



Transfer and Multi-task Learning in QSAR Modeling: Advances and Challenges

Rodolfo S. Simões¹, Vinicius G. Maltarollo², Patricia R. Oliveira¹ and Kathia M. Honorio^{1,3*}

¹ School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, Brazil, ² Department of Pharmaceutical Products, Faculty of Pharmacy, Federal University of Minas Gerais, Belo Horizonte, Brazil, ³ Center for Natural and Human Sciences, Federal University of ABC, Santo André, Brazil

Medicinal chemistry projects involve some steps aiming to develop a new drug, such as the analysis of biological targets related to a given disease, the discovery and the development of drug candidates for these targets, performing parallel biological tests to validate the drug effectiveness and side effects. Approaches as quantitative study of activity-structure relationships (QSAR) involve the construction of predictive models that relate a set of descriptors of a chemical compound series and its biological activities with respect to one or more targets in the human body. Datasets used to perform QSAR analyses are generally characterized by a small number of samples and this makes them more complex to build accurate predictive models. In this context, transfer and multi-task learning techniques are very suitable since they take information from other QSAR models to the same biological target, reducing efforts and costs for generating new chemical compounds. Therefore, this review will present the main features of transfer and multi-task learning studies, as well as some applications and its potentiality in drug design projects.

Keywords: drug design, medicinal chemistry, QSAR, machine learning, transfer learning, multi-task learning

OPEN ACCESS

Edited by:

Leonardo G. Ferreira,
University of São Paulo, Brazil

Reviewed by:

Simone Brogi,
University of Siena, Italy
Andrew James Greenshaw,
University of Alberta, Canada

*Correspondence:

Kathia M. Honorio
kmhonorio@usp.br

Specialty section:

This article was submitted to
Experimental Pharmacology and Drug
Discovery,
a section of the journal
Frontiers in Pharmacology

Received: 22 September 2017

Accepted: 22 January 2018

Published: 06 February 2018

Citation:

Simões RS, Maltarollo VG,
Oliveira PR and Honorio KM (2018)
Transfer and Multi-task Learning
in QSAR Modeling: Advances
and Challenges.
Front. Pharmacol. 9:74.
doi: 10.3389/fphar.2018.00074

INTRODUCTION

The drug design process, since the discovery/identification of bioactive compounds until the approval of its clinical use by a regulatory agency, is very complex and demands time and financial support (Tufts Center for the Study of Drug Development [CSDD], 2014). There are several well-known bottlenecks in this process, such as finding out a suitable and validated molecular target, designing and/or discovering of a lead compound, pharmacokinetic and toxicity optimization, besides commercial reasons, efficacy and clinical safety (Khanna, 2012; Medina-Franco et al., 2013). In this scenario, the use of computational techniques in drug discovery is rapidly increasing.

Computer-aided drug design (CADD) techniques are broadly employed in order to reduce costs and time involved in drug design. Among the important CADD techniques, molecular docking, similarity search and QSAR studies could be highlighted. Molecular docking and virtual screening are considered structure-based drug design (SBDD) strategies since it requires 3D structure of a molecular target and consists of predicting a binding mode of molecules and its binding energy (Walters et al., 1998; Shoichet, 2004; Andricopulo et al., 2008). As the docking simulations consider both structures (ligands and targets), its calculations are more computationally expensive. Considering these aspects, similarity searches and pharmacophore modeling are alternatives to faster calculations (Brogi et al., 2009; Tresadern et al., 2009) and are defined as ligand-based drug design (LBDD) strategies since they do not require the biological target structure

(Turki et al., 2017). **Figure 1** illustrates the main steps in a drug design process, including the use of computational tools.

Another LBDD strategy is known as quantitative structure-activity relationships (QSAR) and it has been widely employed in drug design, mainly aiming to predict the biological activity of a compound set against a specific target to optimize the binding affinity (Du et al., 2008; Gertrudes et al., 2012). QSAR models provide accurate predictions of measured endpoints instead of an independent ranking of biological activity. These quantitative approaches have also been used in other tasks, such as optimization of pharmacokinetics and toxicity profile (Maltarollo et al., 2015; Egeghy et al., 2016; Chemi et al., 2017) and virtual screening (Brogi et al., 2013; Melo-Filho et al., 2016; Neves et al., 2016; Zaccagnini et al., 2017).

Several important QSAR studies can be found in literature, which include the description of successful computational methods and algorithms (Sliwoski et al., 2014; Raies and Bajic, 2016), validation techniques (Gramatica and Sangion, 2016), applications (Cherkasov et al., 2014; Fang and Xiao, 2016) as well-challenges and how those have been addressed (Cronin and Schultz, 2003; Arthur, 2008; Dearden et al., 2009; Scior et al., 2009; Wang et al., 2015; Ponzoni et al., 2017).

In many recent studies, machine learning (ML) methods have been largely applied to QSAR analyses. This growth has been mainly motivated by the increasing availability of data in public repositories, the use of numerous and diverse chemical descriptors and the proposal of accurate predictive algorithms, such as support vector machines (SVMs) and artificial neural networks (Gertrudes et al., 2012; Maltarollo et al., 2013; Mitchell, 2014; Lima et al., 2016). A common application of ML techniques in CADD refers to forecast new compound class labels (e.g., “active” versus “inactive”) using models previously derived from available training sets (Lavecchia, 2015). In such specific situation, ML techniques are said to perform a classification learning task. In addition, other sort of learning tasks can also be considered in CADD, such as clustering and ranking (Agarwal et al., 2010).

Despite of the widespread use of ML methods in QSAR modeling, the success of such approaches critically depends on the availability of a great amount of data, which remains challenging in drug discovery. This problem is strongly related to issues involving the quality of public data sources, including imprecise representation of chemical structures and inaccurate activity information (Zhao et al., 2017). Furthermore, the nature of different experimental protocols can usually lead to data belonging to different probability distributions, which makes the use of traditional ML techniques impracticable.

The data sets available in public repositories are usually obtained from single structure-activity relationship (SAR) campaigns. This explains the several particular and linear sets of compounds that are commonly used to generate only specialized QSAR models. In most of cases, biological activities of two datasets are measured under different experimental conditions, making the link among chemical spaces difficult to be analyzed (Richter and Ecker, 2015). Furthermore, a large chemical space has activity cliffs naturally: regions in a structure/activity surface

where there is a discontinuous SAR (Cruz-Monteaudo et al., 2014).

In 2014, a review on QSAR (Cherkasov et al., 2014) stated that the transferability of QSAR models is one of the challenges in QSAR modeling, since the traditional approaches have been typically designed for each target property individually. Aiming to take advantage of diverse but related available experimental data, transfer and multi-task learning techniques have been recently developed. The novelty behind these approaches is related to their ability to exploit knowledge from other related tasks to improve the learning performance, especially when a small data set is available for training.

TRANSFER AND MULTI-TASK LEARNING

For QSAR purposes, the data space under analysis is characterized by biological and chemical properties. In such scenario, changes in the distribution of data force the model to be rebuilt, implying to collect new training data. However, in many real-world applications, it is expensive or impossible to recollect data required to reconstruct these models. In such situations, transfer learning (or knowledge transfer) among related domains would be desirable (Pan and Yang, 2010).

Transfer learning can be defined as the ability of a system to recognize and apply the knowledge learned in previous (source) tasks for the solution of new (target) problems. The development of such approach was motivated by the fact that one can apply the knowledge acquired previously to solve new problems more quickly and with better solutions. The goal here relies on extracting the knowledge obtained by a model from one or more source tasks and to apply it to a target task. However, one of the premises for using transfer learning technique is that the source and the destination domains must be related. In this sense, Tan et al. (2015) suggest that such relationship can be expressed by instances (Bickel et al., 2009) or characteristics (Satpal and Sarawagi, 2007). If no direct relationship is found, the forced transfer will not work, resulting in no improvement or even degenerating the performance in the target domain (Fitzgerald and Thomaz, 2015). Multi-task learning is closely related to knowledge transfer, but they have also a clear distinction. In multi-task approaches, a number of tasks are learned simultaneously, without involving designated source and target tasks. **Figure 2** illustrates the overall schemes for transfer and multi-task learning.

The methods used for transfer learning can be summarized into four categories, depending on which aspect of knowledge will be transferred, i.e., “what to transfer” (Pan and Yang, 2010). The first category refers to instance-based transfer learning, which assumes that some data from the source set can be selected for training in the target set by re-weighting. Importance sampling and instance reweighting are the two most commonly techniques used (Dai et al., 2007). The second category refers to transfer learning methods by feature representation, which focuses on encoding the structural information carried by molecules into a numerical representation that can be effectively exploited by learning processes in other related problems. In this case,

the intuitive idea consists in learning a suitable representation of characteristics for use in the target set, i.e., the transfer learning is coded in the representation of the new characteristics (Raina et al., 2007). The third category refers to the transfer learning techniques by parameters (Lawrence and Platt, 2004), in which it is assumed that the source and the target tasks share some parameters or prior distributions of the hyper-parameters of the respective QSAR models. In this case, knowledge can be transferred between the tasks by discovering these shared parameters or priors. The last category consists of methods that deal with the problem of relational knowledge transfer, which refers to transfer learning in related domains (Mihalkova et al., 2007). In this condition, the knowledge can be transferred by mapping the data from the source set to the destination one. The statistical methods of relational learning are the most applied in this case (Mihalkova et al., 2007; Davis and Domingos, 2009). A scheme illustrating how the transfer learning approaches can be applied to obtain predictive models is presented in **Figure 3**.

To apply transfer learning techniques, it is assumed that two sets of related data are available and the knowledge will be transferred from the dataset with the largest volume to the set with the least amount of available data. However, this assumption in the chemical datasets is not always sufficient, requiring the opinion of an expert to define the source datasets. To overcome this limitation, Girschick et al. (2012) proposed an approach to select a source dataset in a repository containing target-related sets by following a data-driven methodology. The main idea behind such proposal is based on calculating a measure for the activity overlap between the target set and each related set available in PubChem database. As result, a ranking of all related sets according to their similarity to the target set is obtained. In order to find the similarity values, Tanimoto coefficient is calculated using the categorization of the chemical compounds (active/inactive) in each dataset. Therefore, the objective is to select the set that has the distribution of instances (compounds) closest to the distribution (number of instances categorized as active and inactive) of the target set.

One can find out many situations where transfer learning adds benefits, for example, molecules could be classified as active or inactive according to a biological data for a defined endpoint (e.g., IC_{50} values). For this classification task, it is initially necessary to collect several experimentally tested samples and, next, to train a classifier for the collected data with their respective labels. Since the probability distribution of the comments on other endpoints can be very different, a new classifier has to be trained to each dataset in order to maintain a satisfactory performance. To reduce this effort, it would be desirable to use the knowledge from a classification model that is already trained on some related endpoints to improve the classification performance of other tasks with small samples or datasets (Turki et al., 2017). **Table 1** illustrates examples of transfer learning in drug design.

In general, transfer learning approaches have shown to be promising for combining the knowledge previously obtained in related tasks into a single predictive model, whether for classification, regression, or grouping (Pan and Yang, 2010). In particular, researches in medicinal chemistry with focus in drug discovery have been benefited with the use of transfer learning, as can be seen in previous studies (Girschick et al., 2012; Rosenbaum et al., 2013; Saha et al., 2016). Next, applications of transfer and multi-task learning in medicinal chemistry studies will be presented.

SOME APPLICATIONS OF TRANSFER AND MULTI-TASK LEARNING

Many machine learning methods are based on the assumption that similar drugs may share the same side effects, but measuring the similarity of these drugs is still a challenge. However, the use of data from various sources (similar drugs) provides important information for the analysis of side effects and should be integrated for obtaining a highly accurate prediction. Zhang et al. (2016) discussed the problem of predicting side effects

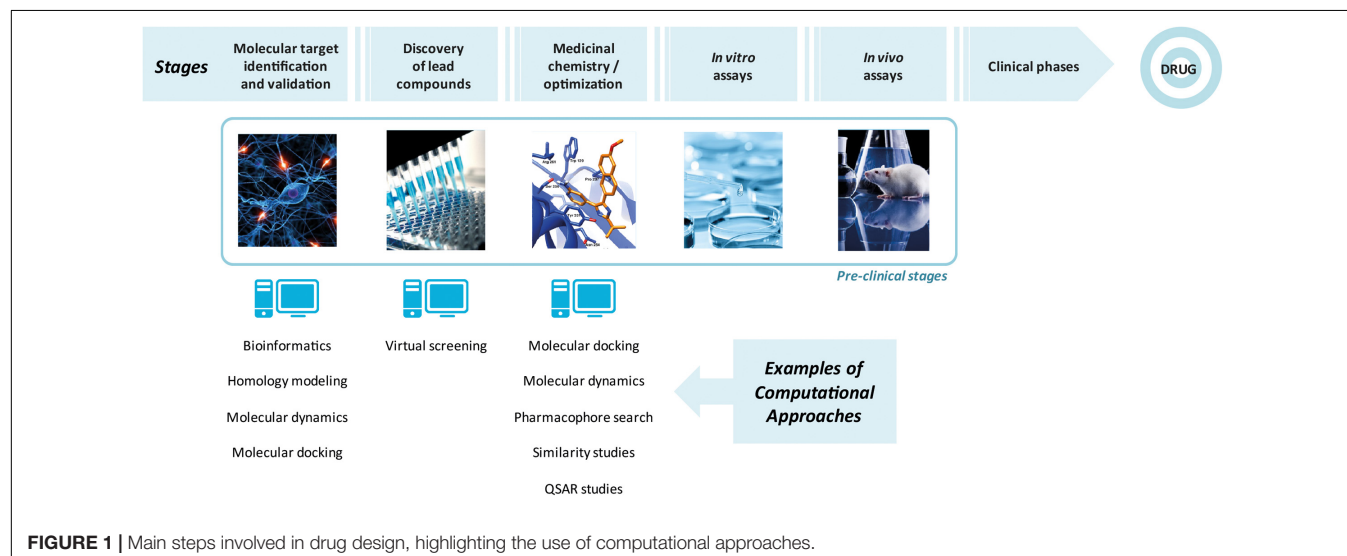
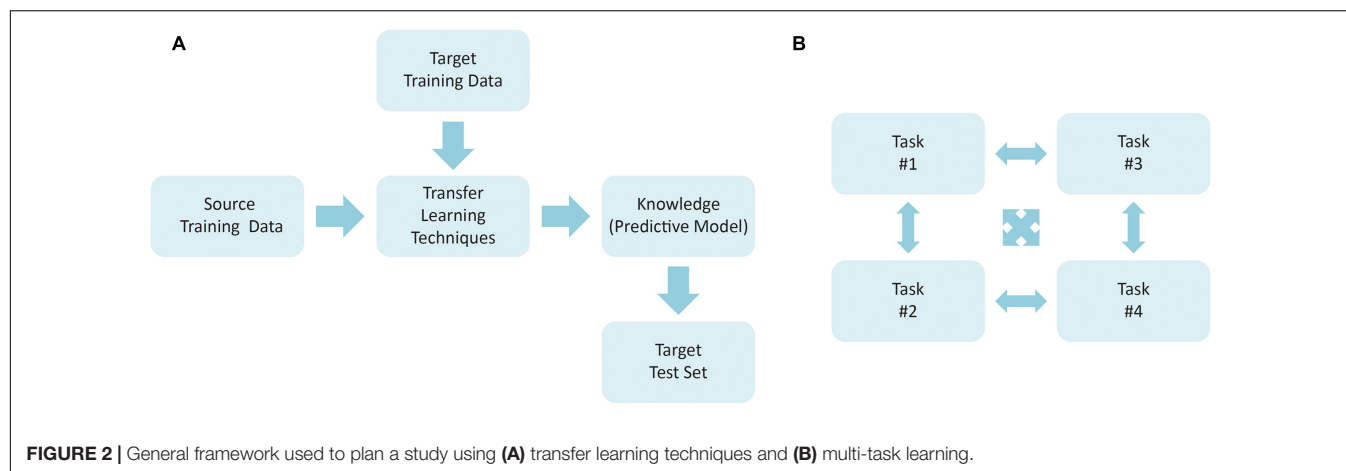
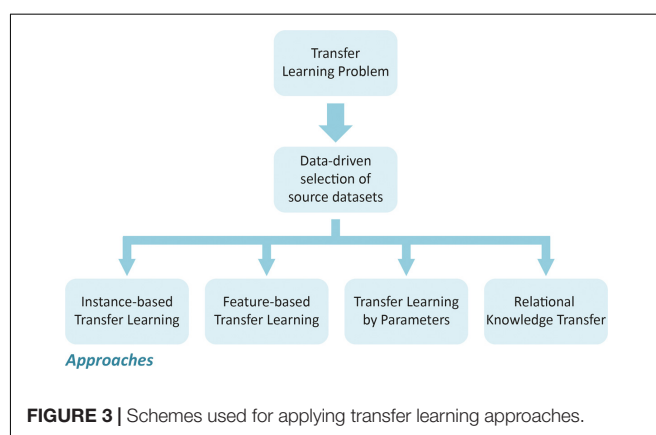


FIGURE 1 | Main steps involved in drug design, highlighting the use of computational approaches.



caused by drugs through linear neighbor approaches and the integration of data from various sources. The authors argued that auxiliary data can bring additional and diverse information (such as drug substructures, drug targets, drug transporters, drug enzymes, drug pathways) that should be integrated to the side-effect prediction, aiming at improving its performance. Analyses on multi-label classification showed that the proposed transfer learning approaches achieved better performance than state-of-art-methods (Pauwels et al., 2011; Liu et al., 2012; Cheng et al., 2013) applied to benchmark datasets.

The task of relating chemical structure to biological activity in QSAR studies is usually based on the notion of chemical similarity to predict the molecular behavior of close compounds. So, techniques that provide similarity measures among chemical compounds are increasingly important (Floris et al., 2014). Lately, relevant solutions have been proposed, which comprise distance learning (Biehl et al., 2014) and inductive transfer (Garcke and Vanck, 2014) methods. Distance learning aims at learning an appropriate distance measure to reflect the underlying relationship between instances in the training set, while inductive transfer refers to the process of transferring knowledge learned from one task into another related task. Girschick et al. (2012) presented an adapted transfer approach, which combines distance learning and inductive transfer by learning the distances on a related task and then transferring them to the target learning task. Additionally, the authors



developed a method for selecting a related task that can be used as source task for transfer learning. This technique consists in applying an activity overlap similarity measure to two datasets to find out a suitable source task. This approach was evaluated on five distinct datasets found in PubChem BioAssay (Wang et al., 2009) repository. The results showed that both proposals worked well for large and small amounts of training data.

The multi-task learning approach (Caruana, 1998) is considered to be closely related to transfer learning, since it attempts to learn multiple tasks simultaneously even when they

TABLE 1 | Examples of potential applications of transfer learning methods in drug design.

Transfer learning approach	Concept	Application in drug design
Instance-based	Uses the same ML technique for modeling but apply some changes to the parameters of the target model.	Source and target datasets have the same endpoint (e.g., same molecular target) but the training data can be collected at different experimental conditions.
Feature representation	Based on some mathematical transformations of data.	Source and target datasets have different but related endpoints, e.g., same classes of molecular target (kinases, nuclear receptors, proteases, etc.).
Parameters	It is assumed that both datasets share some properties.	Source and target datasets have the same or related endpoints.
Relational knowledge transfer	Based on technique for mapping the data in the target domain.	The endpoints of the source and target datasets are different but the domains (the independent variable in QSAR models) are related; e.g., cellular permeability and log P.

are different. Rosenbaum et al. (2013) introduced two multi-task methods and evaluated the performance of such approaches by inferring multi-target QSAR models on a subset of human kinome. The authors assumed that the taxonomical relationship of the kinase targets should correspond to the relatedness of the QSAR problems on these targets. The multi-task techniques were compared to SVMs models independently trained for each target and an SVM model that assumed all targets to be identical. The results demonstrated that the multi-target learning can overperform baseline (pure SVM) methods if knowledge can be transferred from a target with a lot of data to a similar target with little domain knowledge.

Varnek et al. (2009) applied different inductive transfer and multi-task learning approaches to model tissue-air partition coefficients. The authors found that these techniques improved the prediction accuracy of the obtained models when compared to single task learning. Finally, this study indicated that inductive transfer learning is very suitable when single modeling is unable to generate reliable QSAR models using diverse data sets and with small amount of samples.

Brown et al. (2014) presented some challenges involved with chemogenomic data, since high-throughput assays give us a large number of information from multi-ligand and multi-target data (Pereira and Williams, 2007). So, the authors assert that computational techniques, in particular inductive transfer and explicit learning, can help to construct more robust models when compared to target-specific (classical) QSAR ones.

The study of Zhang et al. (2013) discussed the use of single- and multi-task learning to construct QSAR models for predicting the binding affinity of a compound database by estrogen receptors (ERs), which are involved with endocrine disruption by chemicals and the construction of predictive models can contribute to design safer substances. The authors concluded that multi-task learning provided better results for a small dataset (ER β ligands) than single learning, indicating that this approach can be considered as a good tool to understand the action mechanism of endocrine disruption and to predict the ER activity of unknown compounds as endocrine disrupting chemicals.

Another interesting application of multi-task techniques was performed by Liu et al. (2011), which used multi-task learning to construct multi-target QSAR models employing three human immunodeficiency virus (HIV) inhibitor datasets together with other six subsets containing two hepatitis C virus (HCV) inhibitors. The main conclusions of this study included the fact that the integration of all databases (HIV and HCV) improved the rate of the discovery of lead HIV-HCV inhibitors, helping the design of new co-inhibitors for these important infections. Other achievement is related to the successful use (considering efficiency in convergence speed and learning accuracy) of a multi-task learning technique to construct multi-target QSAR models.

DISCUSSION

The main issue of transfer and multi-task learning approaches is to employ the knowledge generated (e.g., features, subset of

variables, weights of equations) from available ML models and other datasets in the construction of models for related endpoints. In this sense, it is possible to use different datasets with the same biological activities but measured at different experimental conditions. Other important consequence of applying transfer and multi-task learning is the decrease on computational costs related to the faster convergence obtained by using the knowledge derived from a model previously built from a related endpoint.

From a literature review taking into account the transfer learning applications on medicinal chemistry, one can note that there is still a great potentiality to be explored in this sense. Other emerging approaches as deep learning methods (Zhang et al., 2017), which basically use complex neural networks architectures, also have promising applications in the era of big data.

Among the main challenges on applying transfer and multi-task learning methods is that they require an artificial intelligence expert to code them since there are no chemical and/or pharmaceutical packages with a graphical user interface. Depend on the source data and on the learning method, transfer and multi-task learning could be also considered as “black-boxes,” making the interpretability of QSAR models difficult. And, finally, the transfer of knowledge could be inappropriately employed if the assumption of “equivalent” endpoints is not valid.

CONCLUSION

Nowadays one can observe increasing number of applications of transfer and multi-task learning in medicinal studies. There are also current challenges in the QSAR field that comprise the integration of different datasets (even from different experiments) aiming the same or similar endpoints (Maltarollo et al., 2017) and the development of universal QSAR models using very large datasets (Alves et al., 2017). Therefore, good examples of dataset that could be benefited from transfer and multi-task learning are: (i) compounds with same endpoint measured under different experimental conditions; (ii) antimicrobial activities against genetically similar microorganisms; (iii) compounds with the same mechanism of action in homologous targets and high degree of similarity in the binding pocket; (iv) non-specific endpoints as toxicity against a cell line or permeability rates determined by different models. In this complex scenario, transfer and multi-task learning techniques can be considered powerful tools for drug design.

AUTHOR CONTRIBUTIONS

RS, KH, VM, and PO designed this article. All authors wrote and revised the manuscript. Also, all authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank FAPESP, CNPq, CAPES, and IBM for funding.

REFERENCES

- Agarwal, S., Dugar, D., and Sengupta, S. (2010). Ranking chemical structures for drug discovery: a new machine learning approach. *J. Chem. Inf. Model.* 50, 716–731. doi: 10.1021/ci9003865
- Alves, V. M., Muratov, E. N., Zakharov, A., Muratov, N. N., Andrade, C. H., and Tropsha, A. (2017). Chemical toxicity prediction for major classes of industrial chemicals: Is it possible to develop universal models covering cosmetics, drugs, and pesticides? *Food Chem. Toxicol.* doi: 10.1016/j.fct.2017.04.008 [Epub ahead of print].
- Andricopulo, A. D., Guido, R. V. C., and Oliva, G. (2008). Virtual screening and its integration with modern drug design technologies. *Curr. Med. Chem.* 15, 37–46. doi: 10.2174/092986708783330683
- Arthur, M. D. (2008). QSAR: dead or alive? *J. Comput. Aided Mol. Des.* 22, 81–89. doi: 10.1007/s10822-007-9162-7
- Bickel, S., Sawade, C., and Scheffer, T. (2009). Transfer learning by distribution matching for targeted advertising. *Adv. Neural Inf. Process. Syst.* 21, 145–152.
- Biehl, M., Hammer, B., and Villmann, T. (2014). “Distance measures for prototype based classification,” in *Brain-Inspired Computing. BrainComp 2013. Lecture Notes in Computer Science*, eds L. Grandinetti, T. Lippert, and N. Petkov (Cham: Springer), doi: 10.1007/978-3-319-12084-3_9
- Broggi, S., Kladi, M., Vagias, C., Papazafiri, P., Roussis, V., and Tafi, A. (2009). Pharmacophore modeling for qualitative prediction of antiestrogenic activity. *J. Chem. Inf. Model.* 49, 2489–2497. doi: 10.1021/ci900254b
- Broggi, S., Papazafiri, P., Roussis, V., and Tafi, A. (2013). 3D-QSAR using pharmacophore-based alignment and virtual screening for discovery of novel MCF-7 cell line inhibitors. *Eur. J. Med. Chem.* 67, 344–351. doi: 10.1016/j.ejmech.2013.06.048
- Brown, J. B., Okuno, Y., Marcou, G., Varnek, A., and Horvath, D. (2014). Computational chemogenomics: Is it more than inductive transfer? *J. Comput. Aided Mol. Des.* 28, 597–618. doi: 10.1007/s10822-014-9743-1
- Caruana, R. (1998). “Multitask learning,” in *Learning to Learn*, eds S. Thrun and L. Pratt (Boston, MA: Springer), 95–133. doi: 10.1007/978-1-4615-5529-2_5
- Chemi, G., Gemma, S., Campiani, G., Broggi, S., Butini, S., and Brindisi, M. (2017). Computational tool for fast *in silico* evaluation of hERG K⁺ channel affinity. *Front. Chem.* 5:7. doi: 10.3389/fchem.2017.00007
- Cheng, F., Li, W., Wang, X., Zhou, Y., Wu, Z., Shen, J., et al. (2013). Adverse drug events: database construction and *in silico* prediction. *J. Chem. Inf. Model.* 53, 744–752. doi: 10.1021/ci4000079
- Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., et al. (2014). QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.* 57, 4977–5010. doi: 10.1021/jm4004285
- Cronin, M. T. D., and Schultz, T. W. (2003). Pitfalls in QSAR. *J. Mol. Struct. Theochem* 622, 39–51. doi: 10.1016/S0166-1280(02)00616-4
- Cruz-Monteagudo, M., Medina-Franco, J. L., Perez-Castillo, Y., Nicolotti, O., Cordeiro, M. N. D., and Borges, F. (2014). Activity cliffs in drug discovery: Dr Jekyll or Mr Hyde? *Drug Discov. Today* 19, 1069–1080. doi: 10.1016/j.drudis.2014.02.003
- Dai, W., Yang, Q., and Yu, Y. (2007). “Boosting for transfer learning,” in *Proceedings of the 24th international conference on Machine learning*, (New York, NY: ACM), 193–200. doi: 10.1145/1273496.1273521
- Davis, J., and Domingos, P. (2009). “Deep transfer via second-order markov logic,” in *Proceedings of the 26th Annual International Conference on Machine Learning*, (New York, NY: ACM), 217–224. doi: 10.1145/1553374.1553402
- Dearden, J. C., Cronin, M. T. D., and Kaiser, K. L. E. (2009). How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). *SAR QSAR Environ. Res.* 20, 241–266. doi: 10.1080/10629360902949567
- Du, Q.-S., Huang, R.-B., and Chou, K.-C. (2008). Recent advances in QSAR and their applications in predicting the activities of chemical molecules, peptides and proteins for drug design. *Curr. Protein Pept. Sci.* 9, 248–259. doi: 10.2174/138920308784534005
- Egeghy, P. P., Sheldon, L. S., Isaacs, K. K., Özkaynak, H., Goldsmith, M. R., Wambaugh, J. F., et al. (2016). Computational exposure science: an emerging discipline to support 21st-century risk assessment. *Environ. Health Perspect.* 124, 697–702. doi: 10.1289/ehp.1509748
- Fang, C., and Xiao, Z. (2016). Receptor-based 3D-QSAR in drug design: methods and applications in kinase studies. *Curr. Top. Med. Chem.* 16, 1463–1477. doi: 10.2174/1568026615666150915120943
- Fitzgerald, T., and Thomaz, A. L. (2015). “Skill demonstration transfer for learning from demonstration,” in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, (New York, NY: ACM), 187–188. doi: 10.1145/2701973.2702728
- Floris, M., Manganaro, A., Nicolotti, O., Medda, R., Mangiardi, G. F., and Benfenati, E. (2014). A generalizable definition of chemical similarity for read-across. *J. Cheminform.* 6:39. doi: 10.1186/s13321-014-0039-1
- Garcke, J., and Vanck, T. (2014). Importance weighted inductive transfer learning for regression. *Lect. Notes Comput. Sci.* 8274, 466–481. doi: 10.1007/978-3-662-44848-9_30
- Gertrudes, J. C., Maltarollo, V. G., Silva, R. A., Oliveira, P. R., Honorio, K. M., and Da Silva, A. B. F. (2012). Machine learning techniques and drug design. *Curr. Med. Chem.* 19, 4289–4297. doi: 10.2174/092986712802884259
- Girschick, T., Ruckert, U., and Kramer, S. (2012). Adapted transfer of distance measures for quantitative structure-activity relationships and data-driven selection of source datasets. *Comput. J.* 56, 274–288. doi: 10.1093/comjnl/bxs092
- Gramatica, P., and Sangion, A. (2016). A historical excursus on the statistical validation parameters for QSAR models: a clarification concerning metrics and terminology. *J. Chem. Inf. Model.* 56, 1127–1131. doi: 10.1021/acs.jcim.6b00088
- Khanna, I. (2012). Drug discovery in pharmaceutical industry: productivity challenges and trends. *Drug Discov. Today* 17, 1088–1102. doi: 10.1016/j.drudis.2012.05.007
- Lavecchia, A. (2015). Machine-learning approaches in drug discovery: methods and applications. *Drug Discov. Today* 20, 318–331. doi: 10.1016/j.drudis.2014.10.012
- Lawrence, N. D., and Platt, J. C. (2004). “Learning to learn with the informative vector machine,” in *Proceedings of the Twenty-first International Conference on Machine Learning*, (New York, NY: ACM), 65. doi: 10.1145/1015330.1015382
- Lima, A. N., Philot, E. A., Trossini, G. H. G., Scott, L. P. B., Maltarollo, V. G., and Honorio, K. M. (2016). Use of machine learning approaches for novel drug discovery. *Exp. Opin. Drug Discov.* 11, 225–239. doi: 10.1517/17460441.2016.1146250
- Liu, M., Wu, Y., Chen, Y., Sun, J., Zhao, Z., Chen, X. W., et al. (2012). Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *J. Am. Med. Inform. Assoc.* 19, e28–e35. doi: 10.1136/amiajnl-2011-000699
- Liu, Q., Zhou, H., Liu, L., Chen, X., Zhu, R., and Cao, Z. (2011). Multi-target QSAR modelling in the analysis and design of HIV-HCV co-inhibitors: an *in-silico* study. *BMC Bioinformatics* 12:294. doi: 10.1186/1471-2105-12-294
- Maltarollo, V. G., Gertrudes, J. C., Oliveira, P. R., and Honorio, K. M. (2015). Applying machine learning techniques for ADME-Tox prediction: a review. *Expert Opin. Drug Metab. Toxicol.* 11, 259–271. doi: 10.1517/17425255.2015.980814
- Maltarollo, V. G., Honorio, K. M., and da Silva, A. B. F. (2013). “Applications of artificial neural networks in chemical problems,” in *Artificial Neural Networks-Architectures and Applications*, ed. K. Suzuki (Rijeka: InTech), 203–223. doi: 10.5772/51275
- Maltarollo, V. G., Kronenberger, T., Wrenger, C., and Honorio, K. M. (2017). Current trends in quantitative structure-activity relationship validation and applications on drug discovery. *Future Sci. OA* 3, FSO214–FSO216. doi: 10.4155/foa-2017-0052
- Medina-Franco, J. L., Giulianotti, M. A., Welmaker, G. S., and Houghten, R. A. (2013). Shifting from the single to the multitarget paradigm in drug discovery. *Drug Discov. Today* 18, 495–501. doi: 10.1016/j.drudis.2013.01.008
- Melo-Filho, C. C., Dantas, R. F., Braga, R. C., Neves, B. J., Senger, M. R., Valente, W. C., et al. (2016). QSAR-driven discovery of novel chemical scaffolds active against *Schistosoma mansoni*. *J. Chem. Inf. Model.* 56, 1357–1372. doi: 10.1021/acs.jcim.6b00055
- Mihalkova, L., Huynh, T., and Mooney, R. J. (2007). “Mapping and revising markov logic networks for transfer learning,” in *Proceedings of the 22nd Association for the Advancement of Artificial Intelligence (AAAI) Conference. Artificial Intelligence*, (Vancouver, BC: AAAI), 608–614.
- Mitchell, J. B. O. (2014). Machine learning methods in chemoinformatics. *WIREs Comput. Mol. Sci.* 4, 468–481. doi: 10.1002/wcms.1183

- Neves, B. J., Dantas, R. F., Senger, M. R., Melo-Filho, C. C., Valente, W. C., De Almeida, A. C., et al. (2016). Discovery of new anti-schistosomal hits by integration of QSAR-Based virtual screening and high content screening. *J. Med. Chem.* 59, 7075–7088. doi: 10.1021/acs.jmedchem.5b02038
- Pan, S. J., and Yang, Q. (2010). A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191
- Pauwels, E., Stoven, V., and Yamanishi, Y. (2011). Predicting drug side-effect profiles: a chemical fragment-based approach. *BMC Bioinformatics* 12:169. doi: 10.1186/1471-2105-12-169
- Pereira, D. A., and Williams, J. A. (2007). Origin and evolution of high throughput screening. *Br. J. Pharmacol.* 152, 53–61. doi: 10.1038/sj.bjp.0707373
- Ponzoni, L., Sebastián-Pérez, V., Requena-Triguero, C., Roca, C., Martínez, M. J., Cravero, F., et al. (2017). Hybridizing feature selection and feature learning approaches in QSAR modeling for drug discovery. *Sci. Rep.* 7:2403. doi: 10.1038/s41598-017-02114-3
- Raies, A. B., and Bajic, V. B. (2016). *In silico* toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 6, 147–172. doi: 10.1002/wcms.1240
- Raina, R., Battle, A., Lee, H., Packer, B., and Ng, A. Y. (2007). “Self-taught learning: transfer learning from unlabeled data,” in *Proceedings of the 24th International Conference on Machine Learning, ICML 2007*, (Corvallis, OR: ACM), 759–766. doi: 10.1145/1273496.1273592
- Richter, L., and Ecker, G. F. (2015). Medicinal chemistry in the era of big data. *Drug Discov. Today* 14, 37–41. doi: 10.1016/j.ddtec.2015.06.001
- Rosenbaum, L., Dörr, A., Bauer, M. R., Boeckler, F. M., and Zell, A. (2013). Inferring multi-target qsar models with taxonomy-based multi-task learning. *J. Cheminform.* 5:33. doi: 10.1186/1758-2946-5-33
- Saha, B., Gupta, S., Phung, D., and Venkatesh, S. (2016). “Transfer learning for rare cancer problems via discriminative sparse gaussian graphical model,” in *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR)*, (Cancun: IEEE - Institute of Electrical and Electronics Engineers, Inc), 537–542. doi: 10.1109/ICPR.2016.7899689
- Satpal, S., and Sarawagi, S. (2007). “Domain adaptation of conditional probability models via feature subsetting,” in *Knowledge Discovery in Databases: PKDD 2007*, eds J. N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenic, and A. Skowron (Berlin: Springer), 224–235. doi: 10.1007/978-3-540-74976-9_23
- Scior, T., Medina-Franco, J. L., Do, Q. T., Martínez-Mayorga, K., Yunes Rojas, J. A., and Bernard, P. (2009). How to recognize and workaround pitfalls in QSAR studies: a critical review. *Curr. Med. Chem.* 16, 4297–4313. doi: 10.2174/092986709789578213
- Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature* 432, 862–865. doi: 10.1038/nature03197
- Sliwoski, G., Kothiwale, S. K., Meiler, J., Lowe, E. W., and Barker, E. L. (2014). Computational methods in drug discovery. *Pharmacol. Rev.* 66, 334–395. doi: 10.1124/pr.112.007336
- Tan, B., Song, Y., Zhong, E., and Yang, Q. (2015). “Transitive transfer learning,” in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (New York, NY: ACM), 1155–1164. doi: 10.1145/2783258.2783295
- Tresadern, G., Bemporad, D., and Howe, T. (2009). A comparison of ligand based virtual screening methods and application to corticotropin releasing factor 1 receptor. *J. Mol. Graph. Model.* 27, 860–870. doi: 10.1016/j.jmkgm.2009.01.003
- Tufts Center for the Study of Drug Development [CSDD] (2014). *Cost to Develop and Win Marketing Approval for a New Drug Is \$2.6 Billion*. Available at: http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study
- Turki, T., Wei, Z., and Wang, J. T. (2017). Transfer learning approaches to improve drug sensitivity prediction in multiple Myeloma patients. *IEEE Access* 5, 7381–7393. doi: 10.1109/ACCESS.2017.2696523
- Varnek, A., Gaudin, C., Marcou, G., Baskin, I., Pandey, A. K., and Tetko, I. V. (2009). Inductive transfer of knowledge: application of multi-task learning and feature net approaches to model tissue-air partition coefficients. *J. Chem. Inf. Model.* 49, 133–144. doi: 10.1021/ci8002914
- Walters, W. P., Stahl, M. T., and Murcko, M. A. (1998). Virtual screening—an overview. *Drug Discov. Today* 3, 160–178. doi: 10.1016/S1359-6446(97)01163-X
- Wang, T., Wu, M. B., Lin, J. P., and Yang, L. R. (2015). Quantitative structure-activity relationship: promising advances in drug discovery platforms. *Expert Opin. Drug Discov.* 11, 1–18. doi: 10.1517/17460441.2015.1083006
- Wang, Y., Bolton, E., Dracheva, S., Karapetyan, K., Shoemaker, B. A., Suzek, T. O., et al. (2009). An overview of the PubChem BioAssay resource. *Nucleic Acids Res.* 38, D255–D266. doi: 10.1093/nar/gkp965
- Zaccagnini, L., Brogi, S., Brindisi, M., Gemma, S., Chemi, G., Legname, G., et al. (2017). Identification of novel fluorescent probes preventing PrP Sc replication in prion diseases. *Eur. J. Med. Chem.* 127, 859–873. doi: 10.1016/j.ejmech.2016.10.064
- Zhang, L., Sedykh, A., Tripathi, A., Zhu, H., Afantitis, A., Mouchlis, V. D., et al. (2013). Identification of putative estrogen receptor-mediated endocrine disrupting chemicals using QSAR- and structure-based virtual screening approaches. *Toxicol. Appl. Pharmacol.* 272, 67–76. doi: 10.1016/j.taap.2013.04.032
- Zhang, L., Tan, J., Han, D., and Zhu, H. (2017). From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discov. Today* 22, 1680–1685. doi: 10.1016/j.drudis.2017.08.010
- Zhang, W., Chen, Y., Tu, S., Liu, F., and Qu, Q. (2016). “Drug side effect prediction through linear neighborhoods and multiple data source integration,” in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (Shenzhen: IEEE)*, 427–434. doi: 10.1109/BIBM.2016.7822555
- Zhao, L., Wang, W., Sedykh, A., and Zhu, H. (2017). Experimental errors in QSAR modeling sets: what we can do and what we cannot do. *ACS Omega* 2, 2805–2812. doi: 10.1021/acsomega.7b00274

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with the authors.

Copyright © 2018 Simões, Maltarollo, Oliveira and Honorio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.