Video Article

# Drug Repurposing Hypothesis Generation Using the "RE:fine Drugs" System

Kelly Regan[1], Soheil Moosavinasab[2], Philip Payne[1], Simon Lin[2]

[1]Department of Biomedical Informatics, The Ohio State University

[2]Research Information Solutions and Innovation, The Research Institute at Nationwide Children's Hospital

Correspondence to: Kelly Regan at Kelly.Regan@osumc.edu

URL: http://www.jove.com/video/54948
DOI: doi:10.3791/54948

Keywords: Medicine, Issue 118, Drug repurposing, Drug discovery, Database, Hypothesis generation, Genome-wide association study (GWAS), Phenome-wide association study (PheWAS)

Date Published: 12/11/2016

Citation: Regan, K., Moosavinasab, S., Payne, P., Lin, S. Drug Repurposing Hypothesis Generation Using the "RE:fine Drugs" System. *J. Vis. Exp.* (118), e54948, doi:10.3791/54948 (2016).

## Abstract

The promise of drug repurposing is that existing drugs may be used for new disease indications in order to curb the high costs and time for approval. The goal of computational methods for drug repurposing is to enable solutions for safer, cheaper and faster drug discovery. Towards this end, we developed a novel method that integrates genetic and clinical phenotype data from large-scale GWAS and PheWAS studies with detailed drug information on the concept of transitive Drug-Gene-Disease triads. We created "RE:fine Drugs," a freely available, interactive dashboard that automates gene, disease and drug-based searches to identify drug repurposing candidates. This web-based tool supports a user-friendly interface that includes an array of advanced search and export options. Results can be prioritized in a variety of ways, including but not limited to, biomedical literature support, strength and statistical significance of GWAS and/or PheWAS associations, disease indications and molecular drug targets. Here we provide a protocol that illustrates the functionalities available in the "RE:fine Drugs" system and explores the different advanced options through a case study.

## Video Link

The video component of this article can be found at http://www.jove.com/video/54948/

## Introduction

The costly and inefficient processes associated with traditional drug discovery approaches, including high-throughput drug and lead compound screening, are contributing to delays in translating research discoveries into therapies for patients [1,2]. An average of 1 billion U.S. dollars and 15-20 years is required to bring a new drug from the bench to the bedside [3]. Further, 52% of drugs fail during development in phase 1 clinical trials, and only 25% of compounds that enter phase 2 proceed into full phase 3 clinical studies [4]. The goal of drug repurposing or drug repositioning is to renew failed drugs and/or find novel indications for approved drugs in order to deliver new therapies to patients faster and with a higher success rate. Drug repurposing may decrease the timeline to make drugs available for use in patients to 3-12 years [5]. Important medical applications for drug repurposing include: diseases with poor prognosis and low survival rates, drug-resistant diseases, underfunded disease research areas and impoverished and underserved patient populations.

Computational drug repurposing is defined as the process of designing and validating automated workflows that can generate hypotheses for new indications for a drug candidate [6]. Existing computational drug repurposing methods have been categorized target-based, knowledge-based, signature-based, network-based, and targeted-mechanism-based, and can be oriented from gene, disease or drug perspectives. Furthermore, computational approaches may further accelerate proof-of-concept validation experiments and small-scale clinical studies for repurposed drug candidates [7]. We have previously reported on "RE:fine Drugs", a freely available, interactive web-based tool for drug repurposing hypothesis generation based on the transitive theory of Drug-Gene-Disease relationships [8]. The overall goal of this method is to systematically integrate diverse types of drug, genetic and clinical data to enable drug repurposing for users from diverse communities, including clinical, industry and regulatory communities. The foundational methods for this system have been previously reported for the use of genome-wide association study (GWAS) and phenome-wide association study (PheWAS) data in drug repurposing research [9,10]. The novel combination of these types of data distinguishes our webtool from other target-based methods [6,11].

The RE:fine Drugs system currently contains 60,911 drug repurposing hypotheses covering 916 drugs, 567 genes and 1,770 diseases. The webtool provides a user-friendly interface for researchers to interactively search drug repurposing hypotheses and prioritize them using diverse criteria. For instance, users can filter drug repurposing hypotheses with support in the biomedical literature and clinical trials database, significant p-values, association odds ratios or by specific indications. The only requirement for this system is Internet access.

## Protocol

# 1. Initiation of Queries from Gene, Drug or Disease Terms

1. Access the homepage for "RE:fine Drugs" at the following link: http://drug-repurposing.nationwidechildrens.org. Begin by entering a query term in the search bar from any of the following three categories: drug (generic drug name), disease (new disease indication) or gene (official HGNC gene symbol).
2. Filter the search bar function to include only "Drug Name", "New Disease Indication", "Gene Symbol" or "All" categories. The search bar includes an auto-fill function for query entries.
3. Put in a keyword, and click on the "Search" button. Sort the table of results by any of the following columns: "Drug", "Registered Indication", "P-Value", "P-Value Adjusted", "Odds Ratio", "Study", "New Indication", "Drug Bank Indication", "# of Medline Abstracts", "# of Clinical Trial Registry", "Potential", "SNP", "Gene" or "action".
4. Navigate to the advanced search option in order to enable the drug information feature. Click on the icon in the "Info" column for a particular drug. Observe a page that lists the all the corresponding information including p-Value for the association, Disease name, Drug name, Gene details (NCBI Gene link) and Drug details (DrugBank link).

# 2. Exploration of Advanced Options

1. Click on the "advanced" button located on the right side of the page, and several options to further refine the results are provided. The advanced search options include modifications to the following: drug, association, disease, potential, gene and action.
2. Export results tables by clicking on the "Export" button on the right side of the page. Click on the "Simple" button in order to fold down the advanced search window.
3. Under the advanced option "drug" tab, specify a particular drug indication or an additional drug name to filter results.
4. Under the "association" tab, filter results by the significance level P-Value, Adjusted P-Value with FDR, effect size (Odds Ratio), and/or Study type (GWAS, PheWAS or Both).
5. Under the "disease" tab, specify a certain disease description for the predicted new use.
6. Under the "potential" tab, filter results according to any of the following criteria: (i) whether the drug indication is contained in the DrugBank database, (ii) number of Medline abstracts with co-occurrence of the drug and the disease, (iii) number of ClinicalTrials.gov database entries with co-occurrence of the drug and the disease and (iv) Repurposing Potential.
   NOTE: The Repurposing Potential option describes the novelty of the discovery: (i) Known: relationship already exists in the DrugBank database, (ii) Strongly supported: some support in both clinical trial registry and the Medline abstracts, (iii) Likely: some support in either clinical trial registry or the Medline abstracts and (iv) Novel: no evidence in clinical trial registry nor in Medline abstracts.
7. Under the "gene" tab, enter a SNP identifier or Gene Symbol to filter results by specific drug target genes.
8. Under the "action" tab, specify the drug action type against the drug target(s), including agonist, antagonist, other, unknown or all (source: DrugBank database).

## Representative Results

In this example, the gene "IL2RB" was entered as a gene-based query, and was automatically recognized as such by the auto-fill function (**Figure 1**). The twelve drug repurposing hypotheses for the IL2RB gene are returned, as shown in **Figure 2**. Detailed information page for a particular drug repurposing hypothesis, "daclizumab" in this case, is provided from the "Info" column (**Figure 3**). The results were filtered on the drug tab was by all results corresponding to the "daclizumab" drug, as shown in **Figure 4**. **Figure 5** shows only those drugs with a known indication for the "transplant" disease term (source: DrugBank database). The association tab allows the user to filter SNP-Phenotype relationships by statistical significance (P-value) and genetic effect size (odds ratio), defined as the ratio of the odds of presence of disease in individuals with a specific genotype (SNP allele) over the odds of presence of disease in individuals without the SNP allele. **Figure 6** shows the results filtered under the association tab within the P-Value range of 0.000001 to 0.05. **Figure 7** shows drug repurposing hypotheses specific for "asthma" based on the new indications we found in the study. **Figure 8** shows results under the potential tab to filter by a minimum number of 5 Medline abstracts containing a co-occurrence of drug and disease terms. In this example, all drugs results under the gene tab are targeted for the "IL2RB" gene, corresponding to the original query term (**Figure 9**). Finally, **Figure 10** shows results filtered under the "action" tab to return all drugs that act as agonists on the IL2RB gene.

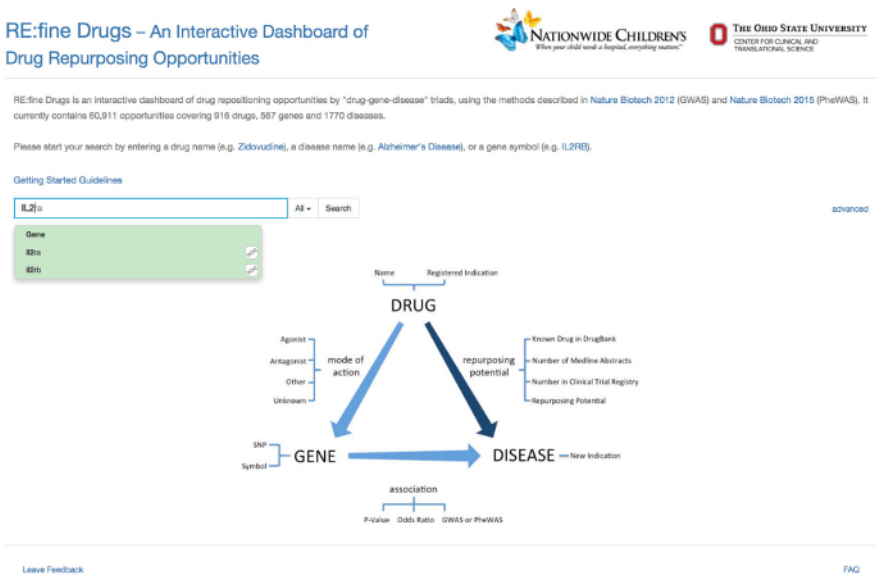


**Figure 1: The RE:fine Drugs interactive dashboard homepage.** Users may begin a query by entering a drug name, new disease indication or gene symbol. Links are also provided for the GWAS and PheWAS reference papers describing methodologies for generating drug repurposing hypotheses. Please click here to view a larger version of this figure.

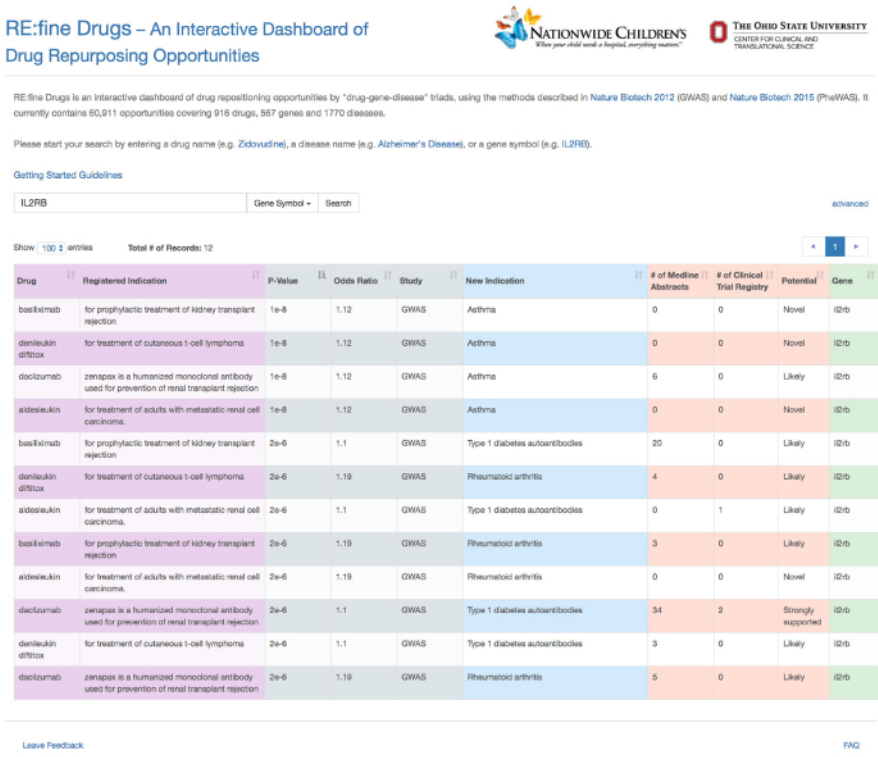


**Figure 2: Auto-fill function for query entries.** As an example, the gene query term "IL2RB" was automatically recognized as a gene term. Please click here to view a larger version of this figure.

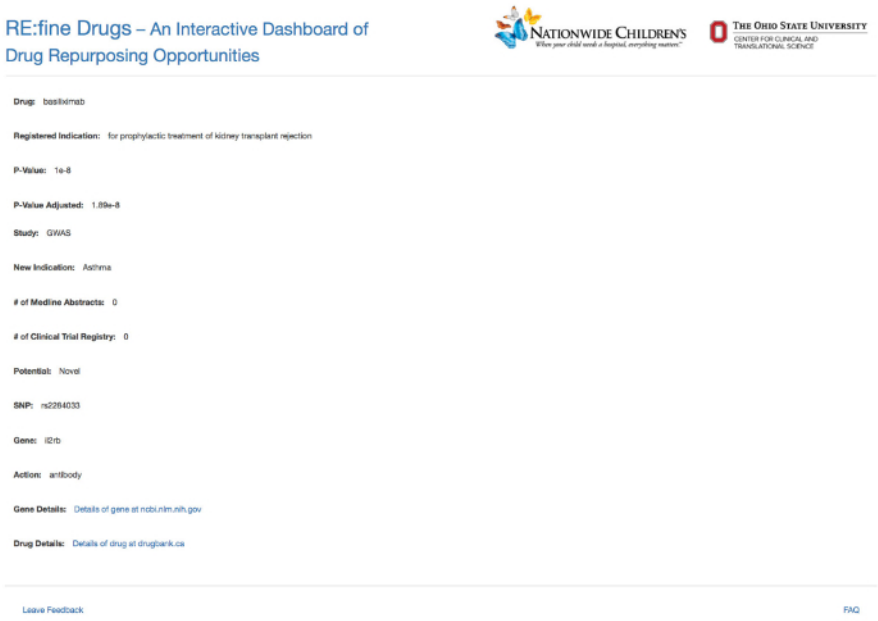


**Figure 3: Drug repurposing results table produced from a gene-based query (*e.g.*, IL2RB).** Twelve drug repurposing hypotheses for the IL2RB gene are produced. Please click here to view a larger version of this figure.

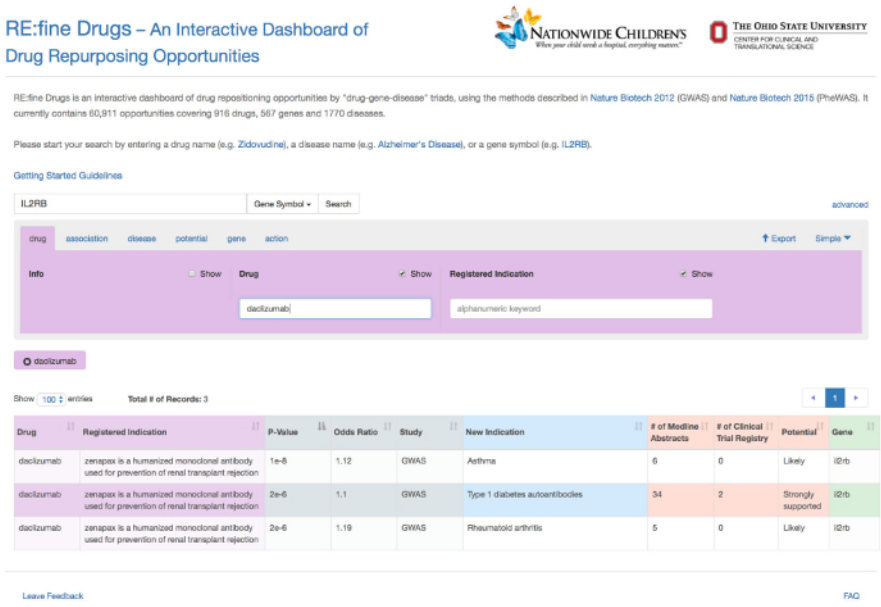


**Figure 4: Information page for individual drugs from results page.** Clicking on the icon from the "Info" column shows detailed information for the drug daclizumab. Please click here to view a larger version of this figure.

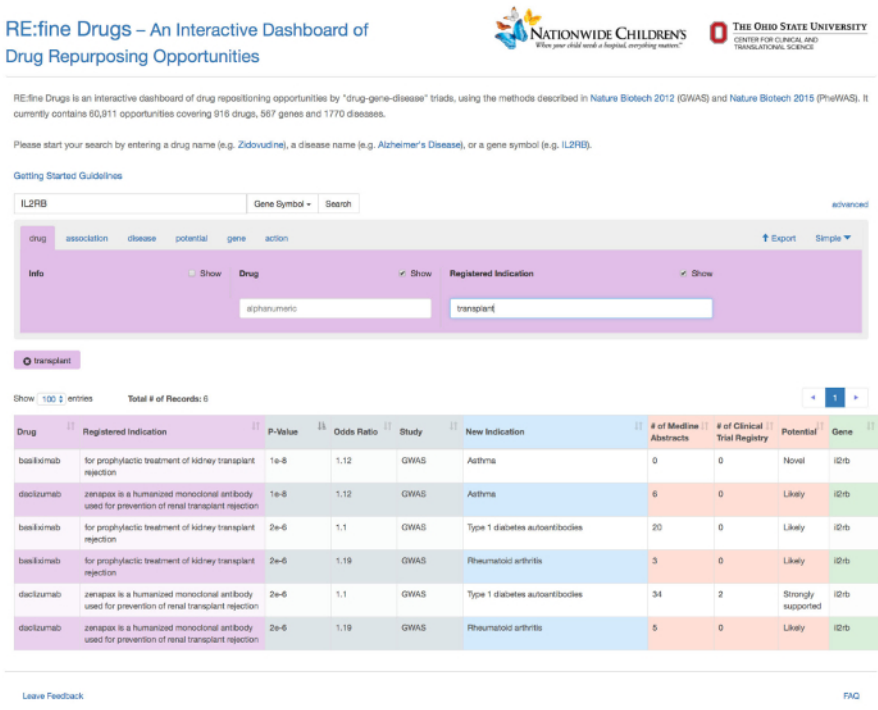


**Figure 5: Advanced search option under drug tab to filter by a specific drug.** In this example, three results are shown for the drug daclizumab. Please click here to view a larger version of this figure.

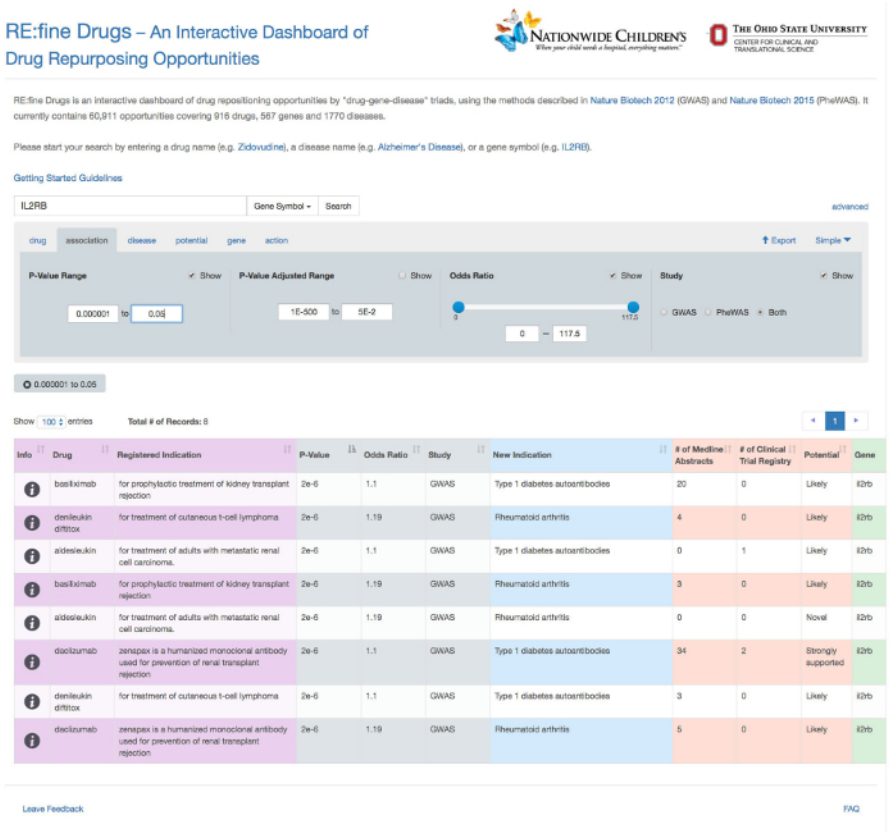


**Figure 6: Advanced search option under drug tab to filter by a specific disease indication from the DrugBank database.** All drugs with a known indication for the disease term "transplant" are shown. Please click here to view a larger version of this figure.

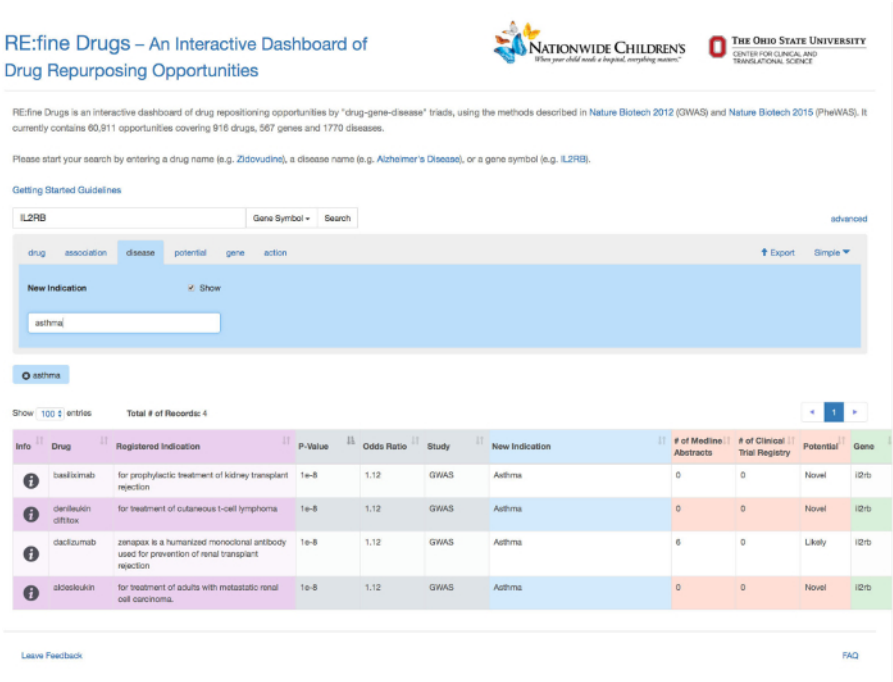


**Figure 7: Advanced search option under association tab to filter by significance level.** In this case, eight results are provided whose association significance level falls within the P-Value range of 0.000001 to 0.05. Please click here to view a larger version of this figure.

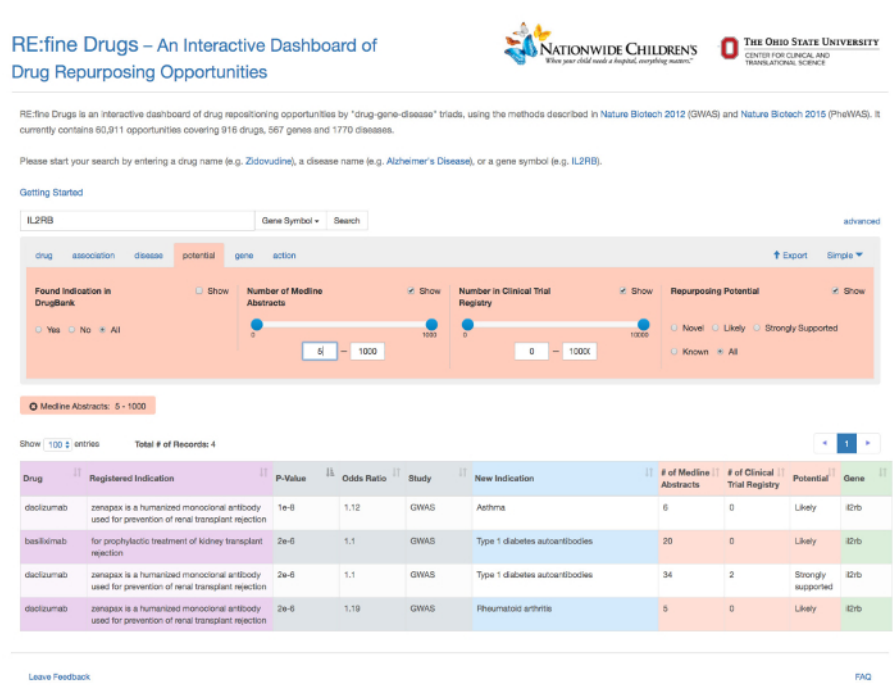


**Figure 8: Advanced search option under disease tab to filter by a specific disease indication we extracted in this study.** In this example, four results are shown for asthma as a new use disease indication. Please click here to view a larger version of this figure.

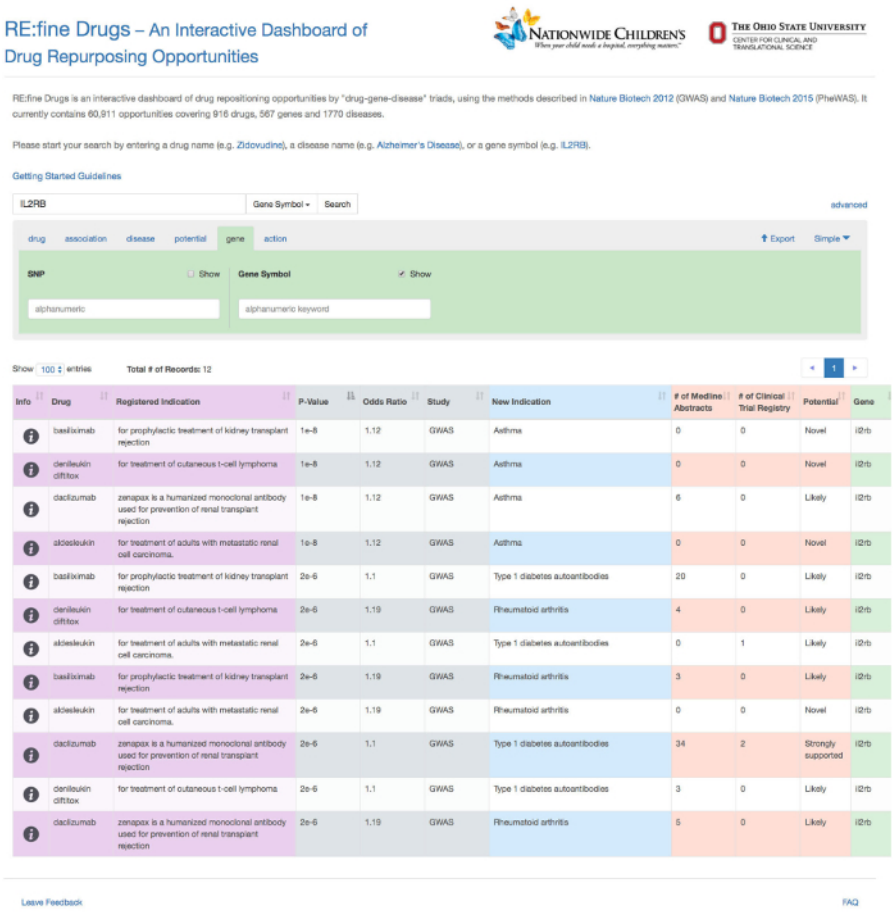


**Figure 9: Advanced search option under potential tab to filter by a co-occurrence of drug and disease terms in Medline abstracts.** In this example, four results are shown that are supported by a minimum number of 5 Medline abstracts containing a co-occurrence of drug and disease terms. Please click here to view a larger version of this figure.

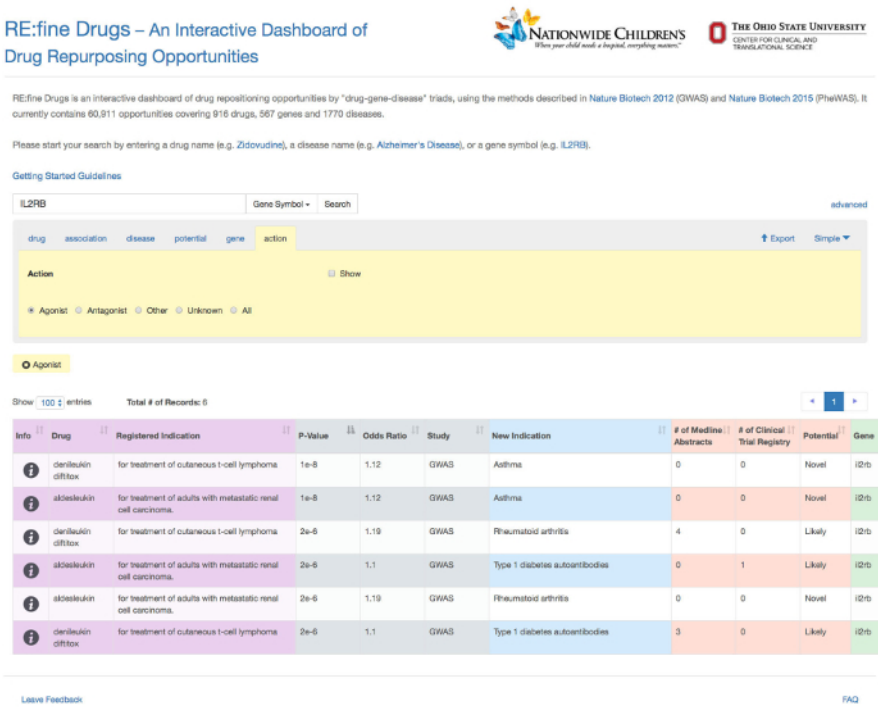


**Figure 10: Advanced search option under gene tab to filter by a specific gene symbol, where all results correspond to the IL2RB gene used for the original query.** In this example, all drugs results under the gene tab are targeted for the "IL2RB" gene, corresponding to the original query term. Advanced search option under action tab to filter by agonist drugs only. In this example, six results are returned for all drugs that act as agonists on the IL2RB gene. Please click here to view a larger version of this figure.

## Discussion

The protocol described here for the RE:fine Drugs interactive dashboard can be modified in different ways according to the user's preferences. This method uniquely integrates GWAS and PheWAS data as a novel paradigm underlying drug repurposing hypothesis generation. Specifically, this system provides access to both 52,966 PheWAS associations and 7,945 GWAS associations with advanced options to filter the results by the study type, effect size and/or significance level. Another advantage of this method over existing computational drug repurposing tools is that queries may be made from drug, gene or disease perspectives.

There are several limitations to this method. Currently, the PheWAS data is limited to primarily adult patient population from five institutions contained in the Electronic Medical Records and Genomics (eMERGE) network with a mean age of 69.5 years [12]. Additionally, the "repurposing potential" feature uses co-occurrence of search terms in Medline abstracts as one of its criteria. It is well known that text mining methods using co-occurrence have limitations with respect to syntactical structure and literature bias. Thus, we recommend this feature be used as a starting point to explore the potential novelty and/or evidence supporting specific drug repurposing hypotheses and recommend additional investigation into the biomedical literature and clinical trial databases.

Future directions for this work not described here would be to extend this database to additional sources of GWAS and PheWAS data as they become available. Similar efforts to systematically translate results from large-scale GWAS studies into drug repurposing hypotheses have been previously published [9,13-14]. It may be useful to compare these different workflows to predict drug candidates from GWAS data in future studies. Additionally, several other methods exist to computationally generate drug repurposing hypotheses from different data sources, including: genomics, transcriptomics, chemical structures, drug side effect profiles, as previously summarized [6,11]. Future methodological advancements could also include automating drug combination predictions and providing information on drug toxicity to guide follow up studies for drug candidates.

Furthermore, the hypotheses generated from RE:fine Drugs may be further validated using electronic health records, before initiating clinical trials [15]. Finally, future studies will be needed to compare this system to other target-based drug repurposing methods.

## Disclosures

The authors declare that they have no competing financial interests.

## Acknowledgements

This work was partially supported by the National Institutes of Health (NIH) Clinical and Translational Science Awards (CTSA) Grant (UL1TR001070) to the Ohio State University's Center for Clinical and Translational Science (CCTS) and the National Library Of Medicine under Award Number T15LM011270.

## References

1. Borisy, A. A. *et al.* Systematic discovery of multicomponent therapeutics. *Proc Natl Acad Sci U S A.* **100** (13), 7977-7982 (2003).
2. Zhang, L. *et al.* High-throughput synergy screening identifies microbial metabolites as combination agents for the treatment of fungal infections. *Proc Natl Acad Sci U S A.* **104** (11), 4606-4611 (2007).
3. Adams, C. P., & Brantner, V. V. Estimating the cost of new drug development: is it really 802 million dollars? *Health Aff (Millwood).* **25** (2), 420-428 (2006).
4. Bunnage, M. E. Getting pharmaceutical R&D back on target. *Nat Chem Biol.* **7** (6), 335-339 (2011).
5. Ashburn, T. T., & Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov.* **3** (8), 673-683 (2004).
6. Hurle, M. R. *et al.* Computational drug repositioning: from data to therapeutics. *Clin Pharmacol Ther.* **93** (4), 335-341 (2013).
7. Jin, G., & Wong, S. T. Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. *Drug Discov Today.* **19** (5), 637-644 (2014).
8. Moosavinasab, S. *et al. "RE:fine Drugs": An Interactive Dashboard to Access Drug Repurposing Opportunities.* Database (Oxford). (2016).
9. Zhang, J. *et al.* Use of genome-wide association studies for cancer research and drug repositioning. *PLoS One.* **10** (3), e0116477 (2015).
10. Rastegar-Mojarad, M., Ye, Z., Kolesar, J. M., Hebbring, S. J., & Lin, S. M. Opportunities for drug repositioning from phenome-wide association studies. *Nat Biotechnol.* **33** (4), 342-345 (2015).
11. Li, J. *et al.* A survey of current trends in computational drug repositioning. *Brief Bioinform.* **17** (1), 2-12 (2016).
12. Denny, J. C. *et al.* Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol.* **31** (12), 1102-1110 (2013).
13. Wang, H., *et al.* Mining drug-disease relationships as a complement to medical genetics-based drug repositioning: Where a recommendation system meets genome-wide association studies. *Clin Pharmacol Ther.* **97** (5), 451-454 (2015).
14. Sanseau, P., *et al.* Use of genome-wide association studies for drug repositioning. *Nat Biotechnol.* **30** (4), 317-320 (2012).
15. Xu, H. *et al.* Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. *J Am Med Inform Assoc.* **22** (1), 179-191 (2015).