Computers in Biology and Medicine 109 (2019) 254–262

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

DrugR+: A comprehensive relational database for drug repurposing, combination therapy, and replacement therapy

Yosef Masoudi-Sobhanzadeh[a], Yadollah Omidi[b], Massoud Amanlou[c], Ali Masoudi-Nejad[a,*]

[a] Laboratory of Systems Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran
[b] Research Center for Pharmaceutical Nanotechnology and Department of Pharmaceutics, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran
[c] Drug Design and Development Research Center, The Institute of Pharmaceutical Sciences (TIPS), Tehran University of Medical Sciences, Tehran, 14176-53955, Iran

ARTICLE INFO

Keywords:
DrugR+
Drug repurposing
Drug repositioning
Database: combination therapy

ABSTRACT

Drug repurposing or repositioning, which introduces new applications of the existing drugs, is an emerging field in drug discovery scope. To enhance the success rate of the research and development (R&D) process in a cost- and time-effective manner, a number of pharmaceutical companies worldwide have made tremendous investments. Besides, many researchers have proposed various methods and databases for the repurposing of various drugs. However, there is not a proper and well-organized database available. To this end, for the first time, we developed a new database based on DrugBank and KEGG data, which is named "DrugR+". Our developed database provides some advantages relative to the DrugBank, and its interface supplies new capabilities for both single and synthetic repositioning of drugs. Moreover, it includes four new datasets which can be used for predicting drug-target interactions using supervised machine learning methods. As a case study, we introduced novel applications of some drugs and discussed the obtained results. A comparison of several machine learning methods on the generated datasets has also been reported in the Supplementary File. Having included several normalized tables, DrugR + has been organized to provide key information on data structures for the repurposing and combining applications of drugs. It provides the SQL query capability for professional users and an appropriate method with different options for unprofessional users. Additionally, DrugR + consists of repurposing service that accepts a drug and proposes a list of potential drugs for some usages. Taken all, DrugR+ is a free web-based database and accessible using (http://www.drugr.ir), which can be updated through a map-reduce parallel processing method to provide the most relevant information.

1. Background

A drug is normally described as a substance that creates temporal pharmacological impacts on the biological target(s) after entering the body [1]. As a result, it prevents or eliminates the intended structural anomalies at the molecular or cellular level [2]. For designing a drug, several rigorous, expensive and time-consuming steps need to be undertaken [3]. Nevertheless, most drug development projects appear to fail during different R&D steps, in large part due to the complexity of the process or lack of the desired clinical outcomes. Hence, only a very few numbers of them may get approved by the US Food and Drug Administration (FDA) [4]. Given the difficulty and complexity of R&D process in pharmaceutical companies, the pharmaceutical industry seeks some other alternative approaches such as drug repurposing that provides new clinical applications for the available drugs [5]. For instance, Minoxidil designed and developed to control hypertension has been proposed to be used in hair loss [6]. Accordingly, a number of governmental settings such as the National Center of Advancing Translational Science (NCATS) and the Medical Research Council (MRC) began to look drug repositioning [7], in large part because (i) the process of drug design takes about 20 years and needs more than 2.5 billion USD whereas drug repositioning reduces time to 3 years and cost to 300 million dollars; (ii) only a few numbers of drug developing projects can get FDA certificate whereas more than 65% of drug repositioning projects are completed; (iii) drug repositioning can revive some drugs that failed in drug development phases but may find new applications; (iv) drug repurposing may change the application of some drugs and propose better uses; (v) many drug developing companies are not eager to invest on the orphan and rare diseases largely because of the lower revenue. In these situations, drug repositioning appears to

* Corresponding author.
E-mail address: amasoudin@ut.ac.ir (A. Masoudi-Nejad).
URL: http://LBB.ut.ac.ir (A. Masoudi-Nejad).

https://doi.org/10.1016/j.compbiomed.2019.05.006
Received 9 April 2019; Received in revised form 26 April 2019; Accepted 6 May 2019
0010-4825/

provide a robust and sound strategy for revalidating new uses for old drugs.

In drug discovery scope, the combination of two or more drug substances possessing different mechanisms of action was shown as an alternative strategy to enhance the success rate of drug repurposing [8]. Further, form the combination therapy viewpoint, the concurrent use of metformin and temozolomide was reported to be beneficial in glioblastoma treatment [9]. Although combination therapy is not applicable to all drugs and might face many striking challenges [10], still such treatment strategy offer many crucial advantages [11], including (i) the profoundly higher synergic effects in comparison with the single drug therapy, (ii) the reduced incidence/emergence of drug resistance mechanisms, (iii) the use of drugs with much lower concentrations and hence trivial adverse reactions, and (iv) the higher success rate of the treatment modality.

With developing new tools and databases, drug repositioning has also been extended [12]. For this purpose, researchers search, analyze, mine, and investigate drug information such as the structural similarity of drugs, their targets and other properties. However, the lack of a suitable database seems to be a sensible pitfall in this field. Hence, we capitalized on a data mining approach based on the "deductive rules" to develop a new database to tackle the limitations associated with the current databases regarding the repositioning of drugs. This newly established database is called "DrugR+", which is freely available at https://drugr.ir. In addition to single drug repositioning, the concept of drugs combination was also developed for repositioning in DrugR + that is termed as "*combined drug replacement* (CDR)". Having used the CDR capacity of the database, a drug (e.g., D1) is replaced with drugs with no/trivial adverse reactions, in which all drugs mined by DrugR + would be capable to associate with the same biological target (s) of D1 and show same/similar mechanism of action, as shown in Fig. S1 (Supplementary File). In addition to CDR, DrugR + supports the combination therapy, in which the performance of a drug can be enhanced upon the combination with drug(s).

2. Related works

Researchers have designed many databases in order to satisfy various goals in drug scope. In this section, we investigate databases which are used in drug repurposing applications. Based on their applications and roles in drug repositioning, as shown in Fig. S2 (Supplementary File), we divided the databases into four categories as follows:

(i) Basic databases: Some databases such as DrugBank [13] and KEGG [14] have provided useful and basic information about drugs, targets, pathways, etc. In various researches for repurposing existing drugs, these databases have been being used. However, they have not been designed for the repositioning of available drugs. Moreover, these databases have some limitations, which might make them somewhat inconvenient for the repositioning of drugs. Therefore, researchers should mine the databases, and acquire their favorable results. Also, they may face lots of technical problems that might reduce the speed of the mining process.

(ii) Repurposed databases: Several databases hold news, articles, and results that are obtained by other drug repositioning studies. RepoDB [15], Excelra [16], Drug repurposing hub (DRH) [17], and TTD [18] are some paradigms of this category. Although these databases establish information about true and false repurposed drugs and play a key role in decision making; they appear to be less helpful in providing repositioning capabilities.

(iii) Similarity-based databases: Based on similarities between a drug and another drug, or between a target and another target; several databases or web services have been developed for the drug repurposing. For example, PREDICT [19] and RepurposeDB [20] are placed in this category. In drug repositioning, the similarity is one of the effective factors. Meanwhile, there are other factors which have key roles for introducing a new application of existing drugs. In other words, the similarity is not the only determinant factor.

(iv) Target based databases: There are other types of databases, which consider drugs and their various targets, and hence, propose new applications. DMAP [21], DrugSig [22], and DDW [23] are placed in these classes of databases, in which each of them has been developed for a specific application such as the orphan, lysosomal, and a small number of gene-driven diseases.

Given the benefits and limitations of the mentioned databases, we developed the DrugR + database. There are several advantages in using DrugR+, including (i) providing a suitable and easy way for searching and getting information on drugs without any technical problems, (ii) offering drug repurposing capabilities such as combination, single, and CDR repurposing, (iii) granting possibility for the selection of different types of targets, and hence, obtaining a list of drugs for determined drug or diseases, and finally (iv) providing frequent updates on the new drugs along with new features of the previous drugs.

3. Construction and content

DrugBank, which contains different useful information about drugs, is one of the biggest databases in the drug area and consists of a user-friendly website. It includes drugs information, their various targets such as protein, carriers, transporters, and many other valuable data [24,25]. For academic purposes, DrugBank team frequently releases its raw data in an XML structure except for some data (https://www.drugbank.ca/releases/latest). However, there are some problems in exploiting the released file. First, the size of the data is too big, while some special software applications are needed. Second, the file is so difficult to understand and the data obtaining is a long process. Third, the structure of the obtained data is not consistent with the standard of SQL and cannot be imported to the SQL database management system (DBMS). Fourth, the properties or XML tags of drugs vary from a drug to another one. Fifth, high programming skill is needed for getting information on the flat file. Sixth, it is replete of duplicate data. Therefore, we developed the DrugR + database to resolve the aforementioned limitations, to benefit from the CDR capacity, and to provide a suitable database to be pragmatically used for drug repurposing. In other words, the DrugR + database has been developed based on the published file and designed constructively to be a sound and robust tool for drug discovery. Fig. 1 shows several steps used for developing the DrugR + database. The developments phases follow a Map-reduce base method. In the first step of development, the flat file is taken from the DrugBank and KEGG. In the second step, it is divided into small files. For every drug, a file is created using python programing language. For this purpose, the flat file is traversed line by line, and information of a drug is then stored in a separate file. In the third step, the small files are analyzed, and entities and their relationships are determined. In other words, conceptual modeling is done [26]. Based on the entities and their relationships, database normalization is done in the third normal form (3NF) level [27]. Finally, the normalized entities are implemented in SQL DBMS. For every entity, a table is created. For relationships, tables are made based on their cardinalities. Indices on the attributes, which are critical in the performance of searches, are also generated. The statistical summary of the current version of the DrugR+ is depicted in Fig. 2.

After database implementation, the small files are traversed, and their data are inserted in the relevant tables using a MAP-reduce parallel processing method [28]. With considering time-consuming and the database updating, parallel processing is employed. This enables DrugR + to be updated in the shortest time. The python programing language is used in the fifth phase. Finally, data controlling is done, and the database is uploaded to the server.

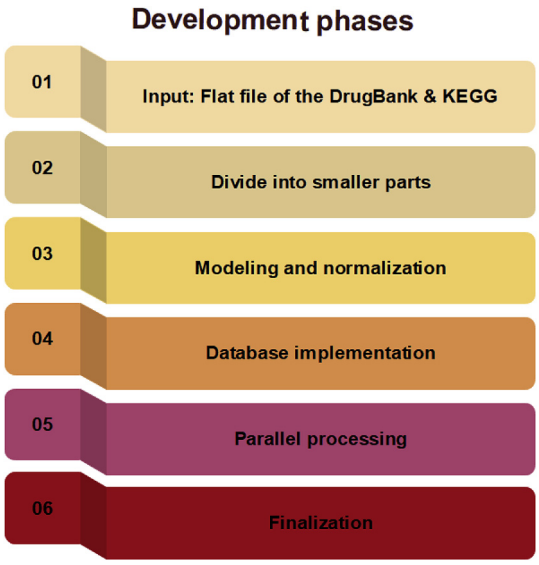

Fig. 1. Phases in developing DrugR+. In the first step, flat files are obtained from DrugBank and KEGG web interfaces. Then, the files are divided into smaller parts. Also, the files are analyzed, and entities and their relationships are determined. For enhancing operations speed, database normalization is done through the third normal form (3NF). After that, the acquired conceptual model is implemented. Then, the small files are processed using available processors, and their information is transferred to the related tables. Finally, the database is uploaded to the server.

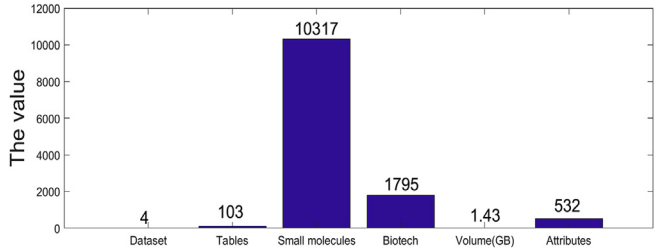

Fig. 2. Statistical summary of DrugR+: DrugR + includes 103 tables, 532 attributes of drugs including properties of drugs and their various targets, and 12112 drugs. The size of the current version of DrugR+ is 1430 MB. These values are related to the current version of the database and are continuously updated. Therefore, these values may be changed in future releases.

3.1. Datasets generating

In addition to the MSSQL database, four types of datasets have been mined for predicting of drugs and their targets because there are not proper datasets for acquiring a model using machine learning approaches [29] such as an artificial neural networks [30], a support vector machines [31], decision trees [32], or others [33]. For this purpose, we created four datasets which are based on chemical information of drugs and genomic information of targets. These targets, which have critical roles in live organisms, are categorized into four classes, including enzymes (ENZ), ion channel proteins (ICP), G protein-coupled receptors (GPCR), and nuclear receptor proteins (NRP). To obtain each of the mentioned datasets, the method based on known drug-target interactions is used. In Fig. 3, the pseudo code of the proposed method is presented.

The chemical information about drugs and genomic information of targets are obtained from KEGG and DrugR+, respectively. For calculating similarities between a drug and another drug, Eq. (1) and Eq. (2) are used [34]. In computing a similarity, the pharmaceutical effects of drugs on 17109 molecular properties, which is presented as an array like (F_1, F_2, …, F_{17109}), are considered. Also, Smith-waterman alignment algorithm is used for computing score between targets [35].

$$SIM(D, D') \frac{\sum_{i=1}^{n} w_i F_{iF_i'}}{\sqrt{\sum_{i=1}^{n} w_i F_i^2} \sqrt{\sum_{i=1}^{n} w_i F_i'^2}} \quad (1)$$

$$w_i = \exp(-d_i^2/(\sigma^2 h^2)) \quad (2)$$

Where, SIM, n, w_i, d_i, σ, and *h* are the similarity between two drugs such as D and D′, a number of molecular properties, the weight of the ith feature, the frequency of the ith feature, the standard deviation of the frequencies, and a constant value 0.1, respectively. In the Supplementary File, the results of the three types of learners on the generated datasets exist and have been compared against each other.

4. Web interface

The web interface of DrugR + consists of some main parts including SEARCH, CDR, DATASETS, and DB STRUCTURE. These parts are shown in Fig. 4.

4.1. Search strategies

The web interface of DrugR + provides an easy way of searching all of the existing data. In the search section, two strategies are available for both unprofessional and professional users, as follows:

(i) Unprofessional users can select their tables and then can confine search space using a proper way. They can also export the acquired results into an excel file.
(ii) Professional users can express their queries in the specified box. In this section, all of the SQL queries like *group by*, nested queries, and many other queries, which are relevant to information retrieval, can be stated. However, data manipulating operations such as *update* and *delete* are not permitted. Like the first search strategy, professional users can also export their favorable results into an excel file.

4.2. CDR

As we know, the current data comprise hidden information which can be obtained by data-mining methods. In addition to the professional and unprofessional search strategy, CDR, which searches existing data and mines latent information based IF-THEN rules, is another brilliant capability of the DrugR+. For an entered drug, this section suggests a list of several drugs if they are available. The proposed list may be used in three ways. One or more drugs of the suggested list may be nominated for combining or replacing with the selected drug in order to reach the repurposing of drugs, combination therapy, or CDR in the first, second, and third ways, respectively. In using this capability, a user can limit mining operations based on FDA-approved drugs, unapproved drugs, or both of them. Furthermore, a user must select target types which are divided into enzymes, carrier proteins, transporter proteins, and other types of proteins; and can confine search space for known-action determining a target which is the cause of a disease or not (e.g., side effects or unknown targets). Fig. 5 shows a simple example, which describes how the proposed lists are generated and how a proper structure of a database may lead to serviceable results. In some studies, a similar systematic method has been proposed for the repositioning of drugs [36,37]. For acquiring the list, the below steps are followed:

(i) For a given drug, information such as drug-disease, drug-target, mechanism of action of the drug on the targets, and main targets and off-targets are extracted.
(ii) The drugs, which have not common target with the entered drug, are filtered.

Generating the datasets for drug-target interactions

Dataset = [];
For every existing drug-target interaction such as (D, T):
    DR= (A set of drugs which their target set includes T) – D;
    TA= (A set of targets of D) – T;
    DR_SIM= An array which consists of the similarity effect score between D and DR;
    TA_SIM= An array which consists of the alignment score between T and TA;
    N= length (DR); D_MI= min (DR); D_MA= max (DR); D_AVG= mean (DR);
    M= length (TA); T_MI= min (TA); T_MA= max (TA); T_AVG= mean (TA);
    R= [N, D_MI, D_MAX, D_AVG, M, T_MI, T_MA, T_AVG];
    Dataset= Dataset U R;
Return Dataset

Fig. 3. Pseudo-code for generating the datasets. In generating the datasets, genomics information of targets and similarity information of drugs are considered. Among every pair of drugs, the similarity score is calculated and are stored in a matrix. Furthermore, smith-waterman alignment scores among every pair of targets are computed. Finally, some features shown in the pseudo-code are obtained, and the datasets are provided.

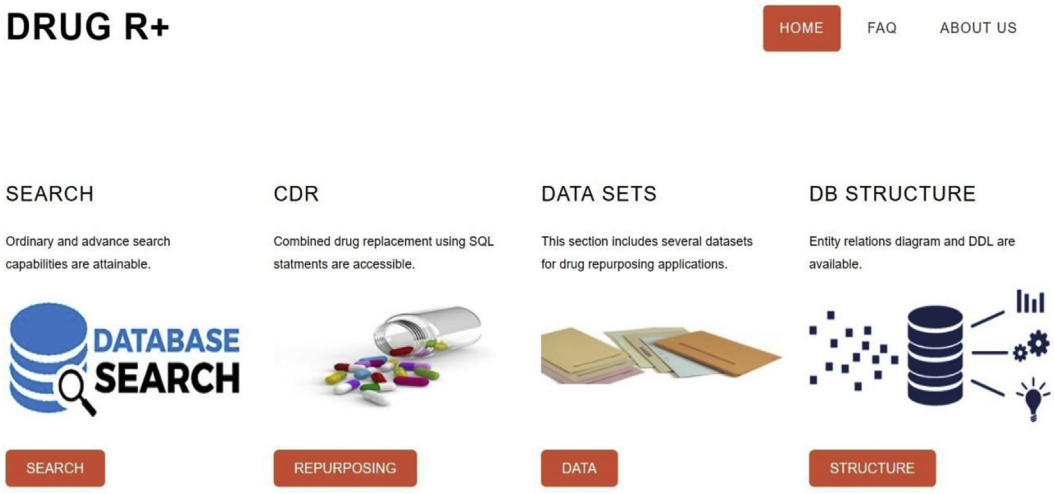

Fig. 4. The first page of the web interface of DrugR+. It consists of four main parts including search, combined drug repurposing (CDR), datasets, and database structures. For professional and unprofessional users, two different search strategies exist. In the CDR section, a list of proposed drugs is produced for an entered drug, and some URL links are generated from DrugR + to DrugBank. The dataset section includes four datasets which can be used for predicting drug-target interactions using machine learning approaches. Besides, entity relationships diagram (ERD), python codes, and SQL scripts are available in the database structure section.

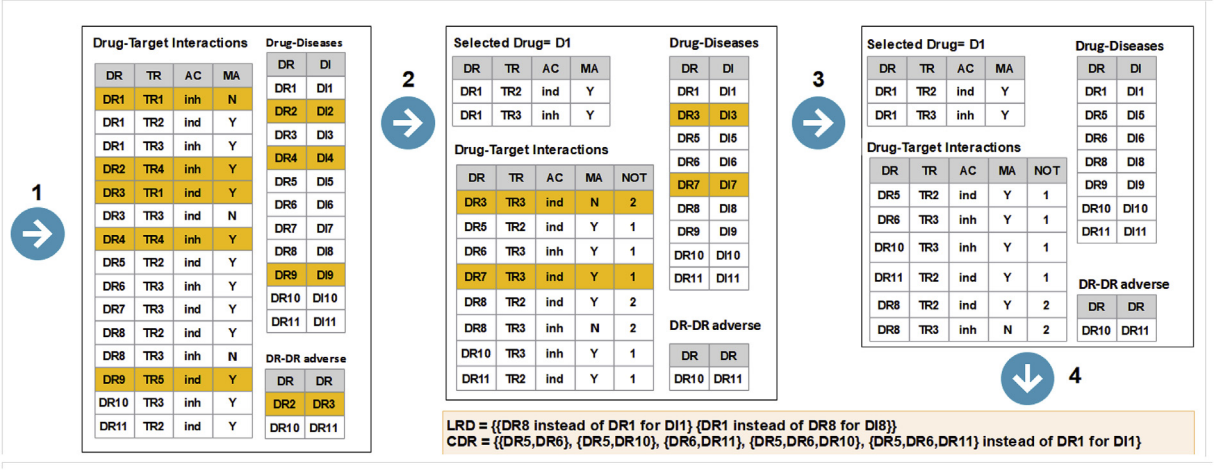

Fig. 5. A simple example of describing generated results. For a given drug (DR) such as DR1, this example explains how the results are produced. In the first step, four tables are joined together and information (like drug names, their targets (TR), the mechanism of action of drugs on targets (MC), and to be the main target of a drug (MA) or not) are extracted. Besides the mentioned information, drug-disease and drug-drug adverse reactions exist. Then, the data, which are not the main target of the selected drug or have not the common target with it, are omitted from three obtained tables. The colored rows are those which are eliminated in the next step. For instance, DR2 is eliminated because it has no common target with DR1. The total number of targets (NOT) including main targets and side effects are also calculated for each of the drugs. In the third step, the drugs, which their mechanism of actions on the targets is not identical with DR1, are disregarded. Next, the remaining drugs are ranked based on the number of side effects. For example, DR3 is deleted because it induces TR3 whereas DR1 inhibits that. Finally, some sets including a list of repurposed drugs (LRD) and CDR up to 3 members are proposed. Given the adverse reaction between DR10 and DR11, they are not considered as a set in the CDR list.

(iii) Mechanism of action (MC) of drugs on targets is investigated. Next, medications, whose MC differ from the MC of the entered drug, are ignored.
(iv) For all remaining drugs, the total number of side-effects are calculated. Then, they are sorted based on their side effects.
(v) The obtained list is probed for determining repurposed drugs in single and combination manners.
(vi) For all of the acquired combination lists, drug-drug adverse reactions (DDAR) are scrutinized, and the lists having DDAR are omitted.

4.3. Datasets

The web interface of the DrugR + incorporates the mined datasets, which have been categorized into four classes, including ENZ, ICP, GPCR, and NRP. For every category, a distinct text file, which contains known drug-target interactions, has been generated. Besides known drug-target interactions, another text file, which consists of unknown drug-target interactions, has been produced for every class of drug-target interactions. In obtaining the datasets, the pseudo code has been employed, as shown in Fig. 3. After creating the eight databases, handheld and automatic investigation of data have been done, and the data with missing values have been eliminated. Using known drug-target interactions, a model can be built and tested by machine learning methods. After that, the model is applied to unknown drug-target interaction dataset in order to predict interactions. The dataset section of DrugR+ is presented in Fig. 6.

4.4. Data access

For academic purposes, users can freely access the DrugR+ and get their favorable data using various manners like simple or nested queries, CDR, etc. In order to enter the web interface, there is no need for logging. In other words, all parts of the DrugR + are accessible when a user enters. Furthermore, the obtained results can be exported as an excel file. In the DrugR+, operations such as select, join, nested queries, data grouping, and every operation which does not update, delete or insert information are permitted. Also, we provide a username and password for expert users who want to connect to the DrugR + using MSSQL software. Like web interface users, these users can only express the operations which do not change database information.

4.5. Code and documentation

The DrugR+ is an open source dataset which has been developed in MSSQL server DBMS. Its web interface has also been developed in C# programing language based on ASP.net technology. In addition to the web interface, the structure of the database, data definition language (DDL) of DrugR+, python codes which break big data into small files and traverse them for inserting their information into the database, and the datasets are accessible through GitHub repository (https://github.com/LBBSoft/DrugR-plus) under MIT license. There is also a help document that describes how users can employ DrugR + database. Besides, the help file includes several examples of SQL statements.

4.6. Feedback and technical issues

The DrugR + has been tested and used by various users, and their feedbacks have been applied. However, it is possible that some bugs or errors occur in different situations. As pointed out in the FAQ section of the web interface, users can report these problems or their personal opinions using provided technical support emails. Frequently, the FAQ section will be updated based on the comments and questions received.

4.7. Updating and the hosting server properties

To update DrugR+, the necessary steps shown in Fig. 1 are followed. In other words, the previous version is archived, and the new version is created and uploaded to the server. To this end, a map-reduce parallel processing method is used. Hence, the time of generating a new database is minimized. Currently, the total time spent on generating and uploading the database based on DrugBank and KEGG data is less than an hour. Besides the database, the web interface of the DrugR+ is also regularly updated, and new capabilities are included. The database and its web interface have been placed on a server with 1 TB of hard disk, 128 GB of RAM, 20 CPU cores, and a windows operating system. This server can provide services to at least 50 users simultaneously.

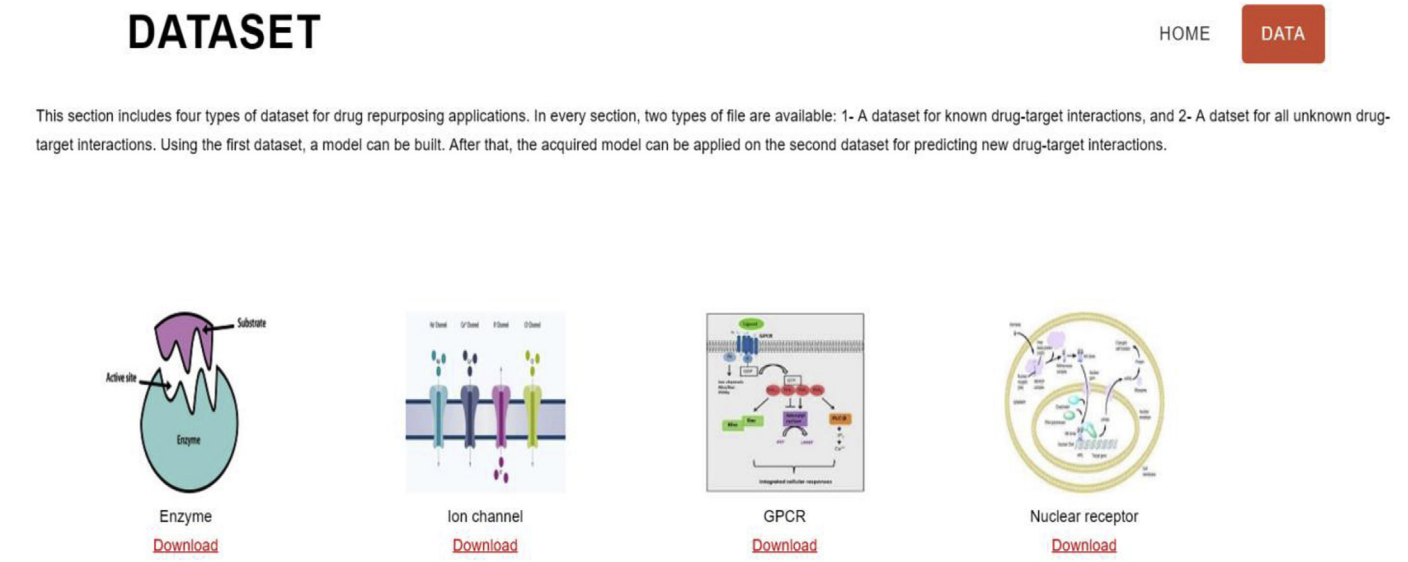

Fig. 6. The dataset section of DrugR+. The dataset section includes four datasets including drug and enzyme (Enzyme), drug and ion-channel (ION), drug and G-protein coupled receptor (GPCR), and drug and nuclear receptor protein interactions. In the supplementary file, the results of some known machine learning performance on the datasets have been reported.

Table 1
Acquired results for the five selected drugs.

Drug name	Therapeutic App	LRD	LCT	LCDRT
Fluoxetine	Antidepressant agent	fluvoxamine	(Fluoxetine + olanzapine), (Fluoxetine + diltiazem)	–
Temozolomide	Brain tumor	–	–	(oxaliplatin + mitotane)
Armodafinil	Wakefulness	diethylpropion	–	–
Phenmetrazine	Appetite depressant	Bupropion	–	(reboxetine + fencamfamine + armodafinil)
Degarelix	Prostate cancer	Ganirelix, teverelix	–	–

5. Utility and discussion

5.1. Drug repurposing using DrugR+

In this section, the DrugR + capabilities are shown. We also describe how a user can use it in drug repositioning applications. For repurposing drugs based on their targets, several stored procedures are available in the DrugR+. The stored procedures, which mine the database to propose a list of potential drugs for a given drug based on main, enzyme, carrier, and transporter targets, can confine search space by different parameters which are determined by a user. For example, users can restrict search space for relevant targets, off-targets, unknown targets, or all of them.

As an empirical example, we selected five popular and reputed drugs such as fluoxetine, temozolomide, armodafinil, phenmetrazine, and degarelix. Then, we proposed three lists of combination therapy (LCT), repurposed drugs (LRD), and CDR therapy (LCDRT) for each of them if they exist. In addition to the selected drugs' name; therapeutic application, LCT, LRD, and LCDRT are available in Table 1. In order to obtain the results, the web interface of DrugR + has been employed. The proposed lists were analyzed, and the refined lists were acquired.

5.2. Discussion

With regarding drug-target interactions, the DrugR + database proposes a list of drugs which may be used for LRD, LCT, and LCDRT. In the mining of the lists using various predefined stored procedures, several issues are considered, including drug-target actions, the mechanism of action of drugs on targets, and drug-drug inverse reactions. Furthermore, the predefined stored procedures can limit the search space using the parameters which are determined by a user. To examine the capabilities of the DrugR+, the five drugs selected based on their popularities and applications were investigated. The drugs and their results are described in five sections, as follows:

(i) Fluoxetine is an antidepressant drug which is placed in the SSRIs group of drugs for treating depression. Despite this issue, it has several side effects which may lead to other problems in a body [38]. Therefore, pharmacists designed a new drug called fluvoxamine instead [39]. As shown in Table .1, the DrugR + proposes that fluoxetine can be replaced by fluvoxamine. The acquired result shows the ability of the DrugR+ in finding other applications of drugs. For instance, we can use this capability to finding effective drugs for orphan or rare diseases [40]. For LCT of fluoxetine, DrugR + shows that its functionality can be enhanced when fluoxetine is combined with olanzapine. This is an atypical antipsychotic or diltiazem which is benzothiazepine derivative. In some studies, a successful combination of fluoxetine and olanzapine has been reported [41]. For LCDRT of fluoxetine, a proper combinational replacement list was not found.

(ii) Temozolomide is an alkylating agent, which is used in the treatment of cancerous brain tumors. For the LRD and LCT of temozolomide, the proposed list by the DrugR+ was investigated, while suitable alternative drugs for replacing or candidate drugs for combining with it were not found. Nevertheless, oxaliplatin and mitotane, whose mechanisms of actions are similar to temozolomide, were available in the proposed list by the DrugR+. Oxaliplatin and mitotane have applications in the treatment of colorectal cancer and adrenocortical cancer, respectively [42]. As shown in Table .1, these drugs are placed in the LCDRT section of temozolomide. By investigating studies, we can find some cases that discuss the multimodal abilities of the mentioned drugs in chemotherapy [43].

(iii) Like fluoxetine and temozolomide, armodafinil used to treat excessive daytime sleeping is an FDA approved drug. For the LRD of armodafinil, the DrugR + suggests diethylpropion that is an appetite suppressant drug. In other words, we propose that diethylpropion can be used for treating excessive daytime sleeping instead of armodafinil. As a side effect, it has been reported that diethylpropion increases wakefulness in several studies [44]. For the LCT and LCDRT of armodafinil, reasonable combinations were not found.

(iv) Phenmetrazine is an FDA approved and appetite depressant drug. As an alternative option, the DrugR + offers that it can be replaced by bupropion which is used for ceasing smoking. By investigating the proposed list for phenmetrazine, we did not find a suitable drug which could be combined with it. Unlike the LCT for phenmetrazine, the combination of reboxetine, fencamfamine, and armodafinil seems to be a proper replacement, in large part because of inhibiting its targets and identically affecting its binding sites. Although fencamfamine enhances dependency and abuse, reboxetine can reduce this anomaly. Furthermore, the selected drugs have a minimum number of side effects, and the effects of the drugs on side effects will also be minimized since the least amount of them is used [45].

(v) Degarelix has been designed and developed for treating advanced prostate cancer. Like the four other drugs, the DrugR + suggests a list of drugs with considering the mechanism of action of degarelix. By inspecting the proposed list, we can reach to two drugs named galinerix and teverelix which are used for controlling ovulation and benign prostate cancer, respectively [46]. While ganilerix is an FDA approved drug, teverelix is in phase II clinical trials. In some studies, ganilerix applications in prostate cancer have been expressed [47]. For the LCT and LCDRT of degarelix, appropriate and reasonable list of drugs was not acquired.

5.3. DrugR+ in comparison with other databases

DrugBank and KEGG, which include the most valuable data for drug applications, provides an appropriate platform for searching data and integrated them from various resources like PubChem. By investigating researches, it is observed that they have been used in most of them, and researchers have acquired new achievements based on their data. Every database has its own structure which makes it suitable for some applications. Based on the usage, the databases contain capabilities such as API, web interface, etc. The DrugR + has several key distinctions in comparison with other databases, as follows:

(i) It incorporates published data of DrugBank and KEGG, while there are some other data such as the chemical structure of drugs, which exist in DrugBank and KEGG, and are not available in the DrugR+. To deal with this, some navigating links have been provided from

the DrugR + to the DrugBank. Furthermore, the DrugR + includes four datasets which can be applied to predict drug-target interactions by machine learning approaches.

(ii) It is a relational database which follows mathematics concepts of relational algebra. Therefore, the DrugR+'s users can express their SQL queries and join their tables in various states. Although it seems that only expert users can exploit the DrugR+ in their projects, there are some other capabilities for inexpert users. For this purpose, the users only select a table and get its information. Then, they can limit the obtained results by clicking on the provided filter buttons easily. For expert users, the DrugBank has been created an API which needs a commercial license.

(iii) DrugBank presents all of the information of an entered drug (e.g., targets, descriptions, manufactures, and various information of a drug) or target, whereas the DrugR + database does not support this capability because of its organized structure and displays a specific part of the information (e.g., only targets of a drug or only manufactures of drugs). Despite this, users can acquire results which are not supported by the DrugBank. For example, a user can obtain a list of all external drug IDs to use them in their projects.

(iv) The DrugR + provides LDR and CDR capabilities which are not available in the DrugBank. These systematic capabilities, which generate the results based on several predefined and embedded stored procedures, exist because of proper organizing of data in the DrugR+.

Although the DrugR + has some features that others lack them, there are some limitations, which constitute future directions, as follows:

(i) It offers a list of drugs using several stored procedures based on drug-target interactions, drug-drug adverse reactions, and mechanism of actions of drugs on the targets. Then, it ranks the list based on the side effects of drugs. Despite this, the proposed list needs to be enriched while another improved list, which might be much more helpful to users, need to be suggested. In creating the list, the optimization algorithms can be useful because obtaining an optimal list of suggested drugs is an NP-hard problem [48].

(ii) It is only based on the DrugBank and KEGG data except for the created dataset. Although DrugBank and KEGG are comprehensive databases and are updated frequently, the other databases include some other useful information which can enhance the success rate of repurposing of drugs.

(iii) Several popular machine learning methods like a support vector machine, an artificial neural network, and a decision tree have been applied to the generated datasets. Further, their results have been reported in the Supplementary File. Hence, the state of the art machine learning methods such as transfer learning, deep learning, and other novel methods can lead to better results.

In addition to the mentioned differences, as shown in Table 2, there exist capabilities of databases from the repurposing applications point of view. The first column from the left side, shows databases such as CheMBL (CHM) [49], DrugBank (DrugB), therapeutic target database (TTD) [50], PubChem (PCHM) [51], and DrugR+. Due to their brilliant role in the drug repurposing applications, they have been chosen. All of the databases consist of some abilities, including being freely available and updated, consisting of advanced search, and possessing classified data. Unlike DrugBank which some of its data is not fully downloadable, the other databases publish their data completely. From providing repurposing tool and expressing standard query language (SQL) aspects, the DrugR + seems to be more suitable than the other databases. Like TTD, the DrugR + does not have web API. Instead, it has supplied read-only access to the database for all of the users. Thus, their programs can exploit this capability and obtain the results in a tabular manner. After searching and acquiring results, they can be downloaded from CHM, PCHM, and DrugR+, whereas DrugB and TTD do not include this capability.

6. Conclusion

DrugR+ is a well-organized database that provides repositioning capabilities such as individual and combined drug repurposing. Nested queries, data exporting, and various search methods are also available in this database. For professional and unprofessional users, this database empowers the users to obtain desired results. In designing and developing of the DrugR+, speed up the operation, different aspects have been considered, including database normalization, indexing, and other technical issues. Furthermore, it includes some optimized stored procedures which are applied to obtain a potential list of drugs. It also consists of four datasets which can be used for predicting drug-target interactions using machine learning methods. In order to propose the list, the DrugR + examines drug-target interactions, their mechanism of action, the adverse reaction between drugs, and their side effects. After acquiring a list of drugs, users can analyze the list and introduce some drugs for replacing a drug with another one, combination therapy, and the combination of drugs instead of a drug. To present the proof-of-technology by using the DrugR + database in terms of drugs repurposing, we selected five drugs which are used for treating various diseases. For the proposed lists, we successfully made some analysis, and selected drugs for repurposing goals.

Declaration conflict of interests

Authors have no competing interests to declare.

Ethics approval and consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

DrugR + has been implemented in C# programing language with ASP.net technology and is available at (https://github.com/LBBSoft/DrugR-plus).

Table 2
A comparison of databases' capabilities.

DB	Available	Update	Tool	Advanced search	Query (SQL)	Export	API	Full download	Classified structure	Free
CHM	✓	✓		✓		✓	✓	✓	✓	✓
DrugB	✓	✓		✓			✓		✓	✓
TTD	✓	✓		✓				✓	✓	✓
PCHM	✓	✓		✓		✓	✓	✓	✓	✓
DrugR+	✓	✓	✓	✓	✓	✓		✓	✓	✓

CHM: CheMBL; DrugB: Drug-Bank; TTD: Therapeutic Target database; PCHM: PubChem; DB: Database.

Competing interests

Not applicable.

Disclaimer

Not applicable.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contribution

YMS: Conceptualization, database development, formal analysis, investigation, writing-manuscript. **YO**: Database testing, validation, Conceptualization, writing-manuscript. **MA**: Software testing, validation, Editing-manuscript. **AMN**: Conceptualization, Supervision, Project administration, Editing the manuscript. All authors have read and approved the manuscript.

Acknowledgments

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.05.006.

References

[1] P. Smirnov, V. Kofia, A. Maru, M. Freeman, C. Ho, N. El-Hachem, G.-A. Adam, W. Ba-alawi, Z. Safikhani, B. Haibe-Kains, PharmacoDB: an integrative database for mining in vitro anticancer drug screening studies, Nucleic Acids Res. 46 (2017) D994–D1002.
[2] T.L. Blundell, B.L. Sibanda, R.W. Montalvão, S. Brewerton, V. Chelliah, C.L. Worth, N.J. Harmer, O. Davies, D. Burke, Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery, Philos. Trans. R. Soc. Lond. B Biol. Sci. 361 (2006) 413–423.
[3] J.A. Goldstein, L.A. Bastarache, J.C. Denny, J.M. Pulley, D.M. Aronoff, Preg OMICS—leveraging systems biology and bioinformatics for drug repurposing in maternal-child health, Am. J. Reprod. Immunol. (2018) e12971.
[4] Y. Donner, S. Kazmierczak, K. Fortney, Drug repurposing using deep embeddings of gene expression profiles, Mol. Pharm. 15 (2018) 4314–4325.
[5] N. Dovrolis, G. Kolios, G. Spyrou, I. Maroulakou, Laying in silico pipelines for drug repositioning: a paradigm in ensemble analysis for neurodegenerative diseases, Drug Discov. Today 22 (2017) 805–813.
[6] A. Talevi, Drug repositioning: current approaches and their implications in the precision medicine era, Expert Review of Precision Medicine and Drug Development 3 (2018) 49–61.
[7] M. Reichman, P.B. Simpson, Open innovation in early drug discovery: roadmaps and roadblocks, Drug Discov. Today 21 (2016) 779–788.
[8] W. Sun, P.E. Sanderson, W. Zheng, Drug combination therapy increases successful drug repositioning, Drug Discov. Today 21 (2016) 1189–1195.
[9] J.E. Lee, J.H. Lim, Y.K. Hong, S.H. Yang, High-dose metformin plus temozolomide shows increased anti-tumor effects in glioblastoma in vitro and in vivo compared with monotherapy, Cancer Research and Treatment: Off. J. Kor. Cancer Assoc. 50 (2018) 1331.
[10] N.M. Maruthur, E. Tseng, S. Hutfless, L.M. Wilson, C. Suarez-Cuervo, Z. Berger, Y. Chu, E. Iyoha, J.B. Segal, S. Bolen, Diabetes medications as monotherapy or metformin-based combination therapy for type 2 diabetes: a systematic review and meta-analysis, Ann. Intern. Med. 164 (2016) 740–751.
[11] D.R. Vogus, V. Krishnan, S. Mitragotri, A review on engineering polymer drug conjugates to improve combination chemotherapy, Curr. Opin. Colloid Interface Sci. 31 (2017) 75–85.
[12] A.S. Brown, C.J. Patel, A review of validation strategies for computational drug repositioning, Briefings Bioinf. 19 (2016) 174–177.
[13] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2017) D1074–D1082.
[14] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs, Nucleic Acids Res. 45 (2016) D353–D361.
[15] A.S. Brown, C.J. Patel, A standard database for drug repositioning, Scientific data 4 (2017) 170029.
[16] N. Arora, P.K. Maurya, P. Kacker, Translational Research in Drug Discovery and Development, Translational Bioinformatics and its Application, Springer, 2017, pp. 55–87.
[17] S.M. Corsello, J.A. Bittker, Z. Liu, J. Gould, P. McCarren, J.E. Hirschman, S.E. Johnston, A. Vrcic, B. Wong, M. Khan, The drug repurposing hub: a next-generation drug library and information resource, Nat. Med. 23 (2017) 405.
[18] H. Yang, C. Qin, Y.H. Li, L. Tao, J. Zhou, C.Y. Yu, F. Xu, Z. Chen, F. Zhu, Y.Z. Chen, Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information, Nucleic Acids Res. 44 (2015) D1069–D1074.
[19] A. Gottlieb, G.Y. Stein, E. Ruppin, R. Sharan, PREDICT: a method for inferring novel drug indications with application to personalized medicine, Mol. Syst. Biol. 7 (2011) 496.
[20] K. Shameer, B.S. Glicksberg, R. Hodos, K.W. Johnson, M.A. Badgeley, B. Readhead, M.S. Tomlinson, T. O'Connor, R. Miotto, B.A. Kidd, Systematic analyses of drugs and disease indications in RepurposeDB reveal pharmacological, biological and epidemiological factors influencing drug repositioning, Briefings Bioinf. 17 (2017) 656–676.
[21] H. Huang, T. Nguyen, S. Ibrahim, S. Shantharam, Z. Yue, J.Y. Chen, DMAP: a Connectivity Map Database to Enable Identification of Novel Drug Repositioning Candidates, BMC Bioinformatics, BioMed Central, 2015, p. S4.
[22] H. Wu, J. Huang, Y. Zhong, Q. Huang, DrugSig: a resource for computational drug repositioning utilizing gene expression signatures, PLoS One 12 (2017) e0177743.
[23] R. Holland, Drug discovery world, drug discovery and development news, Drug Discov. World 17 (2016) 9–14.
[24] R. Ferdousi, R. Safdari, Y. Omidi, Computational prediction of drug-drug interactions based on drugs functional similarities, J. Biomed. Inform. 70 (2017) 54–64.
[25] R. Safdari, R. Ferdousi, K. Aziziheris, S.R. Niakan-Kalhori, Y. Omidi, Computerized techniques pave the way for drug-drug interaction prediction and interpretation, Bioimpacts: BI 6 (2016) 71.
[26] V.C. Storey, I.-Y. Song, Big data technologies and management: what conceptual modeling can do, Data Knowl. Eng. 108 (2017) 50–67.
[27] N. Golov, L. Rönnbäck, Big Data normalization for massively parallel processing databases, Comput. Stand. Interfac. 54 (2017) 86–93.
[28] B. Cai, L. Li, Z. Xiang, W. Xue, B. Yang, Transmission of Map/Reduce Data in a Data Center, Google Patents, 2017.
[29] Y. Masoudi-Sobhanzadeh, H. Motieghader, A. Masoudi-Nejad, FeatureSelect: a software for feature selection based on machine learning approaches, BMC Bioinf. 20 (2019) 170.
[30] G. Huang, G.-B. Huang, S. Song, K. You, Trends in extreme learning machines: a review, Neural Network. 61 (2015) 32–48.
[31] J. Díaz, J. Acosta, R. González, J. Cota, E. Sifuentes, À. Nebot, Modeling the control of the central nervous system over the cardiovascular system using support vector machines, Comput. Biol. Med. 93 (2018) 75–83.
[32] V. Shalaeva, S. Alkhoury, J. Marinescu, C. Amblard, G. Bisson, Multi-operator Decision Trees for Explainable Time-Series Classification, International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 2018, pp. 86–99.
[33] K.Y. Michael, J. Ma, J. Fisher, J.F. Kreisberg, B.J. Raphael, T. Ideker, Visible machine learning for biomedicine, Cell 173 (2018) 1562–1565.
[34] H. Iwata, R. Sawada, S. Mizutani, Y. Yamanishi, Systematic drug repositioning for a wide range of diseases with integrative analyses of phenotypic and molecular data, J. Chem. Inf. Model. 55 (2015) 446–459.
[35] S.K. Zahid, L. Hasan, A.A. Khan, S. Ullah, A novel structure of the smith-waterman algorithm for efficient sequence alignment, digital information, networking, and wireless communications (DINWC), 2015 Third International Conference on, IEEE, 2015, pp. 6–9.
[36] Y. Pan, T. Cheng, Y. Wang, S.H. Bryant, Pathway analysis for drug repositioning based on public database mining, J. Chem. Inf. Model. 54 (2014) 407–418.
[37] R. Sawada, H. Iwata, S. Mizutani, Y. Yamanishi, Target-based drug repositioning using large-scale chemical–protein interactome data, J. Chem. Inf. Model. 55 (2015) 2717–2730.
[38] A.D. Kaye, Tramadol, pharmacology, side effects, and serotonin syndrome: a review, Pain Physician 18 (2015) 395–400.
[39] P. Blier, Neurobiology of depression and mechanism of action of depression treatments, J. Clin. Psychiatry 77 (2016) e319.
[40] C.P. Austin, C.M. Cutillo, L.P. Lau, A.H. Jonker, A. Rath, D. Julkowska, D. Thomson, S.F. Terry, B. de Montleau, D. Ardigò, Future of rare diseases research 2017–2027: an IRDiRC perspective, Clin. Transl. Sci. 11 (2018) 21–27.
[41] H.C. Detke, M.P. DelBello, J. Landry, R.W. Usher, Olanzapine/Fluoxetine combination in children and adolescents with bipolar I depression: a randomized, double-blind, placebo-controlled trial, J. Am. Acad. Child Adolesc. Psychiatry 54 (2015) 217–224.
[42] Y. Yamamoto, N. Sasaki, K. Kumagai, S. Takeuchi, T. Toyooka, N. Otani, K. Wada, Y. Narita, K. Ichimura, H. Namba, Involvement of Intracellular Cholesterol in Temozolomide-Induced Glioblastoma Cell Death, Neurologia medico-chirurgica, 2018 ra. 2018-0040.
[43] C.A. Vargo, L.A. Ray, H.B. Newton, Neurological Complications of Chemotherapy, Cancer Neurology in Clinical Practice, Springer, 2018, pp. 275–310.
[44] S.-C. Chang, W.W. Shen, Stimulants, wakefulness-promoting agents, and non-stimulant attention deficit hyperactivity disorder medications, J. Exp. Clin. Med. 5 (2013) 210–216.
[45] H. Greely, B. Sahakian, J. Harris, R.C. Kessler, M. Gazzaniga, P. Campbell, M.J. Farah, Towards responsible use of cognitive-enhancing drugs by the healthy,

Nature 456 (2008) 702.
[46] S. Sperduti, L. Riccetti, C. Lazzaretti, M. Simoni, L. Casarini, In vitro characterization and comparison of commercial GnRH antagonists, 20th European Congress of Endocrinology, BioScientifica, 2018.
[47] L. Boccon-Gibod, P. Iversen, B.-E. Persson, Degarelix 240/80 mg: a new treatment option for patients with advanced prostate cancer, Expert Rev. Anticancer Ther. 9 (2009) 1737–1743.
[48] Y. Masoudi-Sobhanzadeh, H. Motieghader, World Competitive Contests (WCC) algorithm: a novel intelligent optimization algorithm for biological and non-biological problems, Informatics in Medicine Unlocked 3 (2016) 15–28.
[49] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, ChEMBL: a large-scale bioactivity database for drug discovery, Nucleic Acids Res. 40 (2011) D1100–D1107.
[50] X. Chen, Z.L. Ji, Y.Z. Chen, TTD: therapeutic target database, Nucleic Acids Res. 30 (2002) 412–415.
[51] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, PubChem substance and compound databases, Nucleic Acids Res. 44 (2015) D1202–D1213.