DB01394/DB08901/DB06176
III	<i>SRC, PDGFRA, HSP90AA1, CSK</i>	DB01254/DB08901/DB02424
IV	<i>PPAT, STAT5B, AURKA, MAPK14, HSP90B1, HDAC4, HSPA8</i>	DB01254/DB05220/DB02424/DB06176

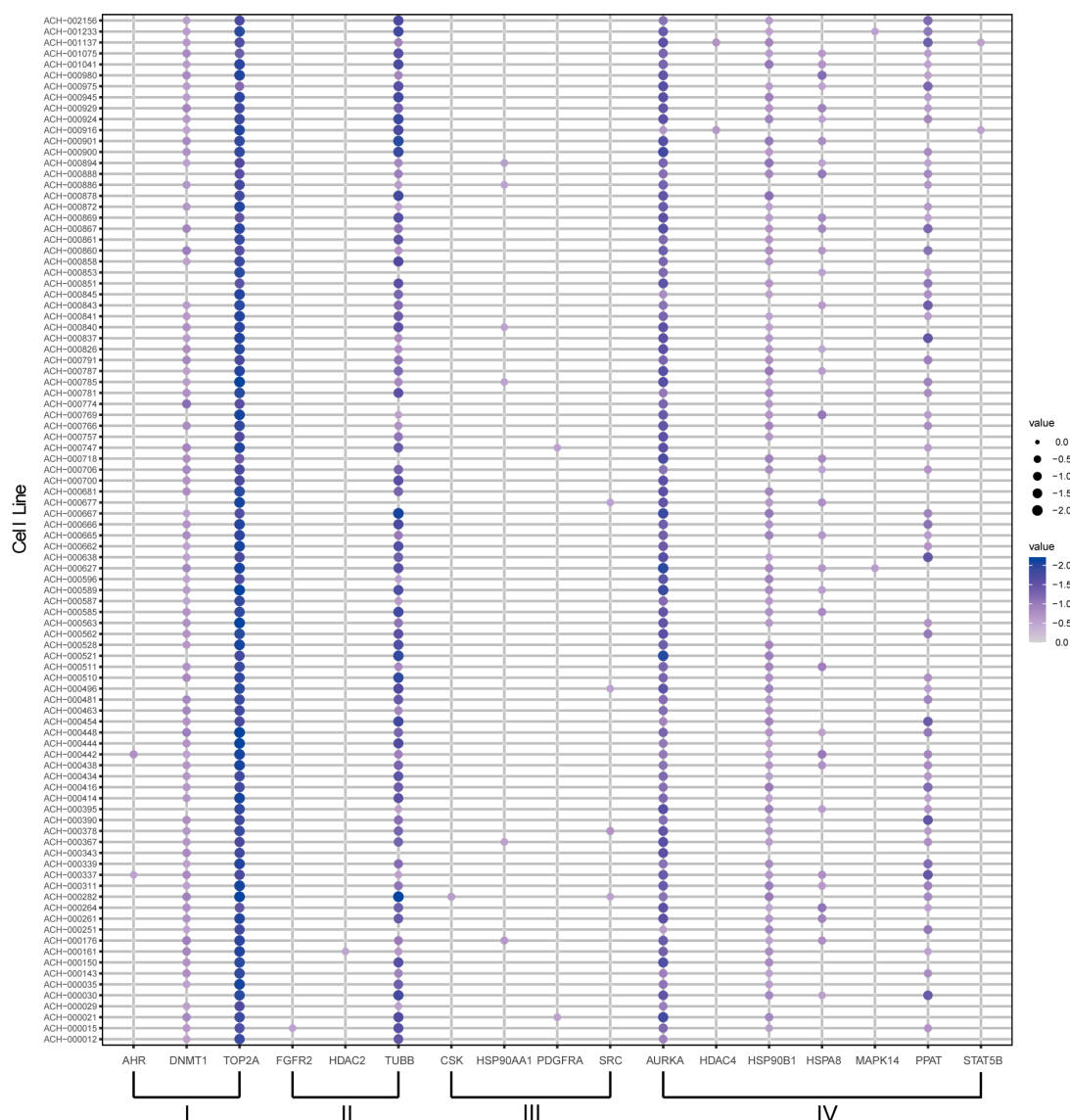


Figure 5. The survival-essential gene aspect. NSCLC cell lines corresponding to survival-essential genes.

Podofilox (DB01179) is a topical agent used for the treatment of external genital warts and perianal warts. *TOP2A* and *TUBB* targeted by podofilox were located in Groups I and II, respectively, and these two genes were identified as survival-essential genes in 94 cell lines and 88 cell lines, respectively. The description of podofilox in DrugBank is that it may have antineoplastic properties, as do some of its congeners and derivatives. (DrugBank: podofilox). Guo et al. proved that podophyllotoxins (podofilox), including epipodophyllotoxin derivatives, can act on a diverse array of drug targets in cancer cells and, thus, possess potent activity against various forms of cancer cell lines, including drug-resistant forms [60]. Likewise, flucytosine (DB01099) targets *DNMT1* in Group I, which was considered to be a survival-essential gene in 80 cell lines. It has been combined with newer azole antifungal agents. It also plays an important role in a new approach to the treatment of cancer [61]. Colchicine (DB01394) targets *TUBB* in Group II, which was considered a survival-essential gene in 88 cell lines. It was concluded that an untapped potential exists for exploiting the colchicine scaffold as a pharmacophore, with the possibility of increasing its affinity for tubulin isotypes overexpressed in cancer and decreasing it for normal cells, thereby widening the therapeutic window [62].

Based on the performance of the drug targets of these drugs in knockout screens for survival-essential genes, as well as descriptions in DrugBank and various studies suggesting that although these drugs have not been directly demonstrated in the literature to have therapeutic effects in NSCLC, their pharmacological effects make them promising potential therapeutic drugs for NSCLC.

From a survival perspective, we can conclude that the genes targeted by certain unproven drugs (DB01099, DB01179, DB01394, etc.) in the literature may be survival-related genes. Thus, these unproven drugs prove their importance not only from a functional point of view but also from the point of view of targeting genes essential for survival, suggesting that such drugs may be potential therapeutic drugs.

2.3. Drug Research of Similar Neoplasms in MeSH Branches

Usually, similar diseases tend to have the same susceptible genes or therapeutic targets [63,64] and similar drugs combined with similar drug targets [65]. MeSH provides a consistent way to find content with different terms but the same concepts. MeSH organizes its descriptors into a hierarchy. Diseases under the same branch have a higher similarity. To further confirm the validity of the DDCM approach, one of the sub-branches of neoplasms was chosen. These four diseases were central nervous system cancer, brain cancer, supratentorial cancer, and pituitary cancer in the same branch based on the stepwise relationship of the MeSH branch (Figure 6A). All four diseases have some overlap in terms of known therapeutic drugs and potential therapeutic drugs. Some of these potential therapeutic drugs were even confirmed in certain aspects. The co-drugs analysis of the MeSH sub-branch has also demonstrated the validity of the DDCM approach from the other side.

Considering the genetic aspect, the overlap among the drug-target genes of the screened potential therapeutic drugs, the target genes of the known drugs, and the pathogenic genes of the diseases was examined for the four diseases. It could be found that the degree of overlap between the screened potential therapeutic drugs and the drug targets of the known drugs for these four similar diseases was relatively large, while the degree of overlap with the pathogenic genes was smaller (Figure 6B(a–d)). This indicated that the therapeutic mechanisms of both screened potential therapeutic drugs and known therapeutic drugs were similar, while they were different from the pathogenic mechanisms of the diseases. In addition, the drug targets of the known therapeutic drugs, the drug targets of the screened potential therapeutic drugs, and the pathogenic genes of the four diseases overlap to some extent (Figure 6C–E), indicating that the four diseases were similar in their respective therapeutic and pathogenic mechanisms. This was further illustrated by the similarity of the four diseases, since they were taken from the same branch of MeSH. Combined with the previous results, similar diseases have drugs in common, so potential therapeutic drugs for these diseases and known therapeutic drugs should each have some degree of overlap.

The analysis of known therapeutic drugs and potential therapeutic drugs for these diseases showed that there was indeed some degree of overlap between the known therapeutic drugs for these diseases. Also, by predicting potential therapeutic drugs for these diseases, common therapeutic drugs could be found between these four cancers (Figure 6F,G).

Among the screened potential therapeutic drugs for the four diseases, plicamycin (DB06810) was one of the screened potential therapeutic drugs common to central nervous system cancer, brain cancer, supratentorial cancer, and pituitary cancer. Plicamycin is an antineoplastic antibiotic produced by *Streptomyces plicatus* (DrugBank: Plicamycin); 6-O-benzylguanine (DB11919) was predicted to be the potential therapeutic drug for brain cancer, central nervous system cancer, and supratentorial cancer [66]. Many current experimental studies have demonstrated the usefulness of the drug in the treatment of brain cancer, central nervous system cancer, and other conditions [67–70]. The co-drug analysis of potential therapeutic drugs for diseases of the same MeSH branch not only strengthened the conclusion of co-drugs for similar diseases but also highlighted the validity of the method laterally.

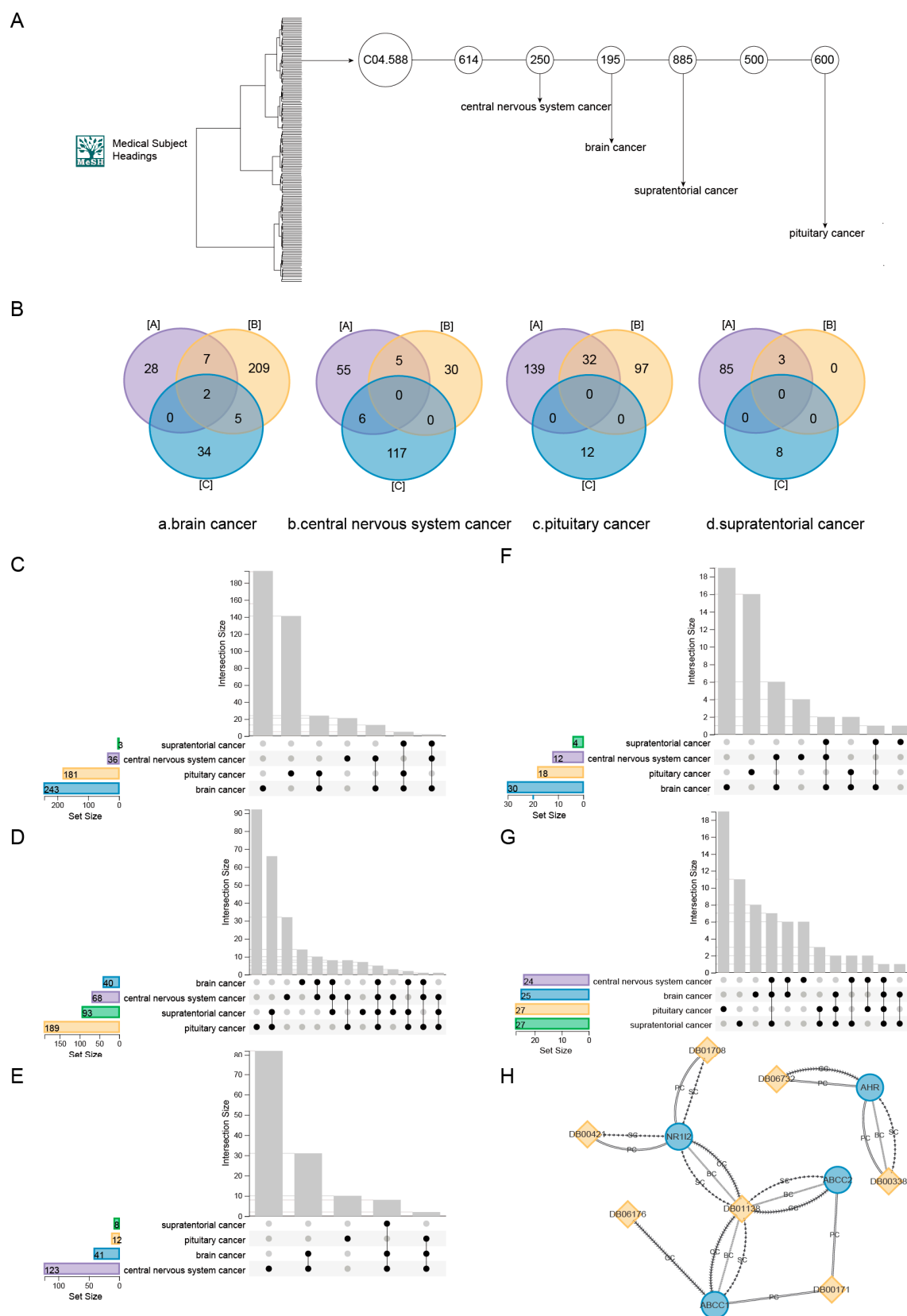


Figure 6. Drug research of similar neoplasms in MeSH branches. (A) The relationship between four cancers (brain cancer, central nervous system cancer, pituitary cancer, and supratentorial cancer) under the C04.588 branch of MeSH. (B) (a–d) The Venn diagrams indicate the degree of overlap of the screened potential therapeutic drug targets, drug targets of known drugs, and disease pathogenic

genes for brain cancer, central nervous system cancer, pituitary cancer, and supratentorial cancer, respectively ([A] Drug-target genes of potential therapeutic drugs predicted by the DDCM approach. [B] Drug-target genes of known drugs for a disease. [C] Pathogenic genes for a disease.). (C–E) The overlap of drug targets of known therapeutic drugs, drug targets of potential therapeutic drugs, and pathogenic genes for each of the four diseases, respectively. (F) Overlap of known therapeutic drugs for the four diseases. (G) Overlap of the number of potential therapeutic drugs predicted by the DDCM method for the four diseases. (H) Four genes traced from four disease perspectives to predict potential therapeutic drug profiles. The blue circle indicates the gene and the yellow diamond indicates the drug. BC, CC, PC, and SC denote the four diseases (brain cancer, central nervous system cancer, pituitary cancer, and supratentorial cancer) of the retrospective approach, respectively.

It was observed that four genes, *AHR*, *ABCC1*, *ABCC2*, and *NR1I2*, were the targets of the screened potential therapeutic drugs for all four diseases with a higher degree of overlap compared to the known drug targets. From these four drug-target genes, potential therapeutic drugs were traced from different disease perspectives (Figure 6H). Among the four genes, we mainly focused on the *AHR* gene. Emerging evidence suggests the promoting role of the *AHR* in the initiation, promotion, progression, invasion, and metastasis of cancer cells [71]. *AHR* activates the expression of multiple phase I and II xenobiotic chemical metabolizing enzyme genes (such as the *CYP1A1* gene) involved in cell-cycle regulation. Beta-Naphthoflavone (DB06732), which targets *AHR*, has been shown to regulate cell differentiation in the central nervous system [72].

3. Discussion

The computational drug-repositioning approach is a successful and promising technique to efficiently discover potential novel therapeutic drugs for diseases and new indications for existing drugs. The potential therapeutic drugs obtained can expand the therapeutic drugs available for the disease and expand the more multifaceted indications of known drugs. We constructed an SVR-based DDCM strategy to predict potential therapeutic drugs for the diseases. The approach was successfully applied in the prediction of potential therapeutic drugs for 59 diseases in the MeSH neoplastic branch. The number of potential therapeutic drugs ranges from 11 to 30. Some well-proven drugs emerged by analyzing the feasibility of potential therapeutic drugs in terms of the literature, function, drug targets, and survival-essential genes, respectively. And we also found similar drugs for diseases in the subbranches of the MESH branch.

In the strategy, SVR calculated continuous correlation scores between diseases and drugs through learning and fitting. The distribution of disease–drug correlation scores predicted by the SVR algorithm was roughly distributed on both sides of zero, but the overall distribution was different from each other. It was believed that in the process of constructing the hybrid matrix, the similarity between diseases and diseases, and the similarity between drugs and drugs was calculated by the design vector, and the number of disease pathogenic genes for each disease or drug-target genes for each drug was different. So, the calculated correlation scores were somewhat different. Thus, we considered the disease–drug correlation scores for each disease separately and conducted subsequent screening operations to derive potential therapeutic drugs instead of performing a cross-sectional comparison to take a percentage threshold as a screening condition for drug candidates. For each disease, after a step of stratified screening, the relevant drug obtained has a higher and more stable correlation with the disease, and we believe that drugs at this point already have the possibility to be used as potential therapeutic drugs for the disease.

Compared with traditional machine-learning methods, such as support-vector machine (SVM), K-nearest neighbor (KNN), and random walk (RW), DDCM methods have certain advantages in terms of performance ($AUC = 9.96 \times 10^{-1}$) (Supplementary Table S4). Similar to the results of Gottlieb et al. [27] in predicting potential therapeutic drugs for colon cancer, the PREDICT method of Gottlieb et al. and the DDCM approach simultaneously

predicted tioguanine (DB00352), fludarabine (DB01073), dasatinib (DB01254), colchicine (DB01394). As well as predicting the potential therapeutic drugs for testicular cancer and comparing the PREDICT method, we found that the commonly predicted drugs were colchicine (DB01394), idarubicin (DB01177), etc. This also proved that the predicted results of our method were convincing to some extent; these drugs might have the potential to treat the corresponding diseases to some extent.

We confirmed the generalizability of the DDCM approach by predicting potential therapeutic drugs for diseases in other branches of the MeSH. In the prediction of potential therapeutic drugs for 49 vascular diseases, the number of drugs predicted ranged from 14 to 26. We likewise found drugs that were well-validated (File S1 Document S1, Supplementary Figures S1–S7, Supplementary Table S3, File S2 Table S6). An analysis of genes associated with cerebrovascular disease revealed that the genes associated with cerebrovascular disease were divided into several gene families, and these drug-targeted gene families have been demonstrated in the literature to be associated with the treatment of cerebrovascular disease. A typical example was tranylcypromine, which can modulate monoaminergic transmission. It suggests that this drug is a promising lead compound for the further development of drugs to be used in therapy for cerebrovascular and neurological diseases. Similarly, in vascular disease, we analyzed potential therapeutic drugs for a particular branch of the disease and found common potential and known therapeutic drugs between the branches. The prediction of vascular diseases further validated the feasibility and generalizability of the DDCM.

4. Material and Methods

To screen potential drug candidates for the treatment of specific diseases, we developed the DDCM approach. A hybrid matrix of disease–drug correlative relationships was constructed by assessing the similarity of multiple levels for diseases and drugs. For a specific disease, the support-vector regression algorithm was used to calculate the disease–drug correlation score and to further screen the potential therapeutic drugs with stable predictions for that disease by setting stability criteria and potential therapeutic criteria. The effectiveness of potential therapeutic drugs was assessed at the levels of the literature, function, drug target, and survival-essential genes.

4.1. Data Retrieval

Drug–target gene data from the DrugBank database (<https://go.drugbank.com/>; version 5.1.8, released on 3 January 2021) [66] were selected. Meanwhile, drug–pathway data were obtained from the Comparative Toxicogenomics Database (CTD, <http://ctdbase.org/>; released on 5 August 2021) [73] and integrated with drug-related pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG, <https://www.genome.jp/kegg/>; version 98.1, released on 1 May 2021) [74]. The drug names in this study were unified with the drug IDs from the downloaded DrugBank to obtain a list of drug names.

Disease–drug correlation data were obtained from the CTD and DrugBank databases, and disease–drug relationship pairs predicted by bioinformatics methods were removed. Next, pathogenetic genes were obtained based on the integration of Online Mendelian Inheritance in Man (OMIM, <https://omim.org/>; released on 28 May 2021) [75], Disease Ontology (DO, <https://disease-ontology.org/>; released on 11 September 2018) [76], and CTD databases from previous studies, where data described as pathogenic genes in the above databases were integrated into the collection of pathogenic genes we used. On the other hand, integrations were obtained for the pathways associated with these diseases in the CTD and KEGG databases. The disease names were aligned with the disease DOIDs in the DO database.

The data types and data sizes in the dataset are given in Supplementary Table S1.

4.2. Strategies for Evaluating the Degree of Disease–Drug Correlation

An important step in the DDCM approach is the computation of drug–disease correlations, which was accomplished by separately computing the similarity between each drug and disease.

4.2.1. Construction of Hybrid Matrix for Input SVR

Here, a hybrid matrix containing both disease and drug similarity information was obtained by using the adjusted cosine similarity (ACS) method to calculate drug similarity and disease similarity. The hybrid matrix by the operation of matrix splicing with the obtained drug similarity and disease-similarity information was gained. The support-vector regression algorithm (SVR algorithm) was then used to measure the degree of correlation between a disease and drugs.

1. Construction of Disease-Similarity Matrix and Drug Similarity Matrix

To construct the hybrid matrix for input to the SVR algorithm, the drug similarity and disease similarity were calculated separately as the basis. The correlations between diseases and drugs were calculated separately using the ACS.

For each disease, a feature list could be constructed whose length was the number of all genes. Each position in the list corresponds to a gene, and the position corresponding to the pathogenic gene for the disease was marked as 1, and the rest of the positions were marked as 0. The ACS of the gene signature lists corresponding to the two diseases was regarded as the similarity of the pathogenic genes of the two diseases. Similarly, the similarity of the disease-pathway data was calculated using the ACS. The mean value of similarity at both the pathway and gene levels was used as the final similarity for two diseases (File S4 3_ACS_can_Dism_gene, 4_ACS_can_Dism_pathway, File S5 3_ACS_vas_Dism_gene, 4_ACS_vas_Dism_pathway).

$$\begin{aligned} Di_g &= \{Di_{g1}, Di_{g2}, Di_{g3}, \dots, Di_{ga}, \dots, Di_{gk}\}, \forall a \in (1, k), \forall Di_{ga} \in \{0, 1\} \\ Di_p &= \{Di_{p1}, Di_{p2}, Di_{p3}, \dots, Di_{pb}, \dots, Di_{ph}\}, \forall b \in (1, h), \forall Di_{pb} \in \{0, 1\} \end{aligned} \quad (1)$$

where Di_g and Di_p represent the vectors of disease genes and disease pathways, respectively, and k and h represent the lengths of the two vectors, respectively.

To avert the similarity measures obtained based on different distance scale functions may produce different data distribution patterns, the next step was considered to be to jointly construct a disease similarity and drug similarity based on two levels based on disease similarity and drug similarity. To ensure that the data from the similarity of the two levels in the hybrid matrix can be compared together, the ACS was again chosen to measure the similarity of the drugs. Similarly, the calculation method of the disease-similarity matrix in the previous section was chosen to calculate the similarity of the two levels of data of the drug-target genes and drug-related pathways, respectively. The mean value of similarity at both the pathway and gene levels was used as the final similarity for two drugs (File S4 5_ACS_can_Drsm_gene, 6_ACS_can_Drsm_pathway, File S5 5_ACS_vas_Drsm_gene, 6_ACS_vas_Drsm_pathway).

$$\begin{aligned} Dr_g &= \{Dr_{g1}, Dr_{g2}, Dr_{g3}, \dots, Dr_{gc}, \dots, Dr_{gc}\}, \forall c \in (1, r), \forall Dr_{gc} \in \{0, 1\} \\ Dr_p &= \{Dr_{p1}, Dr_{p2}, Dr_{p3}, \dots, Dr_{pd}, \dots, Dr_{pt}\}, \forall d \in (1, t), \forall Dr_{pd} \in \{0, 1\} \end{aligned} \quad (2)$$

where Dr_g and Dr_p represent the vector of the drug-target gene and drug pathway, respectively, and r and t represent the length of the two vectors, respectively.

The ACS measure was an improved form of vector-based similarity. ACS can handle the relationship that has not been found yet. The influence of the relationship that has not

been found on the results can be minimized by operating on the mean value. And overall, the ACS scales well. The ACS formula is as follows:

$$R = \frac{\sum_{i=1}^N (A_i - \bar{\alpha})(B_i - \bar{\beta})}{\sqrt{\sum_{i=1}^N (A_i - \bar{\alpha})^2} \sqrt{\sum_{i=1}^N (B_i - \bar{\beta})^2}} \quad (3)$$

where R is the ACS score between two diseases/drugs, A_i is the number of corresponding positions of the gene/pathway vector of one disease/drug, $\bar{\alpha}$ is the mean value of all genes/pathways vectors of the gene/pathway list for one drug/disease, B_i is the number of corresponding positions of the gene/pathway vector of the other disease/drug, $\bar{\beta}$ is the mean value of all genes/pathways vectors of the gene/pathway list for the other disease/drug, and N is the total length of the gene/pathway vector of the disease/drug.

2. Construction of Hybrid Matrix

On the basis of the above-mentioned homogeneity information, in this paper, the drug-similarity matrix and the disease-similarity matrix were innovatively spliced into a heterogeneous hybrid matrix in a specific way. The rows of the matrix represent disease–drug relationship pairs, which were known drugs with their indications. Columns represent drug and disease, respectively. In the column representing the drug, each position in the matrix indicates the similarity of the drug in the disease–drug relationship pair represented by that row to the drug represented by that column. In the column representing the disease, each position in the matrix indicates the similarity of the disease in the disease–drug relationship pair to the disease represented by this column. Finally, a multi-level disease–drug correlation hybrid matrix was formed with the number of disease–drug relationship pairs as the number of rows and the sum of the number of drugs and diseases as the number of columns.

4.2.2. Acquisition of Drug–Disease Correlation Score

The use of support-vector regression (SVR) is the same principle as SVM for classification, with a few small differences. Supervised machine-learning models with associated learning algorithms that can analyze data for classification and regression analysis are called support-vector regression. Here, x is the drug–drug correlation score and disease–disease correlation score in the correlation hybrid matrix mentioned in the matrix above. $y(y \in R)$ is the disease–drug correlation scores of a drug for each disease, and l is the sample size. Then the linear support-vector regression algorithm can be expressed in the following form:

$$y = \langle \omega, x \rangle + \varphi, \omega, x \in R^e, \varphi \in R \quad (4)$$

Based on the known disease–drug correlation relationships we obtained, the ω and φ of the parties in the regression equation were acquired by the learning fit of the SVR algorithm to the scores of each row in the hybrid matrix. From this, we know that, in the prediction, for the unknown pairs of relationships, the following formula is available:

$$f(x) = \langle \omega, x \rangle + \varphi, \omega, x \in R^e, \varphi \in R \quad (5)$$

where $f(x)$ is the objective function that needs to be from the train set, $\langle \cdot, \cdot \rangle$ is the dot product operation in the R^e space, and ω is the weight. The problem of solving $f(x)$ through training can be transformed into a solution; then, the linear SVR formula can be transformed into:

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 \\ s.t. |\langle \omega, x \rangle + \varphi - y_i| \leq \varepsilon, i = 1, 2, \dots, l \end{cases} \quad (6)$$

That is, the difference between the predicted value ($\langle \omega, x \rangle + \varphi_i$) and the actual value y_i is smaller than the constant ε (effect of constant change).

In the actual prediction process, it is often difficult to directly determine the appropriate ε to ensure that most of the data can be within the interval, and SVR expects all training data

to be within the interval. So, the slack variable ξ is added, so that the interval requirement of the function changes, that is, allowing some samples to be outside the interval. After introducing the slack variables, all sample data meet the conditions:

$$|y_i - (\omega x_i + \varphi)| \leq \varepsilon + \xi, \forall i \quad (7)$$

This is the restriction after reflecting into a slack variable, so it is also called soft-interval SVR. For any sample x_i , if it is inside or on the edge of the isolation zone, ξ is 0. Above the isolation zone, $\xi > 0$, $\xi^* = 0$, and below the isolation zone, then $\xi^* > 0$, $\xi = 0$. After introducing the slack variable, the original formula can be written as:

$$\begin{aligned} \min_{\omega, b, \xi_i, \hat{\xi}_i} & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{\delta} (\xi_i, \hat{\xi}_i) \\ \text{s.t.} & \\ & f(x_i) - y_i \leq \varepsilon + \xi_i, \\ & y_i - f(x_i) \leq \varepsilon + \hat{\xi}_i, \\ & \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, \delta. \end{aligned} \quad (8)$$

Here $\varepsilon = 0.1$, as the loss function, is an empirical value that is considered to define the minimum distance to the sample point furthest from the hyperplane.

The hybrid matrix is input into the support-vector regression machine, while the corresponding labels were given, 1 for known disease–drug relationship pairs and 0 for non-known disease–drug relationship pairs. After inputting the matrix, the SVR learns and fits a regression line, so that pairs of drug and disease relationships that satisfy the constraints of the SVR machine get predicted results and their relative degree of correlation.

In summary, after setting the parameters, the corresponding regression equations were finally obtained by fitting each row of the constructed hybrid matrix. For these diseases, the corresponding prediction models can be obtained by training on known data and adjusting the parameters. The prediction model scores the relationship between the disease and some of the drugs to be predicted separately. This score is the so-called disease–drug correlation score, which is also known as $f(x_i)$ in the equation.

4.3. Screening of Potential Therapeutic Drugs

For a disease, a list of drugs was predicted by a support-vector regression algorithm after inputting the hybrid matrix into the SVR. Each drug has a correlation score with the corresponding disease. The drugs were filtered in a stepwise manner to obtain potential therapeutic drugs for the disease. The correlation scores between the drugs and a disease were calculated through random perturbation. Drug candidates for the disease were screened based on the rank of disease–drug correlation scores before and after randomized perturbation experiments, respectively. The stability score was used to measure the stability of the drug with respect to its correlation with the disease. Potential therapeutic drugs were obtained by setting up potential therapeutic criteria to screen for stable drug candidates (Supplementary Figure S8).

4.3.1. Drug Candidates

For a disease, the disease–drug correlation scores were calculated by SVR, and the score distribution ranged from negative to positive, with a greater positive direction indicating a stronger degree of correlation with the drug for that disease. The distribution of disease–drug correlation scores was similar for each disease.

For different diseases, the distributions of disease–drug correlation scores for the drugs were different. So, in the first step, we selected the drugs that ranked in the top 50 (de-redundant: the highest correlation score for a particular drug corresponding to a disease was taken) correlation scores as drug candidates (*DrC*) for a disease. Drugs ranked after the 50th were subsequently discarded.

4.3.2. Stable Drug Candidates

Considering the differences in the composition of the drugs in the training set could affect the screening of drug candidates. For a given disease, correlation scores and the obtained top 50 candidate drugs changed accordingly. Therefore, a random perturbation experiment was performed. Setting stability criteria to screen drugs. Drug candidates with high stability were identified as stable drug candidates.

Three types of random perturbation processes were carried out. Training sets for multiple diseases were constructed by randomly removing data for 1, 2, and 5 different diseases. We considered the top 50 drugs in the training set of these randomization operations as drug candidates for different randomization conditions.

The frequency of the top 50 drugs was counted, and finally, the stability scores of these drugs that emerged during randomization were used to assess the stability of the drugs and obtain stable drug candidates.

For each disease, the given stability quantitative index was the stability score for a candidate drug:

$$S_j = 1 - \sqrt{\frac{\sum_{i=1}^M (s_i - \bar{s})^2}{M - 1}} \quad (9)$$

where S_j is the stability score for a candidate drug, j indicates the type of randomization ($j = 1$ or 2 or 5), M is the number of times of randomization for a certain random type. s denotes the frequency of occurrence of the drug in the top 50 drugs of the list of disease–drug correlations obtained in a particular randomization of a certain randomization type. \bar{s} denotes the average of the frequency of occurrence of the top 50 drugs in the list of disease–drug correlations obtained for the drug in all randomization times obtained for a certain randomization type. The larger the S_j , the more stable the frequency of the drug in the predicted results in this random perturbation.

The stability criteria were specified, and the intersection of drugs that pass the stability interval under several random perturbations with the drug candidates before they undergo perturbations is called the set of stable drug candidates Dr_{S_j} , whose set is denoted by $SDrC$.

$$Dr_{S_j} = \{Top95\%S_j\} \\ SDrC = Dr_{S_1} \cap Dr_{S_2} \cap Dr_{S_5} \cap DrC \quad (10)$$

4.3.3. Potential Therapeutic Drugs Identification

Considering that the stability scores of stable drug candidates fluctuate to some extent, further, the potential therapeutic criteria were set to screen the final potential therapeutic drugs.

For each disease, the stability scores are in different types of random processes. It can be derived from this:

$$S_{SDrC_k} = 1 - \sqrt{\frac{\sum_{i=1}^{M'} (s_i - \bar{s})^2}{M' - 1}} \quad (11)$$

where S_{SDrC_k} is the total stability score of a drug in the set of stable drug candidates, and $M' = 3M + 1$ and M' denotes the sum of all random counts under different random types and the number of predictions before perturbation. k denotes a drug in the set of stable drug candidates. The larger the S_{SDrC_k} , the more stable the drug appears in various conditions, and the more likely it is to be a potential therapeutic drug for the disease.

Similarly, after calculating the total stability score of a drug, the potential therapeutic interval is defined. Stable drug candidates whose total stability score of S_{SDrC_k} in the top 95 percent were considered potential therapeutic drugs.

A collection of drugs that meet the potential therapeutic criteria is called the potential therapeutic drug set ($PtDr$).

$$PtDr = \{Top95\%S_{SDrC_k}\} \quad (12)$$

5. Conclusions

In this study, we presented a novel computational strategy for drug repositioning to identify potential therapeutic drugs for a disease. The approach was successfully applied to the prediction of potential therapeutic drugs for diseases in the MESH branch. And some potential therapeutic drugs have been well-validated from different levels. Compared with traditional machine-learning methods, the DDCM strategy was able to demonstrate drug–disease correlations with continuous scores, predicting potential therapeutic drugs for the disease on a larger scale. However, by collecting large amounts of data to conduct experiments, we found that the operational efficiency of the DDCM may be affected. In the future, we will consider more aspects of drug or disease similarity to improve the confidence of our algorithm. Moreover, we can also use the drug as the subject to predict which different diseases the drug can treat, thus expanding the scope of application of our strategy. In short, DDCM provides a new perspective for the study of drug repositioning, and it can be regarded as a useful guide for drug repositioning and drug discovery.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25105267/s1>. References [77–111] are cited in the supplementary materials.

Author Contributions: Conceptualization, H.Z., L.C. and W.L.; methodology, M.X.; software, M.X.; validation, J.H., Y.W., J.L. and W.H.; formal analysis, M.X.; investigation, M.X. and J.H.; resources, M.X. and J.H.; data curation, M.X. and W.L.; writing—original draft preparation, M.X.; writing—review and editing, H.Z., L.C. and W.L.; visualization, M.X. and W.L.; supervision, H.Z. and L.C.; project administration, H.Z. and L.C.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (32170674) and the Natural Science Foundation of Heilongjiang Province, China (LH2020C05).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Williams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: Progress, challenges and recommendations. *Nat. Rev. Drug Discov.* **2019**, *18*, 41–58. [[CrossRef](#)] [[PubMed](#)]
2. Monticelli, M.; Liguori, L.; Allocca, M.; Bosso, A.; Andreotti, G.; Lukas, J.; Monti, M.C.; Morretta, E.; Cubellis, M.V.; Hay Mele, B. Drug Repositioning for Fabry Disease: Acetylsalicylic Acid Potentiates the Stabilization of Lysosomal Alpha-Galactosidase by Pharmacological Chaperones. *Int. J. Mol. Sci.* **2022**, *23*, 5104. [[CrossRef](#)] [[PubMed](#)]
3. Ochoa-Maganda, V.Y.; Rangel-Castañeda, I.A.; Suárez-Rico, D.O.; Cortés-Zárate, R.; Hernández-Hernández, J.M.; Pérez-Rangel, A.; Chiquete-Félix, N.; León-Ávila, G.; González-Pozos, S.; Gaona-Bernal, J.; et al. Antigiardial Activity of Acetylsalicylic Acid Is Associated with Overexpression of HSP70 and Membrane Transporters. *Pharmaceuticals* **2020**, *13*, 440. [[CrossRef](#)] [[PubMed](#)]
4. Amare, G.G.; Meharie, B.G.; Belayneh, Y.M. A drug repositioning success: The repositioned therapeutic applications and mechanisms of action of thalidomide. *J. Oncol. Pharm. Pract.* **2021**, *27*, 673–678. [[CrossRef](#)] [[PubMed](#)]
5. Ohmoto, A.; Fuji, S. Current status of drug repositioning in hematology. *Expert Rev. Hematol.* **2021**, *14*, 1005–1011. [[CrossRef](#)]
6. Cruz-Burgos, M.; Losada-Garcia, A.; Cruz-Hernández, C.D.; Cortés-Ramírez, S.A.; Camacho-Arroyo, I.; Gonzalez-Covarrubias, V.; Morales-Pacheco, M.; Trujillo-Bornios, S.I.; Rodríguez-Dorantes, M. New Approaches in Oncology for Repositioning Drugs: The Case of PDE5 Inhibitor Sildenafil. *Front. Oncol.* **2021**, *11*, 627229. [[CrossRef](#)] [[PubMed](#)]
7. Keats, T.; Rosengren, R.J.; Ashton, J.C. The Rationale for Repurposing Sildenafil for Lung Cancer Treatment. *Anticancer Agents Med. Chem.* **2018**, *18*, 367–374. [[CrossRef](#)] [[PubMed](#)]
8. Zhang, Y.; Zhao, S.; Fu, Y.; Yan, L.; Feng, Y.; Chen, Y.; Wu, Y.; Deng, Y.; Zhang, G.; Chen, Z.; et al. Computational repositioning of dimethyl fumarate for treating alcoholic liver disease. *Cell Death Dis.* **2020**, *11*, 641. [[CrossRef](#)]
9. Jourdan, J.P.; Bureau, R.; Rochais, C.; Dallemagne, P. Drug repositioning: A brief overview. *J. Pharm. Pharmacol.* **2020**, *72*, 1145–1151. [[CrossRef](#)]

10. Lastres-Becker, I.; García-Yagüe, A.J.; Scannevin, R.H.; Casarejos, M.J.; Kügler, S.; Rábano, A.; Cuadrado, A. Repurposing the NRF2 Activator Dimethyl Fumarate as Therapy Against Synucleinopathy in Parkinson's Disease. *Antioxid. Redox. Signal.* **2016**, *25*, 61–77. [\[CrossRef\]](#)
11. Park, K. A review of computational drug repurposing. *Transl. Clin. Pharmacol.* **2019**, *27*, 59–63. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Badkas, A.; De Landtsheer, S.; Sauter, T. Topological network measures for drug repositioning. *Brief. Bioinform.* **2021**, *22*, bbaa357. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Andronis, C.; Sharma, A.; Virvilis, V.; Deftereios, S.; Persidis, A. Literature mining, ontologies and information visualization for drug repurposing. *Brief. Bioinform.* **2011**, *12*, 357–368. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Tari, L.B.; Patel, J.H. Systematic drug repurposing through text mining. In *Biomedical Literature Mining*; Methods in Molecular Biology Series; Humana Press: New York, NY, USA, 2014; Volume 1159, pp. 253–267. [\[CrossRef\]](#)
15. Yang, H.T.; Ju, J.H.; Wong, Y.T.; Shmulevich, I.; Chiang, J.H. Literature-based discovery of new candidates for drug repurposing. *Brief. Bioinform.* **2017**, *18*, 488–497. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Zhu, Q.; Tao, C.; Shen, F.; Chute, C.G. Exploring the pharmacogenomics knowledge base (PharmGKB) for repositioning breast cancer drugs by leveraging Web ontology language (OWL) and cheminformatics approaches. In *Biocomputing 2024*; World Scientific: Singapore, 2014; pp. 172–182.
17. Jin, S.; Niu, Z.; Jiang, C.; Huang, W.; Xia, F.; Jin, X.; Liu, X.; Zeng, X. HeTDR: Drug repositioning based on heterogeneous networks and text mining. *Patterns* **2021**, *2*, 100307. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Xue, H.; Li, J.; Xie, H.; Wang, Y. Review of Drug Repositioning Approaches and Resources. *Int. J. Biol. Sci.* **2018**, *14*, 1232–1244. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Wu, Z.; Wang, Y.; Chen, L. Network-based drug repositioning. *Mol. Biosyst.* **2013**, *9*, 1268–1281. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Wang, W.; Yang, S.; Zhang, X.; Li, J. Drug repositioning by integrating target information through a heterogeneous network model. *Bioinformatics* **2014**, *30*, 2923–2930. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Cheng, F.; Liu, C.; Jiang, J.; Lu, W.; Li, W.; Liu, G.; Zhou, W.; Huang, J.; Tang, Y. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput. Biol.* **2012**, *8*, e1002503. [\[CrossRef\]](#)
22. Wu, C.; Gudivada, R.C.; Aronow, B.J.; Jegga, A.G. Computational drug repositioning through heterogeneous network clustering. *BMC Syst. Biol.* **2013**, *7* (Suppl. 5), S6. [\[CrossRef\]](#)
23. Jarada, T.N.; Rokne, J.G.; Alhaji, R. A review of computational drug repositioning: Strategies, approaches, opportunities, challenges, and directions. *J. Cheminform.* **2020**, *12*, 46. [\[CrossRef\]](#)
24. Zong, N.; Kim, H.; Ngo, V.; Harismendy, O. Deep mining heterogeneous networks of biomedical linked data to predict novel drug-target associations. *Bioinformatics* **2017**, *33*, 2337–2344. [\[CrossRef\]](#)
25. Zeng, X.; Zhu, S.; Liu, X.; Zhou, Y.; Nussinov, R.; Cheng, F. DeepDr: A network-based deep learning approach to in silico drug repositioning. *Bioinformatics* **2019**, *35*, 5191–5198. [\[CrossRef\]](#)
26. Zhao, K.; So, H.C. Drug Repositioning for Schizophrenia and Depression/Anxiety Disorders: A Machine Learning Approach Leveraging Expression Data. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 1304–1315. [\[CrossRef\]](#)
27. Gottlieb, A.; Stein, G.Y.; Rupp, E.; Sharan, R. PREDICT: A method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* **2011**, *7*, 496. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Yang, M.; Luo, H.; Li, Y.; Wu, F.X.; Wang, J. Overlap matrix completion for predicting drug-associated indications. *PLoS Comput. Biol.* **2019**, *15*, e1007541. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Setoain, J.; Franch, M.; Martínez, M.; Tabas-Madrid, D.; Sorzano, C.O.; Bakker, A.; Gonzalez-Couto, E.; Elvira, J.; Pascual-Montano, A. NFFinder: An online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. *Nucleic Acids Res.* **2015**, *43*, W193–W199. [\[CrossRef\]](#)
30. Le, D.H.; Nguyen-Ngoc, D. Drug Repositioning by Integrating Known Disease-Gene and Drug-Target Associations in a Semi-supervised Learning Model. *Acta Biotheor.* **2018**, *66*, 315–331. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Zhao, B.W.; Hu, L.; You, Z.H.; Wang, L.; Su, X.R. HINGRL: Predicting drug-disease associations with graph representation learning on heterogeneous information networks. *Brief. Bioinform.* **2022**, *23*, bbab515. [\[CrossRef\]](#)
32. Rodriguez, S.; Hug, C.; Todorov, P.; Moret, N.; Boswell, S.A.; Evans, K.; Zhou, G.; Johnson, N.T.; Hyman, B.T.; Sorger, P.K.; et al. Machine learning identifies candidates for drug repurposing in Alzheimer's disease. *Nat. Commun.* **2021**, *12*, 1033. [\[CrossRef\]](#)
33. Rikova, K.; Guo, A.; Zeng, Q.; Possemato, A.; Yu, J.; Haack, H.; Nardone, J.; Lee, K.; Reeves, C.; Li, Y.; et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* **2007**, *131*, 1190–1203. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Chen, Y.; Mao, B.; Peng, X.; Zhou, Y.; Xia, K.; Guo, H.; Jiang, H.; Wang, G.; Zhuang, W. A Comparative Study of Genetic Profiles of Key Oncogenesis-Related Genes between Primary Lesions and Matched Lymph Nodes Metastasis in Lung Cancer. *J. Cancer* **2019**, *10*, 1642–1650. [\[CrossRef\]](#)
35. da Cunha Santos, G.; Shepherd, F.A.; Tsao, M.S. EGFR mutations and lung cancer. *Annu. Rev. Pathol.* **2011**, *6*, 49–69. [\[CrossRef\]](#)
36. Huang, J.Q.; Duan, L.X.; Liu, Q.Y.; Li, H.F.; Hu, A.P.; Song, J.W.; Lin, C.; Huang, B.; Yao, D.; Peng, B.; et al. Serine-arginine protein kinase 1 (SRPK1) promotes EGFR-TKI resistance by enhancing GSK3 β Ser9 autophosphorylation independent of its kinase activity in non-small-cell lung cancer. *Oncogene* **2023**, *42*, 1233–1246. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Chang, H.; Iizasa, T.; Shibuya, K.; Iyoda, A.; Suzuki, M.; Moriya, Y.; Liu, T.L.; Hiwasa, T.; Hiroshima, K.; Fujisawa, T. Increased expression of collagen XVIII and its prognostic value in nonsmall cell lung carcinoma. *Cancer* **2004**, *100*, 1665–1672. [\[CrossRef\]](#) [\[PubMed\]](#)

38. Zhang, P.; Ma, Y.; Lv, C.; Huang, M.; Li, M.; Dong, B.; Liu, X.; An, G.; Zhang, W.; Zhang, J.; et al. Upregulation of programmed cell death ligand 1 promotes resistance response in non-small-cell lung cancer patients treated with neo-adjuvant chemotherapy. *Cancer Sci.* **2016**, *107*, 1563–1571. [\[CrossRef\]](#)
39. Chatterjee, S.; Crozet, L.; Damotte, D.; Iribarren, K.; Schramm, C.; Alifano, M.; Lupo, A.; Cherfils-Vicini, J.; Goc, J.; Katsahian, S.; et al. TLR7 promotes tumor progression, chemotherapy resistance, and poor clinical outcomes in non-small cell lung cancer. *Cancer Res.* **2014**, *74*, 5008–5018. [\[CrossRef\]](#)
40. Ren, M.; Hong, M.; Liu, G.; Wang, H.; Patel, V.; Biddinger, P.; Silva, J.; Cowell, J.; Hao, Z. Novel FGFR inhibitor ponatinib suppresses the growth of non-small cell lung cancer cells overexpressing FGFR1. *Oncol. Rep.* **2013**, *29*, 2181–2190. [\[CrossRef\]](#)
41. Tchaicha, J.H.; Akbay, E.A.; Altabef, A.; Mikse, O.R.; Kikuchi, E.; Rhee, K.; Liao, R.G.; Bronson, R.T.; Sholl, L.M.; Meyerson, M.; et al. Kinase domain activation of FGFR2 yields high-grade lung adenocarcinoma sensitive to a Pan-FGFR inhibitor in a mouse model of NSCLC. *Cancer Res.* **2014**, *74*, 4676–4684. [\[CrossRef\]](#)
42. Geng, N.; Su, J.; Liu, Z.; Ding, C.; Xie, S.; Hu, W. The Influence of KDR Genetic Variation on the Efficacy and Safety of Patients with Advanced NSCLC Receiving First-Line Bevacizumab Plus Chemotherapy Regimen. *Technol. Cancer Res. Treat.* **2021**, *20*, 15330338211019433. [\[CrossRef\]](#)
43. Testoni, E.; Stephenson, N.L.; Torres-Ayuso, P.; Marusiak, A.A.; Trotter, E.W.; Hudson, A.; Hodgkinson, C.L.; Morrow, C.J.; Dive, C.; Brognard, J. Somatically mutated ABL1 is an actionable and essential NSCLC survival gene. *EMBO Mol. Med.* **2016**, *8*, 105–116. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Garmendia, I.; Pajares, M.J.; Hermida-Prado, F.; Ajona, D.; Bértolo, C.; Sainz, C.; Lavín, A.; Remírez, A.B.; Valencia, K.; Moreno, H.; et al. YES1 Drives Lung Cancer Growth and Progression and Predicts Sensitivity to Dasatinib. *Am. J. Respir. Crit. Care Med.* **2019**, *200*, 888–899. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Gold, K.A.; Lee, J.J.; Harun, N.; Tang, X.; Price, J.; Kawedia, J.D.; Tran, H.T.; Erasmus, J.J.; Blumenschein, G.R.; William, W.N.; et al. A phase I/II study combining erlotinib and dasatinib for non-small cell lung cancer. *Oncologist* **2014**, *19*, 1040–1041. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Haura, E.B.; Tanvetyanon, T.; Chiappori, A.; Williams, C.; Simon, G.; Antonia, S.; Gray, J.; Litschauer, S.; Tetteh, L.; Neuger, A.; et al. Phase I/II study of the Src inhibitor dasatinib in combination with erlotinib in advanced non-small-cell lung cancer. *J. Clin. Oncol.* **2010**, *28*, 1387–1394. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Kilvaer, T.K.; Rakaee, M.; Hellevik, T.; Vik, J.; Petris, L.; Donnem, T.; Strell, C.; Ostman, A.; Busund, L.R.; Martinez-Zubiaurre, I. Differential prognostic impact of platelet-derived growth factor receptor expression in NSCLC. *Sci. Rep.* **2019**, *9*, 10163. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Karthik, S.; Sankar, R.; Varunkumar, K.; Ravikumar, V. Romidepsin induces cell cycle arrest, apoptosis, histone hyperacetylation and reduces matrix metalloproteinases 2 and 9 expression in bortezomib sensitized non-small cell lung cancer cells. *Biomed. Pharmacother.* **2014**, *68*, 327–334. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Pei, K.; Zhu, J.J.; Wang, C.E.; Xie, Q.L.; Guo, J.Y. MicroRNA-185-5p modulates chemosensitivity of human non-small cell lung cancer to cisplatin via targeting ABCC1. *Eur. Rev. Med. Pharmacol. Sci.* **2016**, *20*, 4697–4704. [\[PubMed\]](#)
50. Wang, X.; Wang, H.; Jiang, H.; Qiao, L.; Guo, C. Circular RNA circ_0076305 Promotes Cisplatin (DDP) Resistance of Non-Small Cell Lung Cancer Cells by Regulating ABCC1 Through miR-186-5p. *Cancer. Biother. Radiopharm.* **2023**, *38*, 293–304. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Liewer, S.; Huddleston, A. Alisertib: A review of pharmacokinetics, efficacy and toxicity in patients with hematologic malignancies and solid tumors. *Expert Opin. Investig. Drugs* **2018**, *27*, 105–112. [\[CrossRef\]](#)
52. Wang, C.Y.; Lee, M.H.; Kao, Y.R.; Hsiao, S.H.; Hong, S.Y.; Wu, C.W. Alisertib inhibits migration and invasion of EGFR-TKI resistant cells by partially reversing the epithelial-mesenchymal transition. *Biochim. Biophys. Acta Mol. Cell Res.* **2021**, *1868*, 119016. [\[CrossRef\]](#)
53. Jia, M.; Jia, X.; Zhang, D.; Liu, W.; Yi, S.; Li, Z.; Cong, B.; Ma, C.; Li, S.; Zhang, J. CD2⁺ T-helper 17-like cells differentiated from a CD133⁺ subpopulation of non-small cell lung carcinoma cells promote the growth of lung carcinoma. *Ann. Transl. Med.* **2021**, *9*, 687. [\[CrossRef\]](#)
54. Hwang, K.E.; Kim, H.J.; Song, I.S.; Park, C.; Jung, J.W.; Park, D.S.; Oh, S.H.; Kim, Y.S.; Kim, H.R. Salinomycin suppresses TGF- β 1-induced EMT by down-regulating MMP-2 and MMP-9 via the AMPK/SIRT1 pathway in non-small cell lung cancer. *Int. J. Med. Sci.* **2021**, *18*, 715–726. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Yu, X.; Zhang, Y.; Wu, B.; Kurie, J.M.; Pertsemlidis, A. The miR-195 Axis Regulates Chemoresistance through TUBB and Lung Cancer Progression through BIRC5. *Mol. Ther. Oncolytics* **2019**, *14*, 288–298. [\[CrossRef\]](#)
56. Grandel, U.; Heygster, D.; Sibelius, U.; Fink, L.; Sigel, S.; Seeger, W.; Grimminger, F.; Hattar, K. Amplification of lipopolysaccharide-induced cytokine synthesis in non-small cell lung cancer/neutrophil cocultures. *Mol. Cancer Res.* **2009**, *7*, 1729–1735. [\[CrossRef\]](#)
57. Flynn, R. Survival analysis. *J. Clin. Nurs.* **2012**, *21*, 2789–2797. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Behan, F.M.; Iorio, F.; Picco, G.; Gonçalves, E.; Beaver, C.M.; Migliardi, G.; Santos, R.; Rao, Y.; Sassi, F.; Pinnelli, M.; et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* **2019**, *568*, 511–516. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Dempster, J.M.; Boyle, I.; Vazquez, F.; Root, D.E.; Boehm, J.S.; Hahn, W.C.; Tsherniak, A.; McFarland, J.M. Chronos: A cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol.* **2021**, *22*, 343. [\[CrossRef\]](#)

60. Guo, Q.; Jiang, E. Recent Advances in the Application of Podophyllotoxin Derivatives to Fight Against Multidrug-Resistant Cancer Cells. *Curr. Top. Med. Chem.* **2021**, *21*, 1712–1724. [[CrossRef](#)] [[PubMed](#)]
61. Vermes, A.; Guchelaar, H.J.; Dankert, J. Flucytosine: A review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. *J. Antimicrob. Chemother.* **2000**, *46*, 171–179. [[CrossRef](#)]
62. Johnson, L.; Goping, I.S.; Rieger, A.; Mane, J.Y.; Huzil, T.; Banerjee, A.; Luduena, R.; Hassani, B.; Winter, P.; Tuszyński, J.A. Novel Colchicine Derivatives and their Anti-cancer Activity. *Curr. Top. Med. Chem.* **2017**, *17*, 2538–2558. [[CrossRef](#)]
63. Fakhraei, S.; Huang, B.; Raschid, L.; Getoor, L. Network-Based Drug-Target Interaction Prediction with Probabilistic Soft Logic. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2014**, *11*, 775–787. [[CrossRef](#)] [[PubMed](#)]
64. Hu, Y.; Zhou, M.; Shi, H.; Ju, H.; Jiang, Q.; Cheng, L. Measuring disease similarity and predicting disease-related ncRNAs by a novel method. *BMC Med. Genom.* **2017**, *10* (Suppl. 5), 71. [[CrossRef](#)]
65. Trosset, J.Y.; Cavé, C. In Silico Drug-Target Profiling. In *Target Identification and Validation in Drug Discovery*; Methods in Molecular Biology Series; Humana Press: New York, NY, USA, 2019; Volume 1953, pp. 89–103. [[CrossRef](#)]
66. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082. [[CrossRef](#)] [[PubMed](#)]
67. Srivenugopal, K.S.; Rawat, A.; Niture, S.K.; Paranjpe, A.; Velu, C.; Venugopal, S.N.; Madala, H.R.; Basak, D.; Punganuru, S.R. Posttranslational Regulation of O(6)-Methylguanine-DNA Methyltransferase (MGMT) and New Opportunities for Treatment of Brain Cancers. *Mini Rev. Med. Chem.* **2016**, *16*, 455–464. [[CrossRef](#)] [[PubMed](#)]
68. Neville, K.; Blaney, S.; Bernstein, M.; Thompson, P.; Adams, D.; Aleksic, A.; Berg, S. Pharmacokinetics of O(6)-benzylguanine in pediatric patients with central nervous system tumors: A pediatric oncology group study. *Clin. Cancer Res.* **2004**, *10*, 5072–5075. [[CrossRef](#)] [[PubMed](#)]
69. Keir, S.T.; Dolan, M.E.; Pegg, A.E.; Lawless, A.; Moschel, R.C.; Bigner, D.D.; Friedman, H.S. O6-benzylguanine-mediated enhancement of nitrosourea activity in Mer- central nervous system tumor xenografts—implications for clinical trials. *Cancer Chemother. Pharmacol.* **2000**, *45*, 437–440. [[CrossRef](#)] [[PubMed](#)]
70. Friedman, H.S.; Dolan, M.E.; Pegg, A.E.; Marcelli, S.; Keir, S.; Catino, J.J.; Bigner, D.D.; Schold, S.C., Jr. Activity of temozolomide in the treatment of central nervous system tumor xenografts. *Cancer Res.* **1995**, *55*, 2853–2857. [[PubMed](#)]
71. Xue, P.; Fu, J.; Zhou, Y. The Aryl Hydrocarbon Receptor and Tumor Immunity. *Front. Immunol.* **2018**, *9*, 286. [[CrossRef](#)] [[PubMed](#)]
72. Takanaga, H.; Yoshitake, T.; Yatabe, E.; Hara, S.; Kunimoto, M. Beta-naphthoflavone disturbs astrocytic differentiation of C6 glioma cells by inhibiting autocrine interleukin-6. *J. Neurochem.* **2004**, *90*, 750–757. [[CrossRef](#)]
73. Davis, A.P.; Wieggers, T.C.; Johnson, R.J.; Sciaky, D.; Wieggers, J.; Mattingly, C.J. Comparative Toxicogenomics Database (CTD): Update 2023. *Nucleic Acids Res.* **2023**, *51*, D1257–D1262. [[CrossRef](#)]
74. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)] [[PubMed](#)]
75. Amberger, J.S.; Bocchini, C.A.; Schiettecatte, F.; Scott, A.F.; Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **2015**, *43*, D789–D798. [[CrossRef](#)] [[PubMed](#)]
76. Schriml, L.M.; Mittraka, E.; Munro, J.; Tauber, B.; Schor, M.; Nickle, L.; Felix, V.; Jeng, L.; Bearer, C.; Lichtenstein, R.; et al. Human Disease Ontology 2018 update: Classification, content and workflow expansion. *Nucleic Acids Res.* **2019**, *47*, D955–D962. [[CrossRef](#)] [[PubMed](#)]
77. Shimamura, T.; Perera, S.A.; Foley, K.P.; Sang, J.; Rodig, S.J.; Inoue, T.; Chen, L.; Li, D.; Carretero, J.; Li, Y.C.; et al. Ganetespib (STA-9090), a nongeldanamycin HSP90 inhibitor, has potent antitumor activity in in vitro and in vivo models of non-small cell lung cancer. *Clin. Cancer Res.* **2012**, *18*, 4973–4985. [[CrossRef](#)] [[PubMed](#)]
78. Ardizzoni, A.; Pronzato, P.; Repetto, L.; De Palma, M.; Canobbio, L.; Gulisano, M.; Fusco, V.; Rosso, R.; Ganzina, F. Phase II trial of oral idarubicin in advanced non-small cell lung cancer (NSCLC). *Cancer Investig.* **1988**, *6*, 409–411. [[CrossRef](#)] [[PubMed](#)]
79. Klasa, R.J.; Gillum, A.M.; Klem, R.E.; Frankel, S.R. Oblimersen Bcl-2 antisense: Facilitating apoptosis in anticancer treatment. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 193–213. [[CrossRef](#)] [[PubMed](#)]
80. Frankel, S.R. Oblimersen sodium (G3139 Bcl-2 antisense oligonucleotide) therapy in Waldenström’s macroglobulinemia: A targeted approach to enhance apoptosis. *Semin. Oncol.* **2003**, *30*, 300–304. [[CrossRef](#)] [[PubMed](#)]
81. Kruser, T.J.; Traynor, A.M.; Wheeler, D.L. The use of single-agent dasatinib in molecularly unselected non-small-cell lung cancer patients. *Expert Opin. Investig. Drugs* **2011**, *20*, 305–307. [[CrossRef](#)] [[PubMed](#)]
82. Johnson, F.M.; Saigal, B.; Talpaz, M.; Donato, N.J. Dasatinib (BMS-354825) tyrosine kinase inhibitor suppresses invasion and induces cell cycle arrest and apoptosis of head and neck squamous cell carcinoma and non-small cell lung cancer cells. *Clin. Cancer Res.* **2005**, *11*, 6924–6932. [[CrossRef](#)]
83. Gerber, D.E.; Boothman, D.A.; Fattah, F.J.; Dong, Y.; Zhu, H.; Skelton, R.A.; Priddy, L.L.; Vo, P.; Dowell, J.E.; Sarode, V.; et al. Phase 1 study of romidepsin plus erlotinib in advanced non-small cell lung cancer. *Lung Cancer* **2015**, *90*, 534–541. [[CrossRef](#)]
84. Schrumpp, D.S.; Fischette, M.R.; Nguyen, D.M.; Zhao, M.; Li, X.; Kunst, T.F.; Hancox, A.; Hong, J.A.; Chen, G.A.; Kruchin, E.; et al. Clinical and molecular responses in lung cancer patients receiving Romidepsin. *Clin. Cancer Res.* **2008**, *14*, 188–198. [[CrossRef](#)] [[PubMed](#)]
85. Egyházi, S.; Edgren, M.R.; Hansson, J.; Kröckel, D.; Mannervik, B.; Ringborg, U. Role of O6-methylguanine-DNA methyltransferase, glutathione transferase M3-3 and glutathione in resistance to carmustine in a human non-small cell lung cancer cell line. *Eur. J. Cancer* **1997**, *33*, 447–452. [[CrossRef](#)] [[PubMed](#)]

86. Zhao, X.; Zhang, N.; Huang, Y.; Dou, X.; Peng, X.; Wang, W.; Zhang, Z.; Wang, R.; Qiu, Y.; Jin, M.; et al. Lansoprazole Alone or in Combination with Gefitinib Shows Antitumor Activity Against Non-small Cell Lung Cancer A549 Cells in vitro and in vivo. *Front. Cell Dev. Biol.* **2021**, *9*, 655559. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Arbour, K.C.; Machado, E.; Bott, M.J.; Ahn, L.; Tobi, Y.; Ni, A.A.; Yu, H.A.; Shannon, A.; Ladanyi, M.; Perron, V.; et al. Phase 1 Clinical Trial of Trametinib and Ponatinib in Patients with NSCLC Harboring KRAS Mutations. *JTO Clin. Res. Rep.* **2022**, *3*, 100256. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Ng, T.L.; Yu, H.; Smith, D.E.; Boyle, T.A.; York, E.R.; Leedy, S.; Gao, D.; Aisner, D.L.; Van Bokhoven, A.; Heasley, L.E.; et al. Preselection of Lung Cancer Cases Using FGFR1 mRNA and Gene Copy Number for Treatment with Ponatinib. *Clin. Lung Cancer* **2019**, *20*, e39–e51. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Melichar, B.; Adenis, A.; Lockhart, A.C.; Bennouna, J.; Dees, E.C.; Kayaleh, O.; Obermannova, R.; DeMichele, A.; Zatloukal, P.; Zhang, B.; et al. Safety and activity of alisertib, an investigational aurora kinase A inhibitor, in patients with breast cancer, small-cell lung cancer, non-small-cell lung cancer, head and neck squamous-cell carcinoma, and gastro-oesophageal adenocarcinoma: A five-arm phase 2 study. *Lancet Oncol.* **2015**, *16*, 395–405. [\[CrossRef\]](#)
90. Massabeau, C.; Khalifa, J.; Filleron, T.; Modesto, A.; Bigay-Gamé, L.; Plat, G.; Dierickx, L.; Aziza, R.; Rouquette, I.; Gomez-Roca, C.; et al. Continuous Infusion of Cilengitide Plus Chemoradiotherapy for Patients with Stage III Non-Small-cell Lung Cancer: A Phase I Study. *Clin. Lung Cancer* **2018**, *19*, e277–e285. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Li, Y.; Fang, M.; Xu, Z.; Li, X. Tetrathiomolybdate as an old drug in a new use: As a chemotherapeutic sensitizer for non-small cell lung cancer. *J. Inorg. Biochem.* **2022**, *233*, 111865. [\[CrossRef\]](#)
92. Vokes, E.E.; Lyss, A.P.; Herndon, J.E., 2nd; Cooper, B.; Perry, M.C.; Vinciguerra, V.; Mason-Coughlin, K.; Green, M.R. Intravenous 6-thioguanine or cisplatin, fluorouracil and leucovorin for advanced non-small cell lung cancer: A randomized phase II study of the cancer and leukemia group B. *Ann. Oncol.* **1992**, *3*, 727–732. [\[CrossRef\]](#)
93. Gao, H.; Song, Y.; Ma, J.; Zhai, J.; Zhang, Y.; Qu, X. Untargeted metabolomics analysis of omeprazole-enhanced chemosensitivity to cisplatin in mice with non-small cell lung cancer. *Chem. Biol. Interact.* **2022**, *360*, 109933. [\[CrossRef\]](#)
94. Klinge, C.M. Inhibition of non-small-cell lung cancer growth by combined fulvestrant and vandetanib. *Future Oncol.* **2012**, *8*, 529–533. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Cheng, Y.; Li, N.; Eapen, A.; Parajuli, R.; Mehta, R. Somatic BRCA2 Mutation-Positive Concurrent Accessory Male Breast Cancer (BC) and Non-Small Cell Lung Cancer (NSCLC): Excellent Efficacy of Palbociclib, Fulvestrant and Leuprolide in Platinum-Exposed and Endocrine-Refractory BC Associated with Cyclin D1 and FGFR1 Amplification and of Carboplatin, Paclitaxel and Radiation in NSCLC. *Case Rep. Oncol.* **2019**, *12*, 494–499. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Kochetova, O.V.; Avzaletdinova, D.S.; Korytina, G.F.; Morugova, T.V.; Mustafina, O.E. The association between eating behavior and polymorphisms in GRIN2B, GRIK3, GRIA1 and GRIN1 genes in people with type 2 diabetes mellitus. *Mol. Biol. Rep.* **2020**, *47*, 2035–2046. [\[CrossRef\]](#) [\[PubMed\]](#)
97. Gu, L.; Huang, J.; Li, J.; Huang, S.; Li, M.; Gong, L.; Li, T.; Su, L. Association of CALM1 rs3179089 Polymorphism with Ischemic Stroke in Chinese Han Population. *Neuromol. Med.* **2018**, *20*, 271–279. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Jing, H.; Liu, L.; Jia, Y.; Yao, H.; Ma, F. Overexpression of the long non-coding RNA Oprm1 alleviates apoptosis from cerebral ischemia-reperfusion injury through the Oprm1/miR-155/GATA3 axis. *Artif. Cells Nanomed. Biotechnol.* **2019**, *47*, 2431–2439. [\[CrossRef\]](#)
99. Esteban, G.; Bolea, I.; Sun, P.; Solé, M.; Samadi, A.; Marco-Contelles, J.; Unzeta, M. A therapeutic approach to cerebrovascular diseases based on indole substituted hydrazides and hydrazines able to interact with human vascular adhesion protein-1, monoamine oxidases (A and B), AChE and BuChE. *J. Neural Transm. (Vienna)* **2013**, *120*, 911–918. [\[CrossRef\]](#) [\[PubMed\]](#)
100. Zhao, J.; Forsberg, C.W.; Goldberg, J.; Smith, N.L.; Vaccarino, V. MAOA promoter methylation and susceptibility to carotid atherosclerosis: Role of familial factors in a monozygotic twin sample. *BMC Med. Genet.* **2012**, *13*, 100. [\[CrossRef\]](#) [\[PubMed\]](#)
101. Choi-Kwon, S.; Han, K.; Cho, K.H.; Choi, S.; Suh, M.; Nah, H.W.; Kim, J.S. Factors associated with post-stroke anger proneness in ischaemic stroke patients. *Eur. J. Neurol.* **2013**, *20*, 1305–1310. [\[CrossRef\]](#) [\[PubMed\]](#)
102. Dai, S.; Ma, W.; Qin, G.; Wang, L.; Wang, E. Influence of ADRB1, ADRB2, and COMT Genetic Polymorphisms on Postoperative Outcomes of Patients Undergoing Cardiac Valve Surgery. *Clin. Ther.* **2021**, *43*, 349–359.e2. [\[CrossRef\]](#)
103. Kim, B.R.; Kim, H.Y.; Chun, Y.I.; Yun, Y.M.; Kim, H.; Choi, D.H.; Lee, J. Association between genetic variation in the dopamine system and motor recovery after stroke. *Restor. Neurol. Neurosci.* **2016**, *34*, 925–934. [\[CrossRef\]](#)
104. Buga, A.M.; Ciobanu, O.; Bădescu, G.M.; Bogdan, C.; Weston, R.; Slevin, M.; Di Napoli, M.; Popa-Wagner, A. Up-regulation of serotonin receptor 2B mRNA and protein in the peri-infarcted area of aged rats and stroke patients. *Oncotarget* **2016**, *7*, 17415–17430. [\[CrossRef\]](#) [\[PubMed\]](#)
105. Martos, R.; Márton, I. Correlations between dental-oral infections and cardiovascular disease. *Fogorv. Sz.* **2008**, *101*, 101–105. [\[PubMed\]](#)
106. Nitsa, A.; Toutouza, M.; Machairas, N.; Mariolis, A.; Philippou, A.; Koutsilieris, M. Vitamin D in Cardiovascular Disease. *In Vivo* **2018**, *32*, 977–981. [\[CrossRef\]](#)
107. Kim, J.H.; Almuwaqqat, Z.; Hammadah, M.; Liu, C.; Ko, Y.A.; Lima, B.; Sullivan, S.; Alkhoder, A.; Abdulbaki, R.; Ward, L.; et al. Peripheral Vasoconstriction During Mental Stress and Adverse Cardiovascular Outcomes in Patients With Coronary Artery Disease. *Circ. Res.* **2019**, *125*, 874–883. [\[CrossRef\]](#) [\[PubMed\]](#)

108. Amaro, C.A.; González-Cortazar, M.; Herrera-Ruiz, M.; Román-Ramos, R.; Aguilar-Santamaría, L.; Tortoriello, J.; Jiménez-Ferrer, E. Hypoglycemic and hypotensive activity of a root extract of *Smilax aristolochiifolia*, standardized on N-trans-feruloyl-tyramine. *Molecules* **2014**, *19*, 11366–11384. [[CrossRef](#)] [[PubMed](#)]
109. Katrak, P.H.; Cole, A.M.; Poulos, C.J.; McCauley, J.C. Objective assessment of spasticity, strength, and function with early exhibition of dantrolene sodium after cerebrovascular accident: A randomized double-blind study. *Arch. Phys. Med. Rehabil.* **1992**, *73*, 4–9. [[PubMed](#)]
110. Carpiello, B.; Pinna, F.; Tuveri, R. Delusional infestation in a patient with renal failure, metabolic syndrome, and chronic cerebrovascular disease treated with aripiprazole: A case report. *Case Rep. Med.* **2011**, *2011*, 103652. [[CrossRef](#)]
111. Frishman, W.H.; Hotchkiss, H. Selective and nonselective dopamine receptor agonists: An innovative approach to cardiovascular disease treatment. *Am. Heart J.* **1996**, *132*, 861–870. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.