

# Treatment Repurposing using Literature-related Discovery

Ronald N. Kostoff\*

School of Public Policy, Georgia Institute of Technology, UNITED STATES.

## ABSTRACT

This article describes the Literature-Related Discovery technique and its application to Treatment Repurposing (which includes, but goes well beyond, Drug Repurposing). Illustrative results of potential repurposed treatments were shown from a study on preventing and reversing Alzheimer's disease. The detailed query used to generate these results is presented. The approach has the potential to identify voluminous amounts of candidate treatments for repurposing. Additionally, a broad review of the Drug Repurposing literature is provided. A Drug Repurposing database is retrieved and the structure and content are analyzed using Text Clustering and Factor Analysis. Two taxonomies of the Drug Repurposing literature are presented and specific major themes are shown.

**Keywords:** Treatment Repurposing, Drug Repurposing, Drug Repositioning, Literature-Related Discovery, Literature-Based Discovery, Knowledge Discovery.

## Correspondence

**Ronald N. Kostoff**  
School of Public Policy, Georgia Institute of Technology, 13500 Tallyrand Way,  
Gainesville, VA, USA.  
Email: rkostoff@gmail.com

Received: 19-10-2019

Revised: 30-10-2019

Accepted: 03-11-2019

DOI: 10.5530/jscires.8.2.25

## Literature-Related Discovery (LRD) for Treatment Repurposing (TR)

Literature-Related Discovery is a systematic approach to bridging unconnected disciplines based on text mining procedures.<sup>[1-2]</sup> In the LRD context, discovery is linking two or more literature concepts that have heretofore not been linked (i.e., disjoint), in order to produce novel, interesting, plausible and intelligible knowledge. In the medical arena, LRD can be used for identifying potentially new treatments for a disease (aka TR), identifying potentially new contributing factors for a disease, identifying new biomarkers for a disease, etc. By 'new', we mean previously non-existent in the core biomedical literature for the disease of interest.

## Structure of Treatment Repurposing Literature

TR is the application of an existing treatment for one or more diseases to diseases or symptoms of interest other than the disease(s) or symptom(s) for which the treatment was developed (and used) initially. Many comprehensive reviews of one component of TR, drug repurposing/repositioning, have been published recently.<sup>[3-12]</sup> As shown in these reviews, as well as many other more narrowly-focused documents, there are myriad possible categorizations for the TR literature.

For example, some TR papers are:

\*disease-focused;<sup>[13-15]</sup>

\*drug focused;<sup>[16-18]</sup>

\*biomarker-focused;<sup>[19-21]</sup>

\*prediction methodology-focused, etc.

The prediction methodology category includes, but is not limited to, the following sub-categories:

\*text-mining;<sup>[22-28]</sup>

\*machine learning;<sup>[5,7,29-31]</sup>

\*network-based;<sup>[32-35]</sup>

\*semantics;<sup>[36-38]</sup>

\*ligand-binding/ligand-protein docking/binding-site focused;<sup>[39-42]</sup>

\*protein targeting;<sup>[43-45]</sup>

\*transcriptional signature-focused.<sup>[46-48]</sup>

The LRD approach of TR is in the text mining sub-category of prediction methodology-focused approaches. LRD-TR identifies treatments from the biomedical literature that alter combinations of markers (mainly biomarkers) in directions required to reverse target diseases of interest (see Appendix 1 for details).

This review article presents a detailed description of LRD-based TR (which includes Drug Repurposing but goes well beyond drugs to include any type of treatment), in the context of a broad TR literature review. It contains illustrative examples of TR from a 2018 study on preventing and reversing Alzheimer's Disease (AD) (Kostoff RN, Porter AL, Buchtel HA. Prevention and reversal of Alzheimer's disease: treatment protocol. Georgia Institute of Technology. 2018. PDF.

## Copyright

© The Author(s). 2019 This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

<https://smartech.gatech.edu/handle/1853/59311>), as well as an algorithm for generating voluminous TR (see Appendix 1).

Two more objective perspectives on the structure of the TR literature are shown in Appendices 2 and 3. Appendix 2 contains a hierarchical taxonomy of the 2890 Medline record TR literature obtained with the CLUTO text clustering software (CLUTO. 2018. <http://glarosdtdcun.edu/gkhome/views/cluto>, University of Minnesota). This unique taxonomy presents the higher-level and most detailed categories that constitute the core TR biomedical literature. A display of the taxonomy linked to the titles of papers in the most detailed categories can be found in Appendix 2 of (Kostoff RN. Treatment Repurposing using Literature-Related Discovery. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/60507>).

Appendix 3 of the present article contains the results of a factor analysis of the TR literature. Only the factor themes are presented and discussed. Appendix 3 of (Kostoff RN. Treatment Repurposing using Literature-Related Discovery. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/60507>) contains a factor matrix showing the themes and key phrases that had the strongest influence on determining the themes. The phrases under each factor are linked to record titles associated with those phrases.

### Overview of LRD-TR Methodology

Treatments of disease or contributing factors to disease result in changes to myriad markers in the body, including, but not limited to:

\*general biomarkers (e.g., neuroinflammation, neurodegeneration, DNA damage, mitochondrial dysfunction, oxidative stress, neurotransmission dysfunction; olfactory dysfunction, glutamate uptake, glucose homeostasis, etc);

\*specific biomarkers (e.g., ATP, B12, B6, BACE-1, Bax, Bcl-2, BDNF, c-AMP, caspase, folate, GLP-1, GSK-3, etc);

\*symptoms (e.g., insomnia, ataxia, dysphagia, etc);

\*performance (e.g., memory, learning, cognition, etc);

\*behaviors (e.g., apathy, depression, anxiety, aggression, agitation, etc);

\*others.

Appendix 4 of (Kostoff RN. Treatment Repurposing using Literature-Related Discovery. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/60507>) contains extensive examples of myriad markers from the AD study and the directions in which they changed in association with the presence/imposition of AD contributing factors or the provision of AD treatments. These markers and their

directions of change from treatments or contributing factors form the basis of the TR discovery approach.

The TR discovery approach presented in this article consists of a two-stage process:

Stage 1: identify critical markers associated with a disease of interest and identify how the values of those markers change 1) when contributing factors to disease are operable and 2) when treatments are operable.

Stage 2: search the non-disease-of-interest literature for potential treatments that will change the markers of interest in the desired direction.

### Specific Methodology Adapted from AD Study

(Kostoff RN, Porter AL, Buchtel HA. Prevention and reversal of Alzheimer's disease: treatment protocol. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/59311>)

Stage 1: Identify critical markers and their directions of change associated with existing AD contributing factors

The first step in Stage 1 of the AD study (and in a recently-completed study (Kostoff RN. Prevention and Reversal of Peripheral Neuropathy/Peripheral Arterial Disease. Georgia Institute of Technology. 2019. PDF. <http://hdl.handle.net/1853/61865>)) was to 1) identify existing contributing factors (causes) to AD and 2) identify markers (mainly biomarkers) whose changes from the norm were associated with the AD contributing factors. Multiple approaches were used to identify these existing AD contributing factors and their associated markers, since no one approach was fully comprehensive.

### Visual Inspection

A Visual Inspection approach was used initially for the AD study. It started by generating a database of millions of abstract phrases parsed from ~100,000 records that constituted the total AD core Medline literature. Then, tens of thousands of the highest frequency phrases were inspected visually and those that appeared to be contributing factors to AD were selected. During this process and in the subsequent confirmatory process that validated the selection of AD contributing factors, many non-biomedical terms were identified that were closely associated with the existing AD contributing factors (shown in the next section). These non-biomedical terms could then be (and were) used as 'linking terms', to target lower frequency phrases (among the millions of abstract phrases) that had high probability of being/including existing AD contributing factors.

## Linking Term

A number of linking term approaches were used to target records or phrases with high probability of containing existing AD contributing factors. These included:

-MeSH Qualifiers associated strongly with contributing factors (e.g., adverse effects, toxicity, pathogenicity, poisoning);

-Relatively unambiguous MeSH Headings associated strongly with contributing factors (e.g., “Drug-Related Side Effects AND Adverse Reactions”; Abnormalities, Drug Induced; Air Pollutants, Occupational; Amphetamine Related Disorders; Carcinogens; Chemical Warfare Agents; Chemically-Induced Disorders, etc);

-Text terms associated strongly with contributing factors (e.g., -induced; caused by; induced by; -contaminated; exposure to; exposure(s) [at end of phrase]; exposed to; poisoning [at end]; -exposed [at end]; -related; -associated; -infected; abuse\*; toxicity).

These linking terms were especially valuable for accessing low-frequency existing AD contributing factors not accessible from visual inspection of the high-frequency phrases.

-Dot Product

A dot product approach was used to identify phrases that had high probability of being existing AD contributing factors. External lists of toxic substances generated by Federal government organizations, state regulatory agencies and other major organizations were aggregated. The final list of toxic substances was intersected with the full list of millions of abstract phrases in the core AD literature, to identify additional existing AD contributing factors.

The total number of validated existing AD contributing factors identified by the above approaches (from the premier biomedical literature) numbered about 400-600, depending on how the existing AD contributing factors were aggregated. In all the approaches to identifying existing AD contributing factors shown above, the initial existing AD contributing factors selected were confirmed and validated by detailed reading of the relevant abstracts.

During the confirmation and validation process, one or (usually) more record abstracts containing the candidate existing AD contributing factor term were read and other relevant data in the abstract were recorded. These data included biomarkers, symptoms and behaviors impacted by the existing AD contributing factor(s) and the directions in which these markers were moved (increased, decreased, etc). In some/many of these records, one or more existing AD treatment(s) were also identified, as well as the myriad markers associated with the existing AD treatments and the directions of change

these markers experienced as a result of the existing AD treatment(s). These treatment-related data were also recorded.

## Identify critical markers and their directions of change associated with existing AD treatments

The second step in Stage 1 of the AD study (and in the recently-completed study (Kostoff RN. Prevention and Reversal of Peripheral Neuropathy/Peripheral Arterial Disease. Georgia Institute of Technology. 2019. PDF. <http://hdl.handle.net/1853/61865>)) was to 1) identify existing AD treatments and 2) identify markers (mainly biomarkers) whose changes from the norm were associated with the existing AD treatments. Multiple approaches were used to identify these existing AD treatments and their associated markers, since no one approach was fully comprehensive.

## Visual Inspection

A Visual Inspection approach (part of the visual inspection approach described in the previous section) was used, which consisted of reading the thousands of high frequency abstract phrases in the core AD literature and selecting those that appeared to be treatments for AD. During this process and in the subsequent confirmatory process that validated the selection of existing AD treatments, non-biomedical terms were identified that were closely associated with the existing AD treatments (shown in the next section). These non-biomedical terms could then be (and were) used as ‘linking terms’, to target phrases (among the millions of abstract phrases) that had high probability of being/including existing AD treatments.

## Linking Term

Diverse linking term approaches were used to target records with high probability of containing existing AD treatments. These included:

-MeSH Qualifiers associated strongly with treatments (e.g., diet therapy, drug therapy, prevention and control, therapeutic use, therapy, etc);

-Relatively unambiguous MeSH Headings associated strongly with treatments (e.g., Treatment Outcome, Neuroprotective Agents, Nootropic Agents, Plant Extracts, Phytotherapy, Dietary Supplements, Drugs, Chinese Herbal, etc);

-Text terms associated strongly with treatments (treat\*, therap\*, prevent\*, protect\*, improv\*, reduc\*, attenuat\*, ameliorat\*, enhanc\*, revers\*, promot\*, alleviat\*, inhibit\*, remov\*, suppress\*, mitigat\*, restor\*, lower\*, preserv\*, regenerat\*, rescu\*, slow\*).

These linking terms were especially valuable for accessing existing low-frequency AD treatments not accessible from visual inspection of the high-frequency phrases. Some of these linking terms had higher efficiencies of identifying the treatment

consequences of interest than others. Terms like prevent\*, protect\*, improv\*, restor\*, alleviat\*, ameliorat\*, mitigat\*, etc, almost always gave the desired AD markers and the direction in which they changed as a result of treatment. Terms like decreas\* and increas\* (used initially, then abandoned), reduc\*, slow\*, etc, could go either way. The former group of terms had the 'sense' of *improvement*, while the latter group of terms reflected *change* (positive or negative) and may or may not have reflected improvement.

The total number of existing AD treatments identified by the above approaches (from the premier biomedical literature and validated) numbered about 600-700, depending on how the existing AD treatments were aggregated. In all the approaches to identifying existing AD treatments shown above, the initial existing AD treatments selected were confirmed and validated by detailed reading of the relevant abstracts.

The number of AD treatments we identified might seem unrealistically large at first glance, but these levels have occurred in all our disease reversal studies. These studies are based on the following systemic medical principle: *at the present time, removal of cause is a necessary, but not necessarily sufficient, condition for restorative treatment to be effective.* In the AD study, the treatments identified covered research over the past ~thirty years. The study did not exclude treatments that have 'failed' in human clinical trials, for the following reasons. Reading of thousands of abstracts on laboratory experiments and clinical trials of potential AD treatments has shown

- 1) *in vitro* experiments typically performed on neural cells tend to have reasonably positive outcomes, at least for those papers that surface in the peer-reviewed published literature;
- 2) *in vivo* experiments typically performed on rodents (but other small animals as well) tend to also have reasonably positive outcomes, albeit somewhat less than *in vitro* experiments;
- 3) When these potential treatments reach the human clinical trial stage, especially the later phases, the success rates plummet!

The explanation for this discrepancy given most often is the species difference. Humans are different from rodent's *et al.* and their physiological responses to stimuli are different as well. However, the toxic experiential and exposure background differences between humans who live in the sea of toxic exposures in the real world and animals who live in the very controlled environment of the laboratory are rarely, if ever, discussed.

There were many hundreds of potential causes for AD identified in the AD study (ranging from Lifestyle to Occupational/Environmental exposures). For a given individual, some causes have happened in the past and are no longer happening, but their damage trail remains. Other causes are ongoing, have caused damage and continue to cause damage.

Why would anyone expect a human being with such a toxic history to respond to a potential treatment the same way that a laboratory animal raised in a controlled environment would respond to that treatment? Furthermore, why would anyone expect a human being with such a toxic history to respond to a potential treatment the same way that another human being without such a toxic burden would respond to that treatment?

We cannot rule out failure to remove cause as a reason for the massive failure of myriad AD treatments in the clinical trials of the past three decades. That is why we retained even so-called 'failed' treatments in the present full-spectrum study of existing AD treatments. We don't know which treatments failed because 1) they were intrinsically ineffective or 2) their beneficial effects were overwhelmed by the strong negative effects of the ongoing causes remaining operable.

During the confirmation and validation process, one or (usually) more record abstracts containing the candidate existing AD treatment term were read and other relevant data in the abstract were recorded. These data included biomarkers, symptoms and behaviors impacted by the treatment(s) and the directions in which these markers were moved (increased, decreased, etc) associated with the treatment. In some/many of these records, one or more existing AD contributing factor(s) were also identified, as well as the myriad markers associated with the existing AD contributing factor(s) and the directions of change these markers experienced associated with the existing AD contributing factor(s). These contributing factor-related data were also recorded.

Additional markers could have been identified using the same approaches for identifying contributing factors and treatments, but that was not done in the AD study. It was done in the more recent study on peripheral neuropathy (PN)/Peripheral Arterial Disease (PAD) (Kostoff RN. Prevention and Reversal of Peripheral Neuropathy/Peripheral Arterial Disease. Georgia Institute of Technology. 2019. PDF. <http://hdl.handle.net/1853/61865>) and approximately four times as many markers were identified compared to the AD study.

## Stage 2

### Search the non-disease-of-interest literature for potential treatments that will change the markers of interest in the desired direction.

Text mining of the AD biomedical literature (especially records focused on treatments and contributing factors) identified the critical markers associated with AD and identified the directions in which these critical markers needed to change for potential AD alleviation. For example, critical general biomarkers for AD and their desired directions of change included 'reduce oxidative stress', 'alleviate mitochondrial dysfunction', 'prevent apoptosis', etc. Critical specific biomarkers for AD and their desired directions of change included 'reduce BACE1', 'increase Bcl-2', 'enhance ADAM10', etc.

From these markers and their desired directions of change for effective treatment of AD, a query was developed to 1) identify potential AD treatments from 2) treatments used in the non-AD literature (see Appendix 1 for query details). The non-AD biomedical literature was then searched for records including one or more of these AD markers that moved in desired directions as a result of treatments (e.g., reduced Abeta; increased Bcl-2; reduced tau hyperphosphorylation; restricted NFKappaB signaling; reduced inflammation; reduced oxidative stress; enhanced Nrf2, etc).

Searching for records that had a threshold of including *at least one* of these desired marker alterations produced a voluminous retrieval. To keep the records retrieved at a manageable level,

the requirement that a record in the non-AD literature must contain *at least two* AD markers (that moved in the appropriate direction in conjunction with a treatment) to be retrieved was imposed. Even then, the retrieval was voluminous, indicating the wealth of potential AD treatment repurposing possible from an expanded well-resourced study.

As a practical matter, combinations of the more fundamental and less AD-specific linking phrases were used for the treatment repurposing query. The general form of the query was 1) combinations of the markers and their desired directions of change followed by 2) negation of records that contained existing AD treatments.

As an example of the query format:

((increas\* OR enhanc\* OR restor\*) NEAR/3 “norepinephrine”) AND ((increas\* OR enhanc\* OR restor\*) NEAR/3 “Nrf2”)

NOT

((alzheimer\* OR dementia OR “mild cognitive impairment”) OR {existing AD treatments}).

While terms such as ‘reduce Abeta’ or ‘reduce tau phosphorylation’ may be efficient for extracting *existing* AD treatments from the AD literature, they are very inefficient, either in isolation or especially in combination, for AD treatment *repurposing* from the non-AD literature. It is difficult to imagine people doing research in reducing Abeta or reducing tau hyperphosphorylation (much less doing research in both) not emphasizing the AD/dementia applications in their publications.

Finally, there are no restrictions on the numbers of treatments that could be repurposed for any disease of interest. For example, assume that a patient has been diagnosed with a specific disease, characterized by three abnormal biomarker values. The query could be applied to identify/discover 1) one treatment that would bring all three of the biomarkers back to normal, or 2) one treatment that would bring two of the biomarkers back to normal and one treatment that would bring the third biomarker back to normal, or 3) three treatments, each of which would bring one of the biomarkers back to normal. Obviously, the repurposed treatments in 2) and 3) would have to be compatible, but the technique offers a wide variety of options.

### Illustrative Examples of Potential Repurposed Treatments for AD

Appendix 1 contains the details of the actual query used to identify potential repurposed treatments for AD. Since the treatment repurposing described in the AD study was a proof-of-principle demonstration of the latest incarnation of our LRDI approach, only a few examples were provided for il-

lustrative purposes. These examples were based on using *combinations of two biomarkers only* in the query and are contained in the following table:

**Table 1: Illustrative LRD-TR Results from AD Study.**

#### EXAMPLES OF POTENTIAL REPURPOSED AD TREATMENTS

- Fortunellin protects against high fructose-induced diabetic heart injury in mice by *suppressing inflammation and oxidative stress* via AMPK/Nrf-2 pathway regulation<sup>[49]</sup>
- Protective effects of sarains on H2O2-induced mitochondrial dysfunction and oxidative stress; *improving mitochondrial function and decreasing reactive oxygen species* levels; ability to block the mPTP and to *enhance the Nrf2 pathway*<sup>[50]</sup>
- Carboxyamidotriazole alleviates muscle atrophy in tumor-bearing mice by *inhibiting NF-kappaB* and *activating SIRT1*; CAI restricted the NF-kappaB signaling, *downregulated the level of TNF-alpha* in muscle and both *TNF-alpha and IL-6 levels* in serum, directly stimulated SIRT1 activity *in vitro* and *increased SIRT1 content* in muscle<sup>[51]</sup>
- Protective effects and mechanism of meretrix meretrix oligopeptides (MMO) against nonalcoholic fatty liver disease; MMO inhibited the activation of cell death-related pathways, based on *reduced p-JNK, Bax expression, tumor necrosis factor-alpha, caspase-9 and caspase-3* activity in the NAFLD model cells and *Bcl-2 expression was enhanced* in the NAFLD model cells<sup>[52]</sup>
- Extract from *Periostracum cicadae* *inhibits oxidative stress and inflammation* induced by Ultraviolet B Irradiation; *decreased reactive oxygen species (ROS) production*. The extract *attenuates the expression of interleukin-6 (IL-6), matrix metalloproteinase-2 (MMP-2) and MMP-9* in UVB-treated HaCaT cells. Also, *P. cicadae abrogated UVB-induced activation of NF-kappaB, p53 and activator protein-1 (AP-1)*; *accumulation and expression of NF-E2-related factor (Nrf2) were increased*<sup>[53]</sup>

Note that while the query was limited to combinations of two biomarkers only as selection criteria, the actual numbers of biomarkers in the retrieved records that moved in the desired directions for healing were typically greater (sometimes much greater) than two.

### Validation of TR Candidates

The final step involved in converting an existing treatment in the non-AD literature to a repurposed AD treatment is validation that the potential AD treatment has not been associated with AD application in the literature. The following block places the validation of our TR findings in context.

Treatment discovery/repurposing validation (or contributing factor or characteristic validation) is defined as the process of demonstrating that the candidate treatment has not been used or proposed for application to the disease of interest. This will be a function not only of the scope of the disease literature assumed, but which databases are included in the definition of the disease literature.

The AD study used the Pubmed version of Medline to retrieve the core AD literature and used both the Pubmed and Thomson Reuters versions to determine previous use. All the treatment discoveries/repurposing listed above were not present in these versions of the AD literature.

There could be other ways to define the scope of AD. There are also many other databases that could be searched for validation purposes, including other literature indexes, patent databases, books, magazine articles not indexed in Pubmed or Thomson Reuters, etc. Thus, the present validation should be viewed as limited, even though it is the method used and the databases used, by most (if not all) of the literature-based discovery community. For discovery patenting purposes, or other purposes, more extensive validation and larger numbers of databases, may be required.

In summary, each candidate potential AD treatment retrieved using the discovery query required validation before becoming a potential AD treatment. The candidate potential AD treatment was intersected with the core AD literature in late 2017 and was validated only after this intersection showed orthogonality.

For example, the candidate potential AD repurposed treatment “fortunellin” was retrieved because it satisfied the desired query general biomarker combination of *reducing inflammation and oxidative stress*. Fortunellin also had the additional specific biomarkers-based benefits of *reducing the pro-inflammatory cytokines and the expression of p-IkappaB kinase alpha, p-IkappaBalpha and p-nuclear factor-kappaB*, while significantly *enhancing superoxide dismutase, catalase, heme oxygenase-1 and p-AMP-activated protein kinase*. Fortunellin was intersected with the core AD biomedical literature retrieval terms (alzheimer\* OR dementia OR “mild cognitive impairment”) and no records were retrieved, demonstrating that fortunellin could not be found in the core AD literature. Fortunellin was therefore validated as a potential AD repurposed treatment (LRD Discovery).

## CONCLUSION

The TR literature reflects intense interest in the medical community for extracting maximum utility from drugs developed already. TR practitioners come from varied discipline communities and make use of myriad analytical predictive approaches, including text-mining, machine learning, network-based, semantics, ligand-binding/ligand-protein docking/binding-site focused, protein targeting and transcriptional signature-focused. The main diseases studied are cancer, neurodegenerative and infectious. TR may be equally important for the rare diseases, where the modest number of potential patients may not justify the expense to the manufacturers of separate new drug development. The main biomarker targets studied focus on oxidative stress and inflammatory metrics.

The LRD-TR approach has evolved from its initial structure in 2008<sup>[1]</sup> to the more advanced and targeted process described in the present article. It has the capability to generate volumi-

nous TR results for any disease or symptom of interest; the only limitations are study resources.

The same generic process can be applied to identifying contributing factors to a symptom or disease of interest that have not been existent previously in the core biomedical literature of that disease or symptom of interest but have been existent in the core biomedical literature of other diseases or symptoms. The same extrapolation process can be used for myriad markers as well.

For identification of both treatments and contributing factors, research needs to be done on which combinations of biomarkers will be most fruitful in retrieving the largest volume of high-quality TR and new contributing factor candidates. What are the characteristics of such combinations that will maximize marginal utility?

## REFERENCES

1. Kostoff RN. Literature-related Discovery (LRD): Introduction and background. *Technological Forecasting and Social Change*. 2008;75(2):165-85.
2. Kostoff RN, *et al*. Literature-Related Discovery. *Annual Review of Information Science and Technology*. 2009;43(1):241-85.
3. Paranjpe MD, *et al*. Insights into Computational Drug Repurposing for Neurodegenerative Disease. *Trends in Pharmacological Sciences*. 2019;40(8):565-76.
4. Parvathaneni V, *et al*. Drug repurposing: A promising tool to accelerate the drug discovery process. *Drug Discovery Today*. 2019.
5. Zhou L, *et al*. 2019;24(9):1714.
6. Sam E, Athri P. Web-based drug repurposing tools: A survey. *Briefings in Bioinformatics*. 2019;20(1):299-316.
7. Zhao K, So HC. Using Drug Expression Profiles and Machine Learning Approach for Drug Repurposing. *Methods in Molecular Biology*. 2019;1903:219-37.
8. Xue H, *et al*. Review of Drug Repositioning Approaches and Resources. *International Journal of Biological Sciences*. 2018;14(10):1232-44.
9. Karaman B, Sippl W. Computational Drug Repurposing: Current Trends. *Current Medicinal Chemistry*. 2018.
10. Yella JK, *et al*. Changing Trends in Computational Drug Repositioning. *Pharmaceuticals*. 2018;11(2):57.
11. March-Vila E, *et al*. On the Integration of *in silico* Drug Design Methods for Drug Repurposing. *Frontiers in Pharmacology*. 2017;8:298.
12. Sahu NU, Kharkar PS. Computational Drug Repositioning: A Lateral Approach to Traditional Drug Discovery?. *Current Topics in Medicinal Chemistry*. 2016;16(19):2069-77.
13. Walker VM, *et al*. Repurposing antihypertensive drugs for the prevention of Alzheimer's disease: A Mendelian randomization study. *International Journal of Epidemiology*. 2019.
14. Abbruzzese C, *et al*. Drug repurposing for the treatment of glioblastoma multiforme. *Journal of Experimental and Clinical Cancer Research*. 2017;36(1):169.
15. Abbasi J. Repurposing Drugs to Treat Zika. *Jama*. 2016;316(16):1636.
16. Bibi N, *et al*. Inhibitory mechanism of an anticancer drug, Bexarotene against Amyloid beta peptide aggregation: Repurposing via neuroinformatics approach. *Current Pharmaceutical Design*. 2019;25(27):2989-95.
17. Albinana V, *et al*. Repurposing propranolol as a drug for the treatment of retinal haemangioblastomas in von Hippel-Lindau disease. *Orphanet Journal of Rare Diseases*. 2017;12(1):122.
18. Ahmad A, *et al*. The clinically used PARP inhibitor olaparib improves organ function, suppresses inflammatory responses and accelerates wound healing in a murine model of third-degree burn injury. *British Journal of Pharmacology*. 2018;175(2):232-45.
19. Niculescu AB, *et al*. Towards precision medicine for pain: Diagnostic biomarkers and repurposed drugs. *Molecular Psychiatry*. 2019;24(4):501-22.
20. Amar D, *et al*. Integrated analysis of numerous heterogeneous gene expression profiles for detecting robust disease-specific biomarkers and proposing drug targets. *Nucleic Acids Research*. 2015;43(16):7779-89.
21. Kim Y, *et al*. CONCORD biomarker prediction for novel drug introduction to different cancer types. *Oncotarget*. 2018;9(1):1091-106.
22. Zhang N, *et al*. Computational Drug Discovery in Chemotherapy-induced Alopecia via Text Mining and Biomedical Databases. *Clinical Therapeutics*. 2019;41(5):972-80. Abrams ZB, *et al*. Text Mining and Data Modeling of Karyo-

- types to aid in Drug Repurposing Efforts. *Studies in Health Technology and Informatics*. 2015;216:1037.
23. Tari LB, Patel JH. Systematic drug repurposing through text mining. *Methods in Molecular Biology*. 2014;1159:253-67.
  24. Su EW. Drug Repositioning by Mining Adverse Event Data in ClinicalTrials.gov. *Methods in Molecular Biology*. 2019;1903:61-72.
  25. Yang HT, *et al.* Literature-based discovery of new candidates for drug repurposing. *Briefings in Bioinformatics*. 2017;18(3):488-97.
  26. Jang G, *et al.* Literature-based prediction of novel drug indications considering relationships between entities. *Molecular Biosystems*. 2017;13(7):1399-405.
  27. Jang G, *et al.* PISTON: Predicting drug indications and side effects using topic modeling and natural language processing. *Journal of Biomedical Informatics*. 2018;87:96-107.
  28. Huang CH, *et al.* Drug repositioning for non-small cell lung cancer by using machine learning algorithms and topological graph theory. *BMC Bioinformatics*. 2016;17(Suppl 1):2.
  29. Jamal S, *et al.* Checking the STEP-Associated Trafficking and Internalization of Glutamate Receptors for Reduced Cognitive Deficits: A Machine Learning Approach-Based Cheminformatics Study and Its Application for Drug Repurposing. *PLoS One*. 2015;10(6):e0129370.
  30. Nath A, *et al.* Prediction of Human Drug Targets and Their Interactions Using Machine Learning Methods: Current and Future Perspectives. *Methods in Molecular Biology*. 2018;1762:21-30.
  31. Lotfi SM, *et al.* A review of network-based approaches to drug repositioning. *Briefings in Bioinformatics*. 2018;19(5):878-92.
  32. Cheng F, *et al.* Network-based approach to prediction and population-based validation of *in silico* drug repurposing. *Nature Communications*. 2018;9(1):2691.
  33. Berenstein AJ, *et al.* A Multilayer Network Approach for Guiding Drug Repositioning in Neglected Diseases. *PLoS Neglected Tropical Diseases*. 2016;10(1):e0004300.
  34. Alaimo S, *et al.* DTWeb: A web-based application for drug-target interaction and drug combination prediction through domain-tuned network-based inference. *BMC Systems Biology*. 2015;9(Suppl 3):S4.
  35. Tian Z, *et al.* Computational drug repositioning using meta-path-based semantic network analysis. *BMC Systems Biology*. 2018;12(Suppl 9):134.
  36. Mullen J, *et al.* Mining integrated semantic networks for drug repositioning opportunities. *Peer J*. 2016;4:e1558.
  37. Qu XA, *et al.* Inferring novel disease indications for known drugs by semantically linking drug action and disease mechanism relationships. *BMC Bioinformatics*. 2009;10(Suppl 5):S4.
  38. Liu J, *et al.* Combined pharmacophore modeling, 3D-QSAR and docking studies to identify novel HDAC inhibitors using drug repurposing. *Journal of Biomolecular Structure and Dynamics*. 2019;1-15.
  39. Huang T, *et al.* MOST: Most-similar ligand based approach to target prediction. *BMC Bioinformatics*. 2017;18(1):165.
  40. Dash P, *et al.* Three-dimensional models of Mycobacterium tuberculosis proteins Rv1555, Rv1554 and their docking analyses with sildenafil, tadalafil, vardenafil drugs, suggest interference with quinol binding likely to affect protein's function. *BMC Structural Biology*. 2018;18(1):5.
  41. DiMuzio E, *et al.* DockingApp: A user friendly interface for facilitated docking simulations with AutoDock Vina. *Journal of Computer-aided Molecular Design*. 2017;31(2):213-8.
  42. Ozdemir ES, *et al.* Methods for Discovering and Targeting Druggable Protein-Protein Interfaces and Their Application to Repurposing. *Methods in Molecular Biology*. 2019;1903: 1-21.
  43. Huang H, *et al.* Reverse Screening Methods to Search for the Protein Targets of Chemopreventive Compounds. *Frontiers in Chemistry*. 2018;6:138.
  44. Li H, *et al.* Combining *in vitro* and *in silico* Approaches to Find New Candidate Drugs Targeting the Pathological Proteins Related to the Alzheimer's Disease. *Current Neuropharmacology*. 2018;16(6):758-68.
  45. Koudijs KKM, *et al.* Transcriptome Signature Reversion as a Method to Reposition Drugs against Cancer for Precision Oncology. *Cancer Journal*. 2019;25(2):116-20.
  46. Iorio F, *et al.* A Semi-Supervised Approach for Refining Transcriptional Signatures of Drug Response and Repositioning Predictions. *PLoS One*. 2015;10(10):e0139446.
  47. Lee H, *et al.* Drug Repositioning for Cancer Therapy Based on Large-Scale Drug-Induced Transcriptional Signatures. *PLoS One*. 2016;11(3):e0150460.
  48. Zhao C, *et al.* Fortunellin protects against high fructose-induced diabetic heart injury in mice by suppressing inflammation and oxidative stress via AMPK/Nrf-2 pathway regulation. *Biochemical and Biophysical Research Communications*. 2017;490(2):552-9.
  49. Alvarino R, *et al.* Evaluation of the Protective Effects of Saralins on H<sub>2</sub>O<sub>2</sub>-Induced Mitochondrial Dysfunction and Oxidative Stress in SH-SY5Y Neuroblastoma Cells. *Neurotoxicity Research*. 2017;32(3):368-80.
  50. Chen C, *et al.* Carboxyamidotriazole alleviates muscle atrophy in tumor-bearing mice by inhibiting NF-kappaB and activating SIRT1. *Naunyn-Schmiedeberg's Archives of Pharmacology*. 2017;390(4):423-33.
  51. Huang F, *et al.* Protective Effects and Mechanism of Meretrix meretrix Oligopeptides against Nonalcoholic Fatty Liver Disease. *Marine Drugs*. 2017;15(2):31.
  52. Chang TM, *et al.* Extract from Periostracum cicadae Inhibits Oxidative Stress and Inflammation Induced by Ultraviolet B Irradiation on HaCaT Keratinocytes. *Evidence-based Complementary and Alternative Medicine: eCAM*. 2017;2017:8325049.

## Appendix 1: Query for Identifying Potential AD Repurposed Treatments.

The most general form of the TR query can incorporate any number of biomarkers and other markers of interest. For AD, a two-biomarker query was deemed adequate for demonstration purposes. The generic form of the two biomarker AD treatment repurposing query is

(A and B) not (C or D), where

A is a biomarker and its associated desired direction of change

B is another biomarker and its associated direction of change

C is the query used to retrieve the AD core literature

D is a list of existing AD treatments identified in the initial part of the AD study

Thus, the combination (A and B) retrieves ALL records from the full biomedical literature that contain potential AD treatments based on the two desired characteristics A and B, while

(C or D) subtracts those records and existing treatments associated with the AD core literature. The remainder is non-AD records with substances that could be candidate repurposed AD treatments, based on the requirement that A and B must be present.

Twenty of the more than 200 biomarkers identified in the AD study (through text mining techniques) were selected for the query. The query was run in Thompson-Reuters-Medline, since its search engine allows for proximity searching (e.g., [direction] within three words of [biomarker], or [direction] near/3 [biomarker]). In modular form, each query term is shown as follows:

#1 - (reduc\* OR decreas\* OR prevent\* OR attenuat\* OR suppress\* OR alleviat\* OR ameliorat\*) near/3 "oxidative stress"

#2 - (reduc\* OR decreas\* OR prevent\* OR attenuat\* OR suppress\* OR alleviat\* OR ameliorat\*) near/3 "apoptosis"

#3 - ((protect* OR improv* OR enhanc* OR restor* OR preserv*) near/3 “mitochondrial function”) OR ((reduc* OR decreas* OR prevent* OR attenuat* OR suppress* OR alleviat* OR ameliorat*) near/3 “mitochondrial dysfunction”	The final query entered into the search engine (could have been condensed) was:
#4 - (inhibit* OR reduc* OR attenuat* OR decreas*) near/3 “BACE1”	#1 AND (#2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#5 - (modulat* OR attenuat* OR reduc* OR inhibit* OR decreas*) near/3 “gamma-secretase”	#2 AND (#3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#6 - (enhanc* OR increas* OR improv* OR protect*) near/3 “autophagy”	#3 AND (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#7 - (attenuat* OR reduc* OR prevent* OR inhibit*) near/3 (“caspase* activ*” OR “caspase* express”)	#4 AND (#5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#8 - (increas* OR restor* OR enhanc*) near/3 “Bcl-2”	#5 AND (#6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#9 - (attenuat* OR reduc* OR inhibit* OR decreas* OR prevent*) near/3 “NF-kappaB”	#6 AND (#7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#10 - (increas* OR enhanc* OR restor*) near/3 “ADAM10”	#7 AND (#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#11 - (increas* OR restor* OR enhanc*) near/3 “CREB”	#8 AND (#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#12 - (inhibit* OR decreas* OR reduc* OR attenuat*) near/3 “GSK-3”	#9 AND (#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#13 - (increas* OR enhanc* OR restor*) near/3 “GLP-1”	#10 AND (#11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#14 - (increas* OR enhanc* OR restor*) near/3 “ABCA1”	#11 AND (#12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#15 - (increas* OR enhanc* OR restor*) near/3 “nor-epinephrine”	#12 AND (#13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#16 - (increas* OR enhanc* OR restor*) near/3 “Nrf2”	#13 AND (#14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#17 - (increas* OR enhanc* OR restor*) near/3 “seladin-1”	#14 AND (#15 OR #16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#18 - (increas* OR enhanc* OR restor*) near/3 “LRP1”	#15 AND (#16 OR #17 OR #18 OR #19 OR #20)) NOT (#21 OR #22)
#19 - (increas* OR enhanc* OR restor*) near/3 “SIRT1”	
#20 - (attenuat* OR reduc* OR inhibit* OR decreas*) near/3 “beclin1”	
NOT ((#21 - alzheimer* OR dementia OR “mild cognitive impairment”) [Core AD Literature]	
OR	
Existing AD Treatments (see full list of the existing AD treatments that were used in Appendix 6-1 of the AD study)	



(#17 AND (#18 OR #19 OR #20)) NOT (#21 OR #22)

(#18 AND (#19 OR #20)) NOT (#21 OR #22)

(#19 AND (#20)) NOT (#21 OR #22)

Many thousands of candidate repurposed AD treatment records were retrieved. While five were selected and validated, many more were available.

The AD study generated ~200 biomarkers and ~50 symptoms. The 20 terms selected for the above query were all biomarkers. The PN/PAD study, building off the experience from the AD study, has identified ~750 biomarkers and ~250 symptoms. More are still possible.

For a query consisting of 20 terms, there are 190 combinations of two, according to the binomial coefficient. Assume that, for the disease of interest, there are 750 credible biomarkers, as in our ongoing study. Then, using all 750 terms in the query, there would be 280,875 combinations of two.

Many of these combinations, where each member has low record frequency in the literature, would probably yield very little or no retrievals. But, combinations of high frequency terms would yield voluminous results. For more focus, combinations of three or four could be used, but these would have to consist of mid-to high record frequency terms to provide significant retrievals.

## Appendix 2: TR Literature Taxonomy based on Text Clustering.

### Overview of TR text clustering literature taxonomy

The 2890 Medline records that constitute the core TR (mainly drug repurposing) literature were sub-divided into a 32-cluster hierarchical taxonomy using the CLUTO text clustering software [CLUTO, 2018]. Both the top-level aggregated clusters in the taxonomy (Table A2-1) and the 32 elemental clusters (the lowest and most detailed level of the taxonomy-Table A2-2) will be presented in the following sections.

The version of CLUTO used for this analysis does not include fuzzy clustering, so each record is assigned to one cluster only. Many records contained multiple themes and could have been assigned to more than one cluster. Nevertheless, the taxonomy does provide a unique and interesting perspective on the structure of the TR literature.

The themes of the clusters shown in Tables A2-1 and A2-2 are, of necessity, very broad. In Appendix 2 of (Kostoff RN. Treatment Repurposing using Literature-Related Discovery. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/60507>), the titles are provided for each of the 32 elemental clusters shown in Table A2-2, to provide the full spectrum of sub-themes within each elemental cluster and allow the interested reader to identify specific

sub-themes of personal interest within the cluster. The full reference for each title is provided in Chapter 4 of the same referenced document, which will allow the reader to pursue the full text for further information.

Before the details of the 32 'leaf' (lowest level) clusters are presented, a high-level (top three levels) view of the TR text clustering taxonomy is shown in Table A2-1.

<b>Table A2-1: Top-level TR Text Clustering Taxonomy.</b>		
(number of records in each cluster shown in parenthesis ( ))		
SECOND LEVEL	THIRD LEVEL	FOURTH LEVEL
Cluster 58 (1096) Drug repurposing prediction	Cluster 51 (439) - Gene expression; protein interaction network	Cluster 35 (193) - Drug-disease associations; protein interaction networks
		Cluster 41 (246) - Gene expression; genome wide association
	Cluster 56 (657) - Drug-target interaction	Cluster 48 (276) - Drug-target interaction; protein-protein interaction
Cluster 61 (1793) Disease treatment	Cluster 59 (569) - Drug-resistant tuberculosis; infectious diseases	Cluster 7 (114) - Drug-resistant tuberculosis
		Cluster 57 (455) - Viral/bacterial infections; parasites
	Cluster 60 (1224) - chronic disease treatments; cancer; neurodegenerative diseases	Cluster 52 (590) - Cancer treatment
		Cluster 55 (634) - Neurodegenerative disease treatment

The first bifurcated level of the hierarchical taxonomy shows two definite thrust areas: Methods for drug repurposing prediction (Cluster 58) and disease treatments that resulted from drug repositioning (Cluster 61). The next two levels of the hierarchy are self-explanatory.

Table A2-2 relates the eight fourth-level clusters shown above in Table A2-1 to the 32 elemental leaf clusters.

<b>Table A2-2: Lowest Level TR Text Clustering Taxonomy.</b>	
(number of records in each cluster shown in parenthesis ()); this number is hyperlinked to the actual record titles)	
FOURTH LEVEL	LOWEST LEVEL (LEAF CLUSTERS)
Cluster 35 (193) - Drug-disease associations; protein interaction networks	Cluster 4 (47) - drug-disease associations
	Cluster 18 (146) - network-based prediction, especially protein interaction networks
Cluster 41 (246) - Genome-wide associations; gene expression	Cluster 16 (65) - genome-wide associations
	Cluster 21 (181) - gene expression, especially gene expression signatures and gene expression profiles

Cluster 48 (276) - Drug-target interaction prediction; protein-protein interaction	Cluster 13 (100) - drug-target interaction prediction
	Cluster 22 (176) - ligand binding-sites, protein-ligand interactions and protein-protein interactions
Cluster 49 (381) - Computer-aided drug repositioning; drug discovery	Cluster 9 (37) - rare diseases
	Cluster 20 (91) - computational drug repositioning
	Cluster 26 (86) - marketing aspects of drug repurposing
	Cluster 29 (167) - drug development and discovery
Cluster 7 (114) - Drug-resistant tuberculosis	Cluster 7 (114) - drug-resistant tuberculosis
Cluster 57 (455) - Viral/bacterial infections; parasites	Cluster 1 (41) - antiviral treatments for viral infections, especially Ebola virus
	Cluster 0 (21) - antiviral treatments for viral infections, especially Zika virus
	Cluster 15 (80) - antiviral treatments for other viral infections, especially dengue virus, hepatitis B virus, chikungunya virus, human immunodeficiency virus, japanese encephalitis virus, rift valley fever virus, human cytomegalovirus, respiratory syncytial virus, west Nile virus
	Cluster 14 (62) - treatments for parasites, especially trypanosoma cruzi, african trypanosomiasis, trypanosoma brucei, leishmania amazonensi
	Cluster 19 (93) - treatments for parasites, especially plasmodium falciparum, schistosoma mansoni, toxoplasma gondii
	Cluster 6 (45) - antifungal treatments
	Cluster 17 (113) - antimicrobial and antibiotic treatments for infections
	Cluster 23 (93) - repurposing kinase inhibitors, especially for treatment of acute myeloid leukemia
Cluster 52 (590) - Cancer treatment	Cluster 12 (59) - ovarian cancer treatments, especially niclosamide
	Cluster 27 (160) - treatments that destroy cancer cells
	Cluster 25 (171) - anti-cancer treatments
	Cluster 5 (41) - treatments for pancreatic cancer, especially Metformin
	Cluster 11 (66) - breast cancer treatments
	Cluster 2 (48) - Alzheimer's Disease treatments
Cluster 55 (634) - Neurodegenerative disease treatment	Cluster 8 (30) - neurodegenerative disease treatments, especially Parkinson's Disease
	Cluster 24 (117) - treatments for brain disease, especially stroke
	Cluster 3 (33) - drug repurposing patent applications
	Cluster 10 (46) - glioblastoma treatments
	Cluster 31 (151) - anti-inflammatory treatments
	Cluster 28 (97) - treatments for addiction disorders (especially alcohol use) and chronic pain
	Cluster 30 (112) - cancer treatments, especially metronomic chemotherapy

### Appendix 3: TR Literature Taxonomy based on Factor Analysis.

#### Overview of TR literature taxonomy based on Factor Analysis

The previous appendix provided one perspective (text clustering) on the taxonomic structure of the TR literature. The CLUTO text clustering software incorporates all phrases (minus stop-words) and uses one selected algorithm to generate a hierarchical taxonomy. Another approach providing a complementary perspective on the TR literature structure is factor analysis. Here, only pre-selected phrases are used.

The present appendix contains the results of a 37-factor study, where three of the factors define strong themes (high factor loadings) at each end of the phrase list (Factors 5, 8, 11). The main theme of each factor is presented here. A full factor matrix that identifies the main theme of each factor (and shows the key phrases that determine each theme) is presented in Appendix 3 of (Kostoff RN. Treatment Repurposing using Literature-Related Discovery. Georgia Institute of Technology. 2018. PDF. <https://smartech.gatech.edu/handle/1853/60507>). In this reference, the titles of the records associated with each of the key phrases in the 37 themes is presented as well. The factor analysis was conducted using the Vantage Point software ([thevantagepoint.com](http://thevantagepoint.com)).

#### Results of Factor Analysis

While many broad thematic categories were possible, four appeared to be dominant. These included: Repurposing Prediction Methodologies, Diseases, Biomarker Targets and Drug Types. Figure A3-1 lists all the 37 factor themes. The main diseases studied are cancer, neurodegenerative and infectious. The main biomarker targets studied focus on oxidative stress and inflammatory metrics. While drugs of many different classes have been researched for repurposing, the main drug classes as emphasized in Figure A3-1 are inhibitors of myriad signaling pathways. Finally, the main repurposing prediction methodologies studied focus on networks, similarity, machine

1. *In vitro* Experiments of Repurposed Drug Candidates on Cancer Cell Lines
2. Gene Expression Signatures for Predicting Repurposed Drugs
3. Antiviral Protease Inhibitors for Cancer Therapy
4. Neurodegenerative Diseases Biomarkers for Repurposing Targets
- 5a. Antifungal Applications of Non-Antifungal Drugs
- 5b. Network-Based Approaches to Drug Repositioning
6. Antimicrobial Applications of Repurposed Drugs
7. Repurposed Drugs or Antifungal Applications
- 8a. Use of Cellular Signatures Library to Provide Gene Expression Profiles for Drug Repurposing Prediction
- 8b. Ligand-Based Target Inference
9. Binding Site Analysis for Drug Repurposing
10. Antiinflammatory Applications for Repurposed Drugs

- 11a. Biomarkers for Repurposed Drug-Enhanced Apoptosis of Cancer Cells
- 11b. Repurposed Drugs for Oxidative Stress Reduction
12. Repurposed Drugs that Increase or Decrease ROS for Different Applications
13. Similarity Searching of Ligand-Target Sets for Drug Repurposing
14. AMPK Activation for Cancer Treatment, Emphasizing Anti-Diabetic Drug Metformin
15. Repurposing of Cholesterol-Lowering Drugs for Chronic and Infectious Diseases
16. Similarity-Based Methods for Drug Repurposing
17. Machine Learning-Based Drug Repurposing Prediction
18. Tyrosine Kinase Inhibitors Repurposed for Cancer Treatment
19. Chemical Structure Similarity for Repurposing Prediction
20. Network-Based Inference for Predicting Drug-Target Interaction
21. Drug Repurposing for Viral Diseases
22. Drug Repurposing for Anti-Parasitic Applications
23. Phosphodiesterase Inhibitors Repurposed from Predictions of Drug Response Signatures
24. Inhibiting NF-KappaB Signaling for Cancer and Inflammation Treatment
25. Repurposing Based on Comprehensive Multi-Metric Similarity Measures
26. Genome-Wide Association-Based Networks for Repurposing
27. Drug Repurposing for Brain Cancer
28. Repurposing Anthelmintic Drugs for Cancer Treatment
29. Drug Repurposing for Neurodegenerative Diseases
30. Repurposed Drugs Targeting Glutamate Receptors
31. Repurposing Drugs that Target Oxidative and Inflammation Biomarkers VEGF, HO-1, iNOS, Nrf2
32. Computational Drug Repositioning Based on Similarity Networks
33. Repurposing Antipsychotic Drugs
34. Multiple Ligand Simultaneous Docking and Drug Repositioning for Cancer Therapy