We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Quantitative Methods in System-Based Drug Discovery

Yu-Chen Lo , Ren Gui , Hiroshi Honda and
Jorge Z. Torres

Additional information is available at the end of the chapter

Abstract

Modern pharmaceutical industries have faced significant challenges to deliver safe and effective medicines because of significant toxicity and severe side effects of discovered drugs. On the other hand, recent developments and advances in system-based pharmacology aim to address these challenges. In this chapter, we provide an overview of quantitative methods for system-based drug discovery. System-based drug discovery integrates chemical, molecular, and systematic information and applies this knowledge to the designing of small molecules with controlled toxicity and minimized side effects. First, we discuss current approaches for drug discovery and outline their advantages and disadvantages. Next, we introduce basic concepts of systems pharmacology with an emphasis on ligand-based drug discovery and target identification. This is followed by a discussion on structure-based drug design and statistical tools for pharmaceutical research. Finally, we provide an overview of future directions in systems pharmacology that will guide further developments.

Keywords: systems pharmacology, drug discovery, chemoinformatics

1. Introduction

The discovery of effective medicines has been a long-term human endeavor aimed at curing illnesses and improving physiological conditions. Early drug discovery methods often relied on serendipitous findings. For example, penicillin, a substance released by mold, was accidentally discovered to inhibit bacterial growth. However, with the advances of molecular cloning techniques, X-ray crystallography, robotics, and computational aided technology, drugs can now be rationally designed. Additionally, the marriage of combinatorial chemistry and robotics has also created a new drug discovery approach called "high-throughput

screening." In a high-throughput screening campaign, a library consisting of millions of molecules is tested against a disease-relevant target to identify potential drug candidates.

However, the application of modern drug discovery methods has not directly translated into increases in the output of new drugs. Although many of these "designer" drugs have been optimized for binding and specificity, they are all low-weight molecular ligands and are likely to interact with multiple off-targets, which contribute to severe side effects in patients taking these medicines. Consequently, a system drug discovery approach that simultaneously optimizes drug binding, target promiscuity, and safety profile has been proposed. In particular, "poly-pharmacology" is a new drug discovery paradigm that aims to study the interactions of many drugs to many targets as many body problems. The binding profile of a compound can be used to modify a ligand structure to maximize on-target binding while minimizing off-target interactions. More recently, "structural poly-pharmacology" has also been proposed, which utilizes the structural data of proteins or drugs to gain a mechanistic understanding of drug action and side effects [1, 2].

System-based drug discovery can be classified as ligand-based and structure-based drug discovery approaches. In the ligand-based drug discovery approach, ligand structure information takes center stage, and only ligand information is used to derive its multitude of chemical and biological properties. On the other hand, the structure-based approach utilizes the structure of the receptor to identify shape-complementary ligands with optimal interactions. Given a validated disease target with a known crystal structure, the structure-based approach can be utilized to discover ligands that bind to the receptor of interest. On the other hand, the ligand-based approach is useful when the target structure is unknown or when the crystal structure of the target is difficult to obtain. In some cases, the system-based drug discovery approach combines both ligand-based and structure-based approaches to facilitate the drug design and discovery process [3].

2. Ligand-based approaches to drug design

2.1. Introduction

Ligand-based drug design, also known as "knowledge-based" drug design, extracts essential chemical features from drugs to construct a learning model to predict drug properties. It has been proposed that a ligand structure contains all the necessary information to accurately infer its mechanism of action. Several chemical descriptors such as molecular weight and lipophilicity are predictors of important pharmacodynamic and pharmacokinetic properties of ligands. A well-known example is Lipinski's rule of five, which describes a set of chemical properties and rules that can be used to differentiate drugs from nondrug molecules [4]. Likewise, the chemical similarity principle has been widely applied in similarity-based drug design. The chemical similarity principle assumes that if two molecules share similar structures, then they will likely have similar biological properties [5]. This concept is the underlying principle of modern chemical database search techniques used to identify similar compounds with improved bioactivities.

2.2. Outline of ligand-based drug design

The chemical similarity search is an established approach for ligand-based drug discovery. Given a compound with known biological properties, it is possible for a drug designer to identify similar compounds with improved biological properties. To compare two ligand structures, it is essential to develop appropriate data structure representations for chemical structure comparisons. Mathematically, the chemical structure can be represented by graphs where atoms represent vertices and edges represent chemical bonds [6]. Several chemoinformatics algorithm can be used to extract essential characteristic from the chemical graph such as the number of vertices, the number of bonds, path, connectivity, and others. These properties then become chemical features that can be used in feature engineering using complex machine learning techniques for similarity comparison.

The most direct chemical search approach is the nearest neighbor where the chemical feature of a ligand, also known as a "chemical fingerprint," is used to search a compound database to identify the most similar compounds using a predefined distance measure [7]. The most commonly used chemical fingerprints include path-based and substructure-based fingerprints. Using path-based fingerprints, such as Daylight fingerprints or Obabel FP2 fingerprints, potential paths at different bond lengths in a molecular graph of a molecule are used as features for the similarity comparison. On the other hand, substructure-based fingerprints such as MACCS keys use predefined substructures and characterize each molecule based on the presence or absence of a particular substructure using a binary array. Overall, path-based fingerprints offer higher search specificity due to the unique path dependency of the molecular graph. However, substructure-based fingerprints can be used to identify scaffold hopping ligands since the fingerprints do not impose requirements on the connectivity of the scaffolds and functional groups [8]. To quantify the chemical similarity between two fingerprints, several distance metrics have been proposed. The most commonly used chemical similarity metric is the Tanimoto index, which computes the shared feature bits between two fingerprints in the range of 0–1. Although there is no predefined threshold to define the similarity level, a value of 0.7–0.8 has been commonly adapted in many chemical similarity search programs.

2.3. Ligand-based drug design (LBDD) process

The similarity-based drug design process is as follows (**Figure 1**):

1. The target molecule is used as a query for the chemical search.

2. Similar ligands with similar biological properties are identified.

3. Original ligands are modified to suggest new molecules with improved activities.

The advantages of LBDD are as follows:

1. Does not necessitate receptor structures.

2. Low computational intensity and fast database searching.

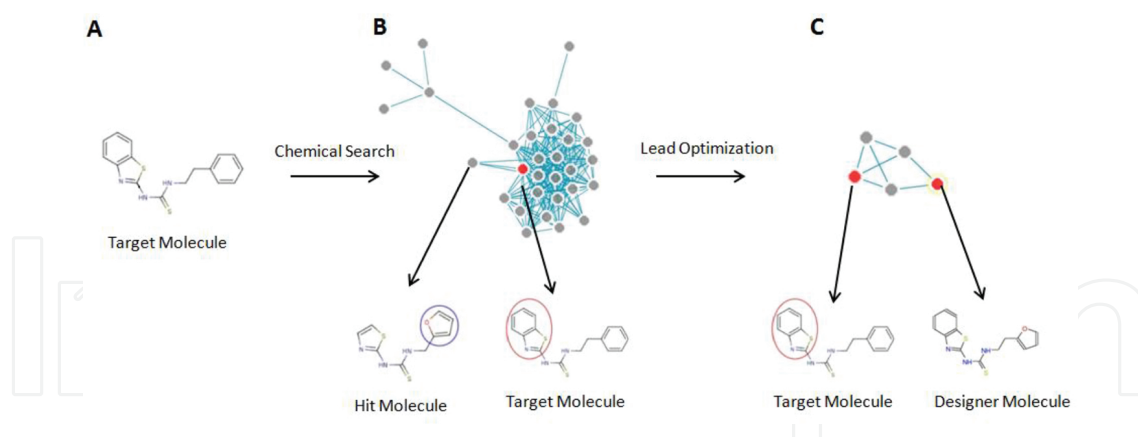3. Allows for large-scale similarity drug design and target prediction.

Figure 1. Ligand-based drug design (LBDD) process.

2.4. Target prediction of drugs

While similarity-based methods are prevalently applied in modern drug discovery programs as an efficient way to transition a hit to a lead, the molecular mechanism of the drug is often unknown, and the adverse reaction cannot be predicted. Consequently, drug-target prediction becomes an important follow-up step. Drug-target prediction can be classified as ligand-based or structure-based methods [9]. In ligand-based target prediction, the molecular target of a drug can be inferred from the target-annotated ligand sharing the highest chemical similarity. Chemical bioactivity databases such as ChEMBL, PubChem and DrugBank, and BindingDB have been developed for this application [10]. However, one major limitation of ligand-based target prediction is that there is no natural cutoff for chemical similarity that clearly defines biological similarity, also known as bioactivity cliffs. Approaches such as similarity ensemble approach (SEA) aim to remedy this by calculating similarity values against a random background using an algorithm similar to BLAST [11]. On the other hand, structure-based target prediction methods identify molecular targets based on the structure of the receptor binding sites. For example, panel docking is a common structure-based approach to identify the most probable target based on the docking score. Alternatively, binding site similarity methods that compare the receptor environment of the target ligand to a database of receptor pockets have also proven to be an effective target prediction approach [12].

More recently, network poly-pharmacology has been proposed as a more comprehensive approach to analyze drug-target interactions. Network pharmacology goes beyond the one drug one target hypothesis to multiple drugs multiple targets hypothesis [13]. The goal of this new paradigm is not only to accurately identify on-targets but also other off-targets. One such approach is the drug-target network, which utilizes a bipartite network to analyze complex drug gene interactions. Alternatively, drug-drug networks or chemical similarity networks have also been proposed [8]. The chemical similarity network clusters drugs based on their structure similarity. This approach can be applied for large-scale compound analysis by clustering diverse chemical structures into distinct scaffolds known as chemotypes. Consequently, each chemotype can be correlated with specific molecular targets. Using a consensus statistics scheme similar to that used for functional prediction in protein-protein interaction

networks, chemical similarity networks have proven useful for target identification from chemical screening campaigns. In addition, structural poly-pharmacology has also gained substantial attention due to the possibility of correlating structural variations to clinical side effects. One example is CSNAP3D, which uses 3D ligand structure similarity to identify simplified scaffold hopping compounds of complex natural products to suggest new drugs with improved pharmacokinetic properties [1].

2.5. Side effect prediction of drugs

Since severe clinical side effects have contributed to drug failure in the late stages of clinical trials, side effect prediction will need to become an integrated part of the drug design process. Knowing the binding affinity of a drug to an array of proteins makes it possible to predict side effects using several statistical methods such as canonical component analysis (CCA), which identify a set of parameters that optimize the correlation between drug binding features and side effect features [14]. In addition, side effect predictions based on chemoinformatics analysis of the compound structures have also been developed. Many of these approaches provide an accurate prediction of the drug side effects and have been applied in the early stages of clinical drug development.

3. Structure-based approaches to drug design

Modern medicinal chemistry methods and molecular modeling have been employed as powerful tools for the study of structure-activity relationships (SAR) [15]. Structure-based drug design is a drug discovery approach by which synthetic compounds are designed from detailed structural knowledge of the active sites of protein targets associated with particular diseases. This field has involved the integrated application of traditional biology and medicinal chemistry along with advances in biomolecular spectroscopic methods such as X-ray crystal-lography, nuclear magnetic resonance (NMR), combinatorial chemistry, computer modeling of molecular structure, and protein biophysical chemistry, to focus on the three-dimensional molecular structure and active site characterization of the proteins that control cellular biology. Structure-based drug design is an improvement over traditional drug screening techniques. By identifying the target protein in advance and by discovering the chemical and molecular structure of the protein, it is possible to design a more optimal drug to interact with the protein.

3.1. Basic concept of structure-based drug design

Enzymes are a subset of receptor-like proteins that are directly responsible for catalyzing the biochemical reactions that sustain life. For example, digestive enzymes act to break down the nutrients of our diet. DNA polymerase and related enzymes are crucial for cell division and replication. Enzymes are genetically programmed to be specific for their appropriate molecular targets. Any errors could have grave consequences. Enzymes ensure the specificity of their targets by forming a molecular environment that excludes interactions with inappropriate molecules. The analogy most often mentioned is that of a lock and key. The enzyme is a

molecular lock, which contains a keyhole that exhibits a very specific and consistent size and shape. This molecular keyhole is termed the active site of the enzyme and allows interaction with only the appropriate molecular targets. Just as a typical lock is much bigger than the keyhole, the receptor is usually much larger than the active site. The receptor, as specified by our DNA, is a folded protein whose major purpose is to form and maintain the size and shape of the active site.

The most important concept in drug design is to understand the methods by which the active site of the receptor selectivity restricts the binding of inappropriate structures. Any potential molecules that can bind to a receptor are called ligands. In order for a ligand to bind, it must contain a specific combination of atoms that present the correct size, shape, and charge composition needed to bind and interact with the receptor. In brief, the ligand must possess the molecular key that binds the receptor lock.

Computer-aided drug design has played an important role in drug discovery and drug development and has become an indispensable tool in the drug industry. For the purposes of discovering and optimizing biologically active compounds, various types of computer-aided drug design software and resources have been used by computational medicinal chemists. Unsurprisingly, many chemical compounds were discovered and optimized by computer-aided drug design methodologies and have reached the late stages of clinical trials.

3.2. Outline of structure-based drug design

Structure-based drug design is a cyclic process which consists of stepwise procedures (**Figure 2**). It begins with a known target structure, and then, *in silico* studies are conducted in order to identify potential ligands. These molecular modeling procedures are followed by the synthesis of the most promising compounds [16]. Next, using diverse experimental platforms, biological properties such as potency, affinity, and efficacy are evaluated [17]. In the end, given that active compounds are identified, the three-dimensional structure of the ligand-receptor complex can be solved. The available structure allows the visualization of the intermolecular features that support the process of molecular recognition. Structural descriptions of ligand-receptor complexes are useful for the investigation of binding conformations, characterization of key intermolecular interactions, characterization of unknown binding sites, mechanistic studies, and the elucidation of ligand-induced confor-mational changes [18].

3.3. Process of structure-based drug design (SBDD)

The structure-based drug design process is as follows:

1. An enzyme that is important in a particular pathological condition is chosen.

2. The three-dimensional structure of the active site of the enzyme is determined, often by X-ray crystallography.

3. A chemical is prepared to fit the active site of the enzyme, which can alter the properties of the enzyme, that is, inactivate the enzyme.

The advantages of structure-based drug design are as follows:

1. Useful results are obtained faster than by traditional drug design methods.

2. The process is less expensive than other drug design methods.

3. The compounds are more specific for the active site and potentially less toxic than compounds prepared by other approaches.
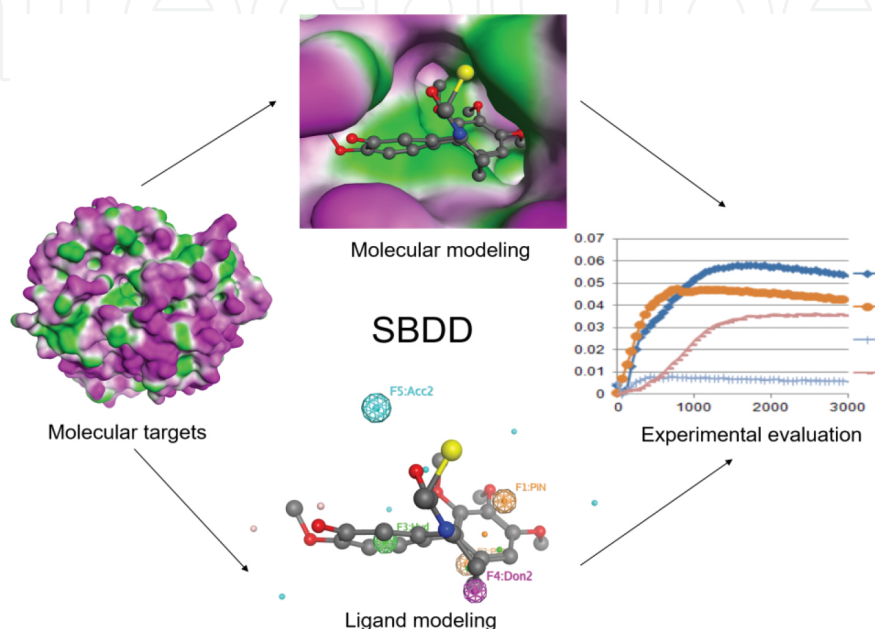

Figure 2. Structure-based drug design (SBDD) process.

3.4. Synthesis of lead compound

The initial drug design phase is followed by the synthesis of the lead compound, quantitative measurements of its ability to interact with the target protein, and X-ray crystallographic analysis of the compound-target complex. This analysis reveals important, empirical information on how the compound actually binds to the target, and the nature and extent of changes induced in the target by the binding. These data, in turn, suggest ways to refine the lead compound to improve its binding to the target protein. The refined lead compound is then synthesized and complexed with the target, and further refined in a reiterative process. If lead compounds are available from other studies, such as screening of combinatorial libraries, these compounds may serve as starting points for this optimization cycle using structure-based drug design.

Once a sufficiently potent compound has been designed and optimized, its activity is evaluated in a biological system to establish the compound's ability to function in a physiological environment. If the compound fails at any stage of the biological evaluation, the design team reviews the structural model and uses crystallography to adjust structural features of the compound to overcome the difficulty. This process continues until a designed compound

exhibits the desired properties. The compound is then evaluated in an experimental disease model. If the compound fails, the reasons for failure (e.g., adverse metabolism, plasma binding, distribution, etc.) are determined, and again new modified compounds are designed to overcome the deficiencies without interfering with their ability to interact with the active site of the target protein. The experimental drug is then ready for conventional drug development (e.g., studies in safety assessment, formulation, clinical trials, etc.). This reiterative analysis and compound modification are possible because of the structural data obtained by X-ray crystallographic analysis at each stage. This capability renders structure-based drug design, a powerful tool for rapid and efficient development of drugs that are highly specific for particular protein target sites.

3.5. Docking methodologies

Several docking procedures exist in the literature, from the use of interactive graphics to manipulate the position of the ligand to completely automated procedures, which are becoming increasingly powerful to screen databases of molecules. Many docking algorithms follow a similar pattern. Usually, the first stage is to represent the molecules by their solvent-accessible surfaces. Beginning with a number of different relative orientations, the two molecules can then be brought together. More sophisticated methods carry out their moves by using a Monte Carlo algorithm to direct both rotations and translations. Rapid convergence to a minimum can be achieved by gradually cooling the simulation temperature as the molecule appears to be descending into a potential well.
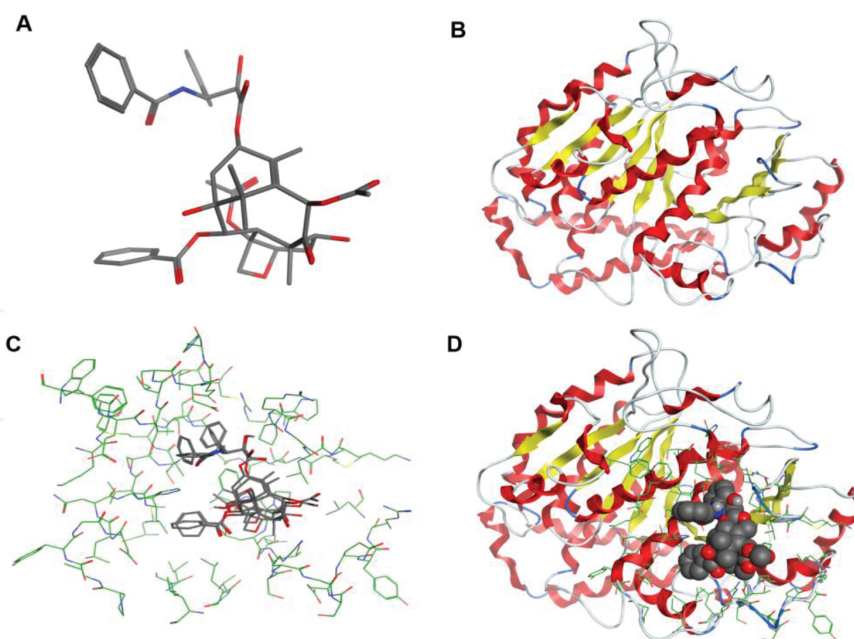

Figure 3. The molecular docking methodology. (A) Prepare the ligand structure, (B) Prepare the receptor structure, (C) Dock the ligand into the receptor surface with multiple potential conformations. (D) The most likely binding mode was identified based on intermolecular interaction between the ligand and the receptor surface. The protein backbone is represented as a cartoon. The ligand and active site residue are shown in stick representation.

Molecular docking is one of the most frequently used methods in structure-based drug design because of its ability to predict, with a substantial degree of accuracy, the conformation of small-molecule ligands within the appropriate target binding site (**Figure 3**) [19]. Following the development of the first algorithms in the 1980s, molecular docking became an essential tool in drug discovery [20]. Highly intensive investigations involving crucial molecular events, including ligand-binding modes and the corresponding intermolecular interactions which stabilize the ligand-receptor complex, could be conveniently performed. Furthermore, molecular docking algorithms execute quantitative predictions of binding energetics, given docked compounds, based on the binding affinity of ligand-receptor complexes [21].

3.6. Computer simulation in drug design

The development of a new drug starts with the design of suitable candidate compounds so-called "ligands" that are selected according to observations about how these compounds are recognized by the target protein and how they bind to it. It is important to know that proteins have dynamic properties: they change their shape in much the way as a machine tool needs to in order to fulfill its function. Now, it is being realized how important it is to have techniques available for studying protein dynamics. Performing experiments is not only expensive but also very time-consuming and, moreover, cannot answer all relevant questions. An alternative is computer-aided simulation of the dynamics of molecules [molecular dynamics (MD) simulations], which is becoming increasingly important to identify the molecular properties that are important and to determine the detailed molecular interactions that are critical for binding.

In some instances, high-performance computing (HPC) is required to cross the threshold where MD simulation becomes a valuable tool for industry. However, in most pharmaceutical companies, HPC is something very new, and supercomputers are simply not available to industrial researchers. With the arrival of affordable high-performance multiprocessor machines and corresponding developments of parallel software, it now becomes possible for industrial researchers to undertake more realistic calculations that were previously out of reach. Scientists at Novo Nordisk, a large Danish pharmaceutical company, are convinced that this new capability will dramatically change the acceptance of MD simulation as a tool in the design of new ligands. During Europort-D, they, for the first time, studied the dynamics of the complex molecular interactions critical for recognition of ligands by their target proteins.

4. Statistical tool for drug design

4.1. Introduction to SPASS for pharmaceutical research

As a sophisticated statistical software, Statistical Package for Social Science (SPSS) can help researchers and users realize their complex statistical tests with result interpretation and the access to data figures, tables, and graphs, which can be quickly and easily displayed in the output view. Even though additional training is usually required before users can maximize its features and the graphing feature is not as simple as an excel spreadsheet, SPSS is still being

used by more and more researchers from other fields for applications specific to their research areas, for example, the application of SPSS in clinical trials, with power and sample size calculation functions, comprehensive statistical tests and graphical tools, such as predictive data trends, forecasting and report generation, and the analysis of complex drug assays, etc. [22].

SPSS is being used majorly in pharmaceutical and medical research, where it can provide techniques for the design, implementation, and development of drug campaigns that are important to pharmaceutical research organizations and also provides data mining solutions based on their existing data on quality control and manufacturing, regulatory safety testing, clinical trials, data mining in drug discovery, compliance, and validation [23]. As many pharmaceutical research organizations have experienced that even small improvements in research or processing affect the efficiency and success of the project, SPSS provides a sophisticated analytical platforms to decrease development costs and unnecessary preclinical and clinical trials by integrating with existing data sources.

Overall, the tools in SPSS provide an efficient mechanism to automatically validate the routine analysis reports. In other words, SPSS empowers researchers with tools, which might bring the rapid and solid return on investments in the near future.

4.2. Outlines of the SPASS package

SPSS is a Windows-based program, which can be used to perform data entry and data analysis by tables and graphs (**Figure 4**). Because SPSS is capable of dealing with huge amounts of data with different modules, it is commonly used in the business world and in the Social Sciences [24]. Familiarity with this software will be useful in the field of drug research where big data,
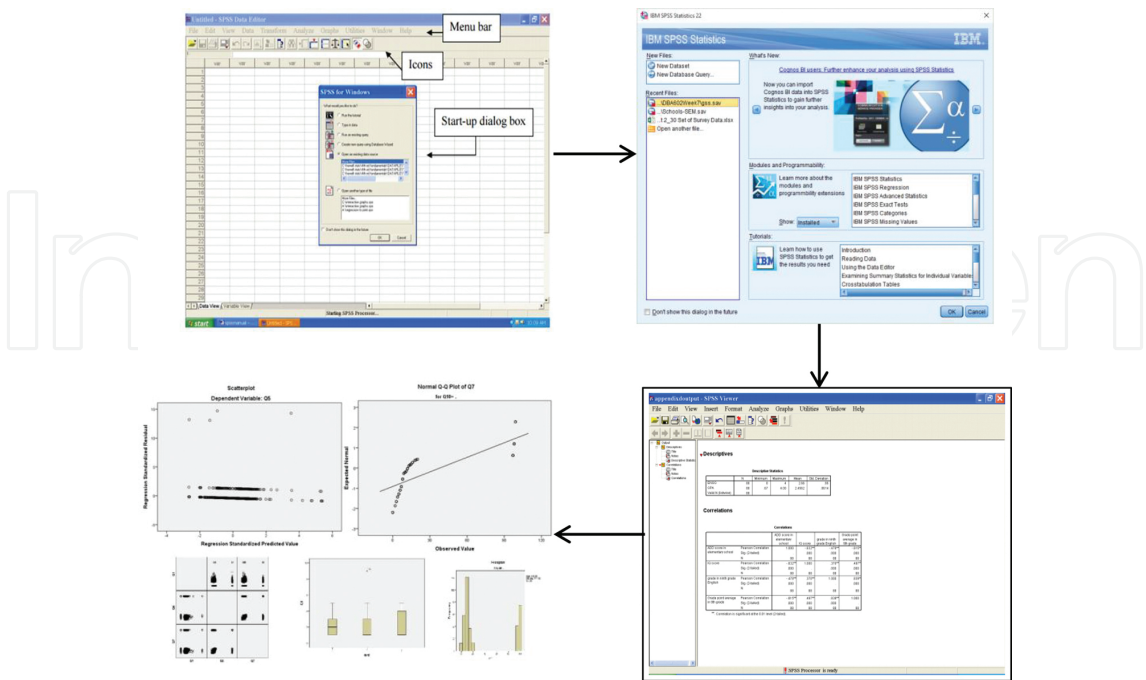

Figure 4. SPSS package in pharmaceutical data analytics application.

such as clinical trials data ranging from simple bioassays and dose-response experiments to long-term survival, and carcinogenicity studies need to be integrated. Unlike Excel, SPSS software can support screen transfers between the data entry view and the output view, which display the results and summarize the data.

4.3. Features of SPSS package

Whether its small and medium enterprises (SME) or large-scale enterprises (LSE), the application of SPSS can provide data accessing in rational data sets, which allows accessing inflexible resources and practices and real-time processing, and mapping of the database [25]. Although other commercial softwares, such as SAS, MATLAB, and others, can also perform such functions, SPSS is a more sophisticated application in statistics analysis. SPSS conducts statistical analysis from basic data, descriptive statistics, such as average and prevalence, and advanced inferential statistics, such as t test, regression model, one-way and factorial analysis of variance (ANOVA), one-way and factorial analysis of covariance (ANCOVA), one-way and factorial multivariate analysis of variance (MANOVA), one-way and factorial multivariate analysis of covariance (MANCOVA), factor analysis, path analysis, and logistic regression [26].

Normally, researchers use SPSS as a tool to collect and analyze data. The data entry screen in SPSS looks like an excel spreadsheet, and users can input variables and quantitative data and save the file as a data file [27]. After data are collected and entered into the data sheet in SPSS, users can also create an output file from the data they used. Then, the users can edit or organize the data in SPSS to check the running results after they choose the module they want. For example, users can check out the frequency distributions of the data to see whether the data set is normally distributed. Furthermore, the researchers can download the tables or graphs directly from the data output view. Because SPSS has the statistical tests built-in to the program, users do not need to do any math calculation or equations to get the results.

5. Conclusion and future direction

Systems pharmacology represents a new paradigm in drug research where ligand and protein information are combined to produce new methods for drug discovery and design. Traditionally, ligand-based approaches utilize information contained within the chemical structure to predict the biological properties of the drug. In particular, chemical similarity database searches can be used to identify molecules with improved activity. However, to be able to design compounds with satisfactory safety profiles, the drug targets will also need to be determined. Target prediction can be performed using ligand or structure data. Both of these approaches require a bioactivity database with pre-characterized activities or functions. If the target structure of the ligand is known, then the structure-based drug discovery approach can be applied. Computer-aided drug design techniques such as molecular docking and molecular dynamic simulation are capable of accurately predicting the binding sites of ligand, docking pose, and binding affinity based on the geometry of the receptor surface. The prediction can then be validated using *in vitro* biochemical assays and X-ray crystallography to validate ligand

binding. Thus, a cyclic drug design process is then continued until the strongest binder is found.

Currently, drug discovery has been mainly focused on in vitro optimization. One future direction in system-based drug discovery will necessarily shift to ADMET prediction to assess drug performance *in vivo*. Although several ADMET properties such as drug-likeness can be predicted by simple rules, a more detailed classification of drug properties can be achieved by more advanced machine learning techniques. Likewise, a holistic drug design process that simultaneously optimizes the drug properties based on *in vitro* and *in vivo* data will also hold promise to optimize drug safety and minimize adverse reactions. Consequently, integration of multilevel data from structure, tissue, and the whole organism will be required to achieve a more accurate prediction. In conclusion, we have presented the essential quantitative methods in system-based drug discovery, and the approaches presented here will stimulate further efforts in the progress of drug discovery and design to engineer safer and more effective drugs.

Author details

Yu-Chen Lo[1*], Ren Gui[2], Hiroshi Honda[2] and Jorge Z. Torres[3]

*Address all correspondence to: bennylo@stanford.edu

1 Department of Bioengineering, Stanford University, The United States

2 Department of Bioengineering, Northwestern Polytechnic University, The United States

3 Department of Chemistry and Biochemistry, University of California, Los Angeles, The United States

References

[1] Lo YC, Senese S, Damoiseaux R, Torres JZ. 3D chemical similarity networks for structure-based target prediction and scaffold hopping. ACS Chemical Biology. 2006;11(8):2244–53. doi:10.1021/acschembio.6b00253. PMID: 27285961.

[2] Barelier S, Sterling T, O'Meara MJ, Shoichet BK. The recognition of identical ligands by unrelated proteins. ACS Chemical Biology. 2015;10(12):2772–2784. doi:10.1021/acschembio.5b00683. PMID: 26421501; PMCID: PMCPMC4684825.

[3] Shoichet BK, Walters WP, Jiang H, Bajorath J. Advances in computational medicinal chemistry: A reflection on the evolution of the field and perspective going forward.

Journal of Medicinal Chemistry. 2016;59(9):4033–4034. doi:10.1021/acs.jmedchem.6b00511. PMID: 27054949.

[4] Lipinski CA. Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. Advanced Drug Delivery Reviews. 2016;101:34–41. doi:10.1016/j.addr.2016.04.029. PMID: 27154268.

[5] Yan X, Liao C, Liu Z, Hagler AT, Gu Q, Xu J. Chemical structure similarity search for ligand-based virtual screening: Methods and computational resources. Current Drug Targets. 2015:17(14):1580–85. PMID: 26521773.

[6] Le SQ, Ho TB, Phan TT. A novel graph-based similarity measure for 2D chemical structures. Genome Informatics. 2004;15(2):82–91. PMID: 15706494.

[7] Raymond JW, Willett P. Effectiveness of graph-based and fingerprint-based similarity measures for virtual screening of 2D chemical structure databases. Journal of Computer-Aided Molecular Design. 2002;16(1):59–71. PMID: 12197666.

[8] Lo YC, Senese S, Li CM, Hu Q, Huang Y, Damoiseaux R, et al. Large-scale chemical similarity networks for target profiling of compounds identified in cell-based chemical screens. PLoS Computational Biology. 2015;11(3):e1004153. doi:10.1371/journal.pcbi.1004153. PMID: 25826798; PMCID: PMCPMC4380459.

[9] Chen X, Yan CC, Zhang X, Zhang X, Dai F, Yin J, et al. Drug-target interaction prediction: Databases, web servers and computational models. Briefings in Bioinformatics. 2016;17(4):696–712. doi:10.1093/bib/bbv066. PMID: 26283676.

[10] Davies M, Nowotka M, Papadatos G, Dedman N, Gaulton A, Atkinson F, et al. ChEMBL web services: Streamlining access to drug discovery data and utilities. Nucleic Acids Research. 2015;43(W1):W612–W620. doi:10.1093/nar/gkv352. PMID: 25883136; PMCID: PMCPMC4489243.

[11] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. Nature Biotechnology. 2007;25(2):197–206. doi:10.1038/nbt1284. PMID: 17287757.

[12] Liu T, Altman RB. Using multiple microenvironments to find similar ligand-binding sites: Application to kinase inhibitor binding. PLoS Computational Biology. 2011;7(12):e1002326. doi:10.1371/journal.pcbi.1002326. PMID: 22219723; PMCID: PMCPMC3248393.

[13] Westerhoff HV. Network-based pharmacology through systems biology. Drug Discovery Today: Technologies. 2015;15:15–16. doi:10.1016/j.ddtec.2015.05.001. PMID: 26464085.

[14] Liu T, Altman RB. Relating essential proteins to drug side-effects using canonical component analysis: A structure-based approach. Journal of Chemical Information and

Modeling. 2015;55(7):1483–1494. doi:10.1021/acs.jcim.5b00030. PMID: 26121262; PMCID: PMCPMC4875781.

[15] Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. British Journal of Pharmacology. 2011;162(6):1239–1249. doi:10.1111/j.1476-5381.2010.01127.x. PMID: 21091654; PMCID: PMCPMC3058157.

[16] Wilson GL, Lill MA. Integrating structure-based and ligand-based approaches for computational drug design. Future Medicinal Chemistry. 2011;3(6):735–750. doi: 10.4155/fmc.11.18. PMID: 21554079.

[17] Fang Y. Ligand-receptor interaction platforms and their applications for drug discovery. Expert Opinion on Drug Discovery. 2012;7(10):969–988. doi: 10.1517/17460441.2012.715631. PMID: 22860803.

[18] Kahsai AW, Xiao K, Rajagopal S, Ahn S, Shukla AK, Sun J, et al. Multiple ligand-specific conformations of the beta2-adrenergic receptor. Nature Chemical Biology. 2011;7(10): 692–700. doi:10.1038/nchembio.634. PMID: 21857662; PMCID: PMCPMC3404607.

[19] Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: A powerful approach for structure-based drug discovery. Current Computer-Aided Drug Design. 2011;7(2):146–157. PMID: 21534921; PMCID: PMCPMC3151162.

[20] Lopez-Vallejo F, Caulfield T, Martinez-Mayorga K, Giulianotti MA, Nefzi A, Houghten RA, et al. Integrating virtual screening and combinatorial chemistry for accelerated drug discovery. Combinatorial Chemistry and High Throughput Screening. 2011;14(6): 475–487. PMID: 21521151.

[21] Huang SY, Zou X. Advances and challenges in protein-ligand docking. International Journal of Molecular Sciences. 2010;11(8):3016–3034. doi:10.3390/ijms11083016. PMID: 21152288; PMCID: PMCPMC2996748.

[22] Williams F. Reasoning with Statistics: How to Read Quantitative Research. 4th ed. Fort Worth: Harcourt Brace Jovanovich College Publishers; 1992, Vol. xiii, 220 p.

[23] Sprinthall RC. Basic Statistical Analysis. 6th ed. Boston: Allyn and Bacon; 2000, Vol. xiv, 626 p.

[24] Norusis MJ, SPSS Inc. SPSS 8.0 Guide to Data Analysis. Upper Saddle River, N.J.: Prentice Hall; 1998, Vol. xviii, 563 p.

[25] Harris MB. Basic Statistics for Behavioral Science Research. 2nd ed. Boston: Allyn and Bacon; 1998, Vol. xvi, 576 p.

[26] Asher HB. Causal Modeling. 2nd ed. Beverly Hills: Sage Publications; 1983, 96 p.

[27] George D, Mallery P. SPSS for Windows Step by Step: A Simple Guide and Reference 9.0 Update. 2nd ed. Boston: Allyn and Bacon; 2000, 357 p.