JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

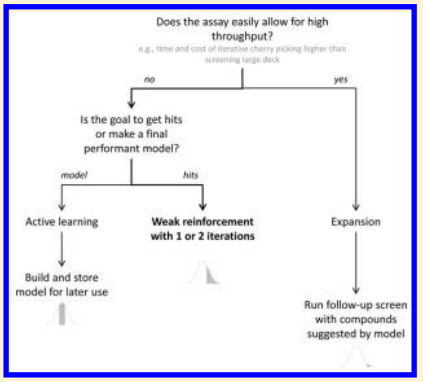
Experimental Design Strategy: Weak Reinforcement Leads to Increased Hit Rates and Enhanced Chemical Diversity

Mateusz Maciejewski,[†,§] Anne Mai Wassermann,[†,§] Meir Glick,[†] and Eugen Lounkine*[,†]

[†]Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

Supporting Information

ABSTRACT: High Throughput Screening (HTS) is a common approach in life sciences to discover chemical matter that modulates a biological target or phenotype. However, low assay throughput, reagents cost, or a flowchart that can deal with only a limited number of hits may impair screening large numbers of compounds. In this case, a subset of compounds is assayed, and *in silico* models are utilized to aid in iterative screening design, usually to expand around the found hits and enrich subsequent rounds for relevant chemical matter. However, this may lead to an overly narrow focus, and the diversity of compounds sampled in subsequent iterations may suffer. Active learning has been recently successfully applied in drug discovery with the goal of sampling diverse chemical space to improve model performance. Here we introduce a robust and straightforward iterative screening protocol based on naïve Bayes models. Instead of following up on the compounds with the highest scores in the *in silico* model, we pursue compounds with very low but positive values. This includes unique chemotypes of weakly active compounds that enhance the applicability domain of the model and increase the cumulative hit rates. We show in a retrospective application to 81 Novartis assays that this protocol leads to consistently higher compound and scaffold hit rates compared to a standard expansion around hits or an active learning approach. We recommend using the weak reinforcement strategy introduced herein for iterative screening workflows.



INTRODUCTION

Pharmaceutical industry and academia use QSAR models extensively to predict bioactivity profiles of compounds in an effort to select the right compounds for *in vitro* or *in vivo* studies. Both of these gains stem from the fact that activity models remove the necessity of bulk screening of large libraries of compounds for desired activity.[1,2] Instead, only a sample of the library is screened and based on the activity patterns in such a training set the model is built. The model reflects the correlation of molecular structure with mode of action (MoA), which can then be applied to the remaining portion of the library to focus on molecules that possess the structural features that, according to the applied model, predispose them to be active. With an increasing number of both industrial as well as academic institutions engaging in ever more complex screening campaigns, it becomes more important to design screening flowcharts that with small numbers of compounds, can cover many diverse actives. Iterative screening approaches are of particular interest, since they allow assembling of more and more focused libraries at each iteration step, learning from the screening results.

Due to their promise to increase the efficiency of the drug discovery process, data modeling strategies have been subject to numerous recent research efforts. These efforts have included studies to increase the effectiveness of the prepared models,[3] studies of different structural descriptors that may be used to capture the structures of molecules,[3] and studies of the orthogonality of these descriptors.[4] Alternative modeling strategies have been also investigated, such as activity-based models[5−7] or phenotype-based models.[8−10] Aside from the data being used, an extensive body of work relating to other aspects of modeling has been conceived, such as optimal modeling methods,[11] ranging from support vector machines,[12,13] through random forest models,[14,15] to naïve Bayes.[16,17]

In this work we focus on the experimental design of iterative screening guided by *in silico* models. We compare three iterative approaches: expansion, active learning, and weak active reinforcement. Expansion utilizes the initial screen to build a model that is applied to the remainder of the library. Active learning has gained popularity in recent years, although the concept is not new.[2] The core idea of active learning is to pick compounds outside of the applicability domain of a model, so that the model can be improved using new experimental evidence. Here we implement a straightforward active learning approach using naïve Bayes models. Finally, weak active reinforcement follows up on compounds that score weakly but nevertheless positively in a naïve Bayes model.

RESULTS

We have evaluated three iterative *in silico* screening strategies that utilize structure-based naïve Bayes classifiers (Figure 1): *expansion*, which is typically applied in iterative screening and serves as a control;[16,17] *active learning*,[2,18−20] for which we devised a robust and straightforward naïve Bayes strategy; and

Received: January 30, 2015
Published: April 27, 2015

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

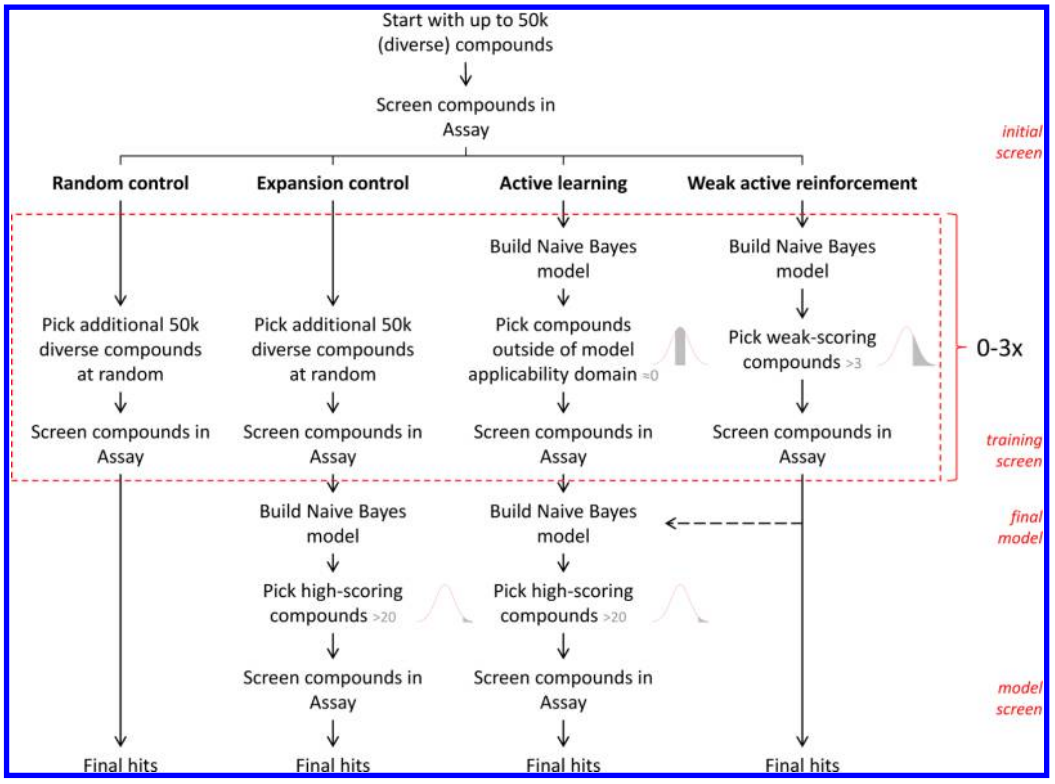


Figure 1. Iterative screening strategies. All strategies start off with up to 50k chemically diverse compounds. This is the initial screen. For the random and enrichment control, we have evaluated screens in which up to 150k additional compounds are screened. For comparison to the iterative approaches, we split the total number of compounds in multiples of 50k. However, in practice, all compounds would be screened in one go. The random control does not employ model building at all, while in the enrichment control a single model is built after screening up to 200k compounds. Active learning and "weak active reinforcement" constitute iterative screening strategies, where compounds are picked based on training results. Active learning employs a final model building step, similar to the enrichment control; by contrast, the "weak active reinforcement" strategy does not require this step.

weak reinforcement, which we introduce here. The analysis was carried out retrospectively for 81 high throughput assays historically run at Novartis.[21] We wanted to know which *in silico* strategy for iterative screening should be employed to maximize the cumulative hit rate (or cumulative scaffold hit rate), i.e., the total number of chemically diverse active compounds identified at the end of all efforts. As an additional control, we assessed the hit rate based on screening 50−200k randomly chosen compounds in increments of 50k compounds.

Expansion Control. Naïve Bayes models (NBM) combined with ECFP-4 fingerprints have become an established tool for predicting compound activity.[22] A typical application of NBM in iterative screening is to expand on found hits. A model is trained on actives vs inactives ("final model" in Figure 1). This model is then applied to the rest of the compound library, and high-scoring compounds (with a Bayes score >20)[23] are subsequently screened. We calculated three metrics to measure performance of this approach: cumulative hit rates, cumulative scaffold hit rates (Figure 2a), and area under the receiver-operator characteristic curves (ROCAUC) (Figure 2b).

Cumulative hit rates measure the fraction of screened compounds or unique scaffolds that were active. Cumulative hits rates for the expansion control were highest (4.7%) with expansion directly applied to the initial screen of 50k compounds. Screening additional compounds up front (up to 200k compounds) did not improve but rather lowered the cumulative hit rate with expansion. By comparison, the hit rates of the random control (screening 50k−200k random compounds) was stable at 1.5%. The same trends were visible for the cumulative scaffold hit rate (Figure 2a).

The ROCAUC measured how well the final model could rank-order compounds in the Novartis library that have not yet been screened. Values close to 1 mean that all true actives are ranked highest, whereas values close to 0.5 mean no enrichment over random selection. On average, the expansion models achieved ROCAUC values of 0.82 after the initial screen and performance steadily increased up to a median ROCAUC value of 0.85 when using 200k compounds to train the model after the third iteration (Table 1, Figure 2b).

Simple Active Learning Implementation Using Naïve Bayes. The central idea of active learning is to iteratively screen compounds that are outside of the model's applicability domain, at each iteration update the model using new data together with the data from previous iterations, and thereby improve the model (Figure 1).[2,18,19] For many models, this requires an explicit metric of the model applicability, with a subsequent selection of a diverse compound set outside of that domain. The advantage of NBM for active learning is that the same score that is used to predict activity can be immediately used to identify the model's applicability domain. NBM training assigns weights to molecular features based on their enrichment among active compounds. Features that are overrepresented among actives receive a large positive score, whereas underrepresented features receive a large negative score.[23] However, most features are not enriched and are assigned a score close to 0. Molecules are scored using a sum over scores of all features present in the compound. Features that were not available during model training do not contribute to the score. Hence, a Bayes score close to 0 is a direct measure of the model applicability domain: compounds that have many features unknown to the model will have a score near 0. The same score

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

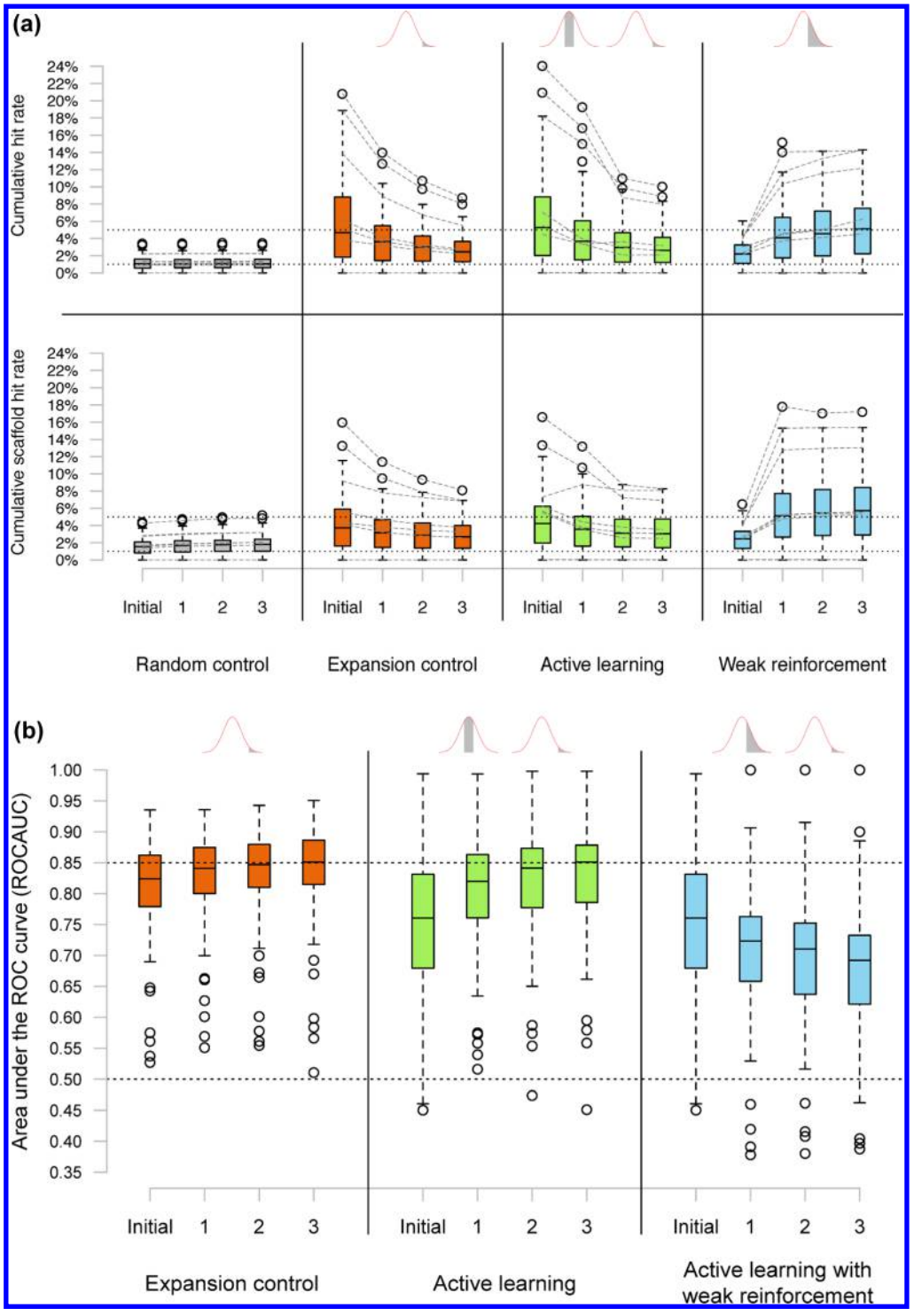


Figure 2. Performance comparison between three strategies and random control. Box plots (**a**) show cumulative compound and scaffold hit rates across 81 assays. Hit rates take into account all hits, i.e., from initial, training, and model screens. The lines show the performance of top three, median three, and bottom three individual assays. (**b**) Final model performance (area under the ROC curve box plots for 81 assays) reflects the ability of the final naïve Bayes model to rank-order hits. A value of 0.5 means no enrichment compared to random selection, and a value of 1 means that all hits are top-ranked by the naïve Bayes model score. For comparison, a final model was built based on the hits from the weak reinforcement strategy. In practice, this model is not utilized, because weak active reinforcement finds the majority of hits already during the iterative training process.

that is used to select actives can therefore directly be used for iterative active learning (Figure 1).

We evaluated the active learning workflow with up to three iterations based on a set of 50k biologically and chemically diverse compounds.[21] Due to the retrospective nature of our analysis, for individual assays, between 21,362 and 39,959 compounds that had been screened were selected for the initial screen (Methods, Supplementary Table S3).

Cumulative hit rates using active learning on these reduced sets were comparable or slightly higher than for the expansion control (5.3% after the initial screen; Figure 2a, Table 1, Supplementary Table S1). On average, active learning suggested

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

Table 1. Iterative Strategy Performance

expansion control			active learning			weak reinforcement		
ROCAUC[a]	Molecule HR[b] (%)	Scaffold HR[c] (%)	ROCAUC[a]	Molecule HR[b] (%)	Scaffold HR[c] (%)	ROCAUC[a]	Molecule HR[b] (%)	Scaffold HR[c] (%)
0.8239	4.68	3.73	0.7606	5.27	4.23	0.7606	5.27	4.23
0.8409	3.62	3.16	0.8201	3.66	5.13	0.7234	4.60	5.13
0.8469	3.00	2.86	0.8414	2.96	5.44	0.7106	5.06	5.44
0.8512	2.44	2.68	0.851	2.62	5.72	0.6921	5.61	5.72

[a]ROCAUC: Area under receiver operating characteristic curve for the final model based on library compounds not yet screened in assay. [b]Molecule HR: median cumulative hit rate across all assays. [c]Scaffold HR: median cumulative scaffold hit rate across all assays.

between 25k (first iteration) and 15k (third iteration) new compounds to screen. ROCAUC scores of the final model trained on all screened compounds increased steadily from 0.75 to 0.85 with increasing numbers of compounds screened (Figure 2b). However, on average, only half of the compounds had to be screened in the active learning scenario to achieve a comparable performance to models built by the expansion control. For example, for the second iteration, 163,295 compounds have been screened in the random control: 150,000 initially plus 13,295 on average suggested by the final model. By contrast, only 82,263 compounds have been screened in active learning: on average, 80,144 from the model training stage plus 2,020 from the final model (Supplementary Table S1).

Weak Reinforcement Strategy. Iterative screening approaches, in particular combined with *in silico* expansion,[21,24] have been shown to successfully recover a large proportion of hits without having to screen the entire compound collection (full deck). Each iteration is primarily designed to find more hits. This is in contrast to the active learning strategy, which deliberately explores diverse chemical space to gather more data for model training.

We wanted to adapt the expansion strategy to this iterative setting. At each iteration, only compounds that were predicted to be active would be screened (Figure 1, "Weak active reinforcement"). For the final screening model we used an empirical cutoff of 20 for the Bayes score.[23] However, active reinforcement did not identify enough compounds scoring that high based on the reduced sets used for evaluation. For 79 out of 81 assays less than 3,000 compounds scored high enough to meet the bar of a Bayes score of 20 or higher (Supplementary Table S2). Hence, a meaningful iterative strategy with top scoring compounds was impossible to realize with this stringent cutoff. Therefore, we lowered the threshold for the training screen models to a Bayes score of 3, which provided >3,000 compounds for more than half of all assays after the second iteration.

In contrast to the expansion control and active learning, cumulative hit rates using the weak reinforcement strategy increased for molecules as well as scaffolds with increasing iterations (Table 1, Supplementary Table S1, Figure 2a). For example, after three iterations comparable numbers of unique scaffolds (approx. 53k, Supplementary Table S1) were screened using active learning and weak active reinforcement. However, the number of unique scaffolds that turned out to be active almost doubled (1674 unique scaffold hits for active learning vs 3007 unique scaffold hits for weak active reinforcement).

Interestingly, when trying to build an additional model on top of these hits, model performance deteriorated with more iterations (Figure 2b) and suggested, on average, less than 100 additional compounds (Supplementary Table S2).

In summary, "weak active reinforcement" found a larger number of more diverse hits during the training procedure than active learning or the expansion control for each compound screened.

DISCUSSION

Naïve Bayes Active Learning. Naïve Bayes models have been established as useful tools for virtual screening[16] and target prediction.[25] Moreover, both feature and compound scores are readily interpretable: a high negative or positive score for a feature means that the feature is under- or overrepresented in a set of reference (usually active) compounds.[23] Here we capitalized on this readily interpretable nature of the naïve Bayes score. Compounds that have an absolute score around 0, in their majority are dominated by features that the model either did not see at all or features that are equally distributed among actives and inactives. Since in an iterative screening scenario the initial screen will cover only a small part of the total chemical space available for screening, it is safe to assume that the majority of newly encountered compounds will be outside of the model's applicability domain, and the rest will get a highly negative or highly positive score. Therefore, utilizing the Bayes score directly enabled us to select compounds outside of the model's applicability domain (scores close to 0) for active learning.

We showed that this very simple strategy produces models whose performance increased with the number of active learning iterations, corroborating previous observations with active learning for target activity prediction.[18] No need for an external measure of the model's applicability domain was necessary. The model performance is comparable to the expansion control, which was trained on more than twice as many compounds. Thus, our naïve Bayes implementation of active learning[2] successfully samples diverse chemical space not yet captured by the screen.

Only Weak Active Reinforcement Is Possible. Usually, building a model during an iterative screen serves the purpose of finding additional hits. We found, however, that selecting high-scoring compounds is nearly impossible beyond a single iteration: the very first model that is trained finds all hits that can be found with that model, and for the subsequent iterations not enough compounds are available for screening. In other words, this approach results in limited diversity that exhausts the chemical matter in the screening collection predicted to the active by the machine learning model. Therefore, we explored a weak active reinforcement approach, where we included compounds that had a very small but, nevertheless, positive Bayes score.

Cumulative Hit Rates Are Highest Using Weak Reinforcement. From the active learning perspective, the goal is to train the best possible model with the least spent resources; we have shown that our simple naïve Bayes active learning protocol satisfies this goal. However, what we were interested in even more was to increase the overall probability of success in an iterative screening setting, yielding as many starting

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

points for further validation experiments as possible. Two criteria had to be satisfied: first, we wanted to increase not just the number of hits in each iteration but the cumulative hit rate, i.e. the total number of hits over all iterations divided by the total number of compounds screened over all iterations; second, we required that the identified hits were diverse, providing many different starting points to go into validation experiments. To assess diversity we used unique scaffold counts and the derived cumulative scaffold hit rates. Other properties that may define a hit worth following up, such as solubility and chemical attractiveness, could be applied at the validation step after iterative screening is done and thus are beyond the scope of this study. As medchem efforts often are guided by clustering of compounds around common scaffolds, we used scaffold uniqueness to assess diversity and overall quality of the retrieved hit list.

Hit rates in active learning were dominated by the final, well-performing model. By design, the training rounds had hit rates comparable to random screening: active learning aims at sampling a diverse chemical space. The weak reinforcement strategy showed the exact opposite characteristic: the final model we built for comparison (Figure 2b) did not add anything new, but the hit rates during training were up to twice as high as with active learning or the baseline. Looking at the entire workflow, weak reinforcement yielded overall highest molecule and scaffold hit rates. Moreover, the last step of building a final model was not necessary in weak active reinforcement.

Weak Reinforcement Samples SAR-Relevant Chemical Space. Naïve Bayes models are classification models, rather than quantitative predictions. As such, the Bayes score, which we utilized here, reflects classification confidence rather than potency. Large positive values mean high confidence in activity, and large negative values mean high confidence that the compound is inactive. This made active learning using Bayes scores straightforward because active learning seeks to test compounds on which the model is not able to make a confident judgment (Bayes score ~0). By contrast, our weak reinforcement strategy emphasizes compounds predicted to be active with a wide confidence range. It may be surprising that screening low confidence compounds enriches for actives. Why does weak reinforcement work and how is low confidence related to the higher diversity of hits that we find?

We believe the key to weak reinforcement's success lies in the way individual molecular features contribute, in naïve Bayes models, to the overall score of a compound, and how this relates to the exploration of chemical space by the model (Figure 3a). A low positive score arises when at least some features of the compound are enriched in actives vs inactives, even if other parts of the molecule are characteristic of inactives or unknown to the model. Although the model is not confident about the activity prediction, with each iteration the size of the positive class increases as more diverse chemical matter is added. Thus, weak active reinforcement explores the structure−activity relationship (SAR) landscape around hits in a permissive manner that allows for finding diverse chemical matter, while still "holding on" to a core pharmacophore defined by few enriched features. In contrast, active learning is not designed to expand the active class and thus adds both active and inactive compounds. Active learning and weak active reinforcement may appear similar in their protocols (Figure 1), but the crucial difference of identifying hits after sampling (active learning) versus during iterations (weak reinforcement) leads to very distinct exploration paths of chemical space and SAR.

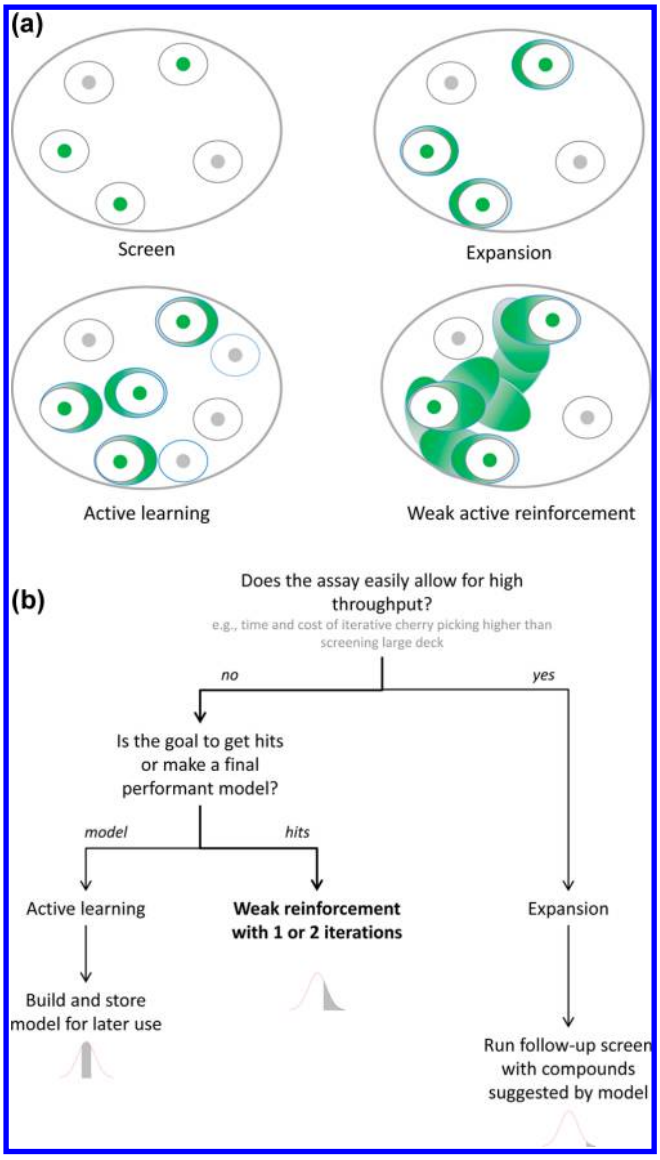


Figure 3. Weak reinforcement assessment for iterative screening. (**a**) Schematic of putative exploration of chemical space by the three methods. Screening identifies hits in chemical space (green dots) as well as inactive compounds (gray dots); Expansion explores chemical space around the hits (green ovals). Active learning, during iterations, finds diverse hits by testing under-sampled regions of chemical space; this is followed by an expansion analogous to enrichment. Weak active reinforcement explores a larger chemical space around the hits already during the iterations; thus it captures more hits but also diverse chemotypes. (**b**) While efficient sampling of chemical space can be useful in systematic profiling efforts, the weak reinforcement strategy introduced here has proven best for iterative screening.

This also explains why the final model built off the confirmed active learning predictions performs so much better on the remainder of the library than a model based on active weak reinforcement hits (Figure 2b, Figure 3a). Active learning does not exhaust the actives in the library, and there are many actives left to find around the seeds picked by active learning. By contrast, weak active reinforcement explores the space around any such seeds during iterative screening. However, because of the low confidence cutoff, diversity of compounds still is ensured. By the end of iteration 3, it has found everything it possibly can find, and thus building a final model is not necessary.

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

Given this relationship between low confidence Bayes scores and SAR exploration around active compounds, one can draw a parallel between low naïve Bayes model confidence and screening weakly active hits as a follow-up (assuming that eventually the activity landscape leads to weaker compounds as we move away from actives). To emphasize this connection, we termed the method "weak active reinforcement", even though we use a classifier and confidence scores rather than quantitative predictions.

CONCLUSION

When deciding on the screening format (iterative or not), both time and resources have to be considered. Some assays (e.g., certain biochemical assays) can be easily run in full deck, screening more than 1 million compounds. Other formats, such as complex phenotypic screens, or electrophysiology assays, can only handle a limited number of compounds at a time. Our expansion control results suggest that a model built on top of a large number of compounds will, with only one additional screen, increase the hit rate substantially, corroborating earlier observations for this kind of expansion.[21] However, if an iterative strategy is more feasible because the time cost is offset by the savings of screening fewer compounds in one round, the weak reinforcement strategy combines sampling the chemical space and expanding around existing hits.

There may be situations where the goal is to have a performant model after iterative screening, e.g., in profiling efforts that aim at building a safety pharmacology model.[26] In these cases, efficient sampling of chemical space is key, and we recommend applying active learning. However, for the particular application to iterative screening workflows, we recommend the use of the weak reinforcement strategy introduced herein (Figure 3b).

METHODS

Data Sets and Study Design. We tested here the performance of two iterative approaches for hit expansion using NBM,[23] "active learning", and "weak reinforcement" and compared it to the performance of a "random control" baseline (Figure 1). For screening, we used primary assays with >1 million tested compounds (81 assays).[21] For each assay the activities of the screened compounds were expressed as Z-scores. Compounds were considered to be active in the assays in which their Z-scores were >3 (activation assays) or < −3 (inhibition assays). Binary naïve Bayesian models of assay activity were built using ECFP-4 descriptors,[27] as implemented in Pipeline Pilot.[28]

Evaluation of iterative screening strategies was carried out retrospectively; for brevity, we use the term "screen" to indicate that we looked up historic activities for a set of compounds. Accordingly, hit rates were calculated based on the binary activity classification. Both individual and cumulative hit rates were calculated; individual hit rates were defined for each iteration step; cumulative hit rates took into account all tested and active compounds throughout the entire workflow. Similarly, scaffold hit rates were defined based on unique Murcko scaffolds of compounds (calculated with Pipeline Pilot[28]). Additionally, ROCAUC scores were calculated for the final model applied to all compounds for which historic assay data were available. ROCAUC scores close to 0.5 indicate no improvement of compound selection compared to random selection, whereas a score of 1.0 indicates best possible model performance (all active compounds are ranked higher than inactive ones). ROCAUC integration was carried out using the *trapz* method from the SciPy[29] library.

Iterative Screening Strategies. We used a set of 50,000 molecules optimized for diversity under a variety of structural, chemical, biological, physicochemical, and other parameters.[21] Given the retrospective character of our study, only a subset of these compounds was screened in all assays. Hence, for each of the 81 assays considered in this study we screened between 21,362 and 39,959 molecules from this diverse set (Supplementary Table S3) in the initial screen.

The following strategy was carried out for each assay independently. Between zero and three iterations of active learning was carried out to derive a *final model* (Figure 1, bottom). In each iteration, we trained a naïve Bayes model on active versus inactive compounds. We applied this model to all compounds in the library and selected compounds that had an absolute naïve Bayes score ≤0.5. If more than 50,000 compounds (for which historic assay data were available) were identified, we calculated Murcko scaffolds and selected a diverse set of compounds by first selecting one representative of each scaffold and then repeating the selection until we reached 50,000 compounds (Figure 1, left). This set was then screened and hits identified (*training screen* in Figure 1). The final model was derived using all screened and hit compounds from the initial screen and iterations (Figure 1, bottom). Compounds scoring ≥20 were then screened.

The weak reinforcement strategy was carried out analogous to the active learning strategy, with the exception that the naïve Bayes score cutoff for the model built during the iterative *training screen* (Figure 1) was ≥3.

The controls added 50,000 compounds in each training screen without any prioritization (Figure 1, left). The final model in the expansion control was calculated and evaluated analogous to the other strategies.

ASSOCIATED CONTENT

Supporting Information

Supplementary Tables 1−3. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00054.

AUTHOR INFORMATION

Corresponding Author

*E-mail: eugen.lounkine@novartis.com.

Present Address

§Pfizer Inc., 610 Main Street, Cambridge, Massachusetts 02139, USA.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

M.M. and A.M.W. were supported by the NIBR Presidential Postdoctoral Fellowship.

REFERENCES

(1) Bajorath, J. Integration of Virtual and High-Throughput Screening. *Nat. Rev. Drug Discovery* **2002**, *1*, 882−894.

(2) Reker, D.; Schneider, G. Active-Learning Strategies in Computer-Assisted Drug Discovery. *Drug Discovery Today* **2015**, *20*, 458−465.

(3) Bender, A.; Glen, R. C. A Discussion of Measures of Enrichment in Virtual Screening: Comparing the Information Content of Descriptors with Increasing Levels of Sophistication. *J. Chem. Inf. Model.* **2005**, *45*, 1369−1375.

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962

(4) Bender, A.; Lenkins, J. L.; Scheiber, J.; Sukuru, S. C. K.; Glick, M.; Davies, J. W. How Similar Are Similarity Searching Methods? A Principal Component Analysis of Molecular Descriptor Space. *J. Chem. Inf. Model.* **2009**, *49*, 108−119.
(5) Petrone, P. M.; Simms, B.; Nigsch, F.; Lounkine, E.; Kutchukian, P.; Cornett, A.; Deng, Z.; Davies, J. W.; Jenkins, J. L.; Glick, M. Rethinking Molecular Similarity: Comparing Compounds on the Basis of Biological Activity. *ACS Chem. Biol.* **2012**, *7*, 1399−1409.
(6) Wassermann, A. M.; Lounkine, E.; Urban, L.; Whitebread, S.; Chen, S.; Hughes, K.; Guo, H.; Kutlina, E.; Fekete, A.; Klumpp, M.; Glick, M. A Screening Pattern Recognition Method Finds New and Divergent Targets for Drugs and Natural Products. *ACS Chem. Biol.* **2014**, *9*, 1622−1631.
(7) Nguyen, H. P.; Koutsoukas, A.; Mohd Fauzi, F.; Drakakis, G.; Maciejewski, M.; Glen, R. C.; Bender, A. Diversity Selection of Compounds Based on "Protein Affinity Fingerprints" Improves Sampling of *Bioactive* Chemical Space. *Chem. Biol. Drug Des.* **2013**, *82*, 252−266.
(8) Scheiber, J.; Jenkins, J. L.; Sukuru, S. C. K.; Bender, A.; Mikhailov, D.; Milik, M.; Azzaoui, K.; Whitebread, S.; Hamon, J.; Urban, L.; Glick, M.; Davies, J. W. Mapping Adverse Drug Reactions in Chemical Space. *J. Med. Chem.* **2009**, *52*, 3103−3107.
(9) Young, D. W.; Bender, A.; Hoyt, J.; McWhinnie, E.; Chirn, G.-W.; Tao, C. Y.; Tallarico, J. A.; Labow, M.; Jenkins, J. L.; Mitchison, T. J.; Feng, Y. Integrating High-Content Screening and Ligand-Target Prediction to Identify Mechanism of Action. *Nat. Chem. Biol.* **2007**, *4*, 59−68.
(10) Laggner, C.; Kokel, D.; Setola, V.; Tolia, A.; Lin, H.; Irwin, J. J.; Keiser, M. J.; Cheung, C. Y. J.; Minor, D. L.; Roth, B. L.; Peterson, R. T.; Shoichet, B. K. Chemical Informatics and Target Identification in a Zebrafish Phenotypic Screen. *Nat. Chem. Biol.* **2011**, *8*, 144−146.
(11) Mitchell, J. B. O. Machine Learning Methods in Chemoinformatics. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 468−481.
(12) Jorissen, R. N.; Gilson, M. K. Virtual Screening of Molecular Databases Using a Support Vector Machine. *J. Chem. Inf. Model.* **2005**, *45*, 549−561.
(13) Heikamp, K.; Bajorath, J. Prediction of Compounds with Closely Related Activity Profiles Using Weighted Support Vector Machine Linear Combinations. *J. Chem. Inf. Model.* **2013**, *53*, 791−801.
(14) Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J. C.; Sheridan, R. P.; Feuston, B. P. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J. Chem. Inf. Model.* **2003**, *43*, 1947−1958.
(15) Bruce, C. L.; Melville, J. L.; Pickett, S. D.; Hirst, J. D. Contemporary QSAR Classifiers Compared. *J. Chem. Inf. Model.* **2007**, *47*, 219−227.
(16) Bender, A. Bayesian Methods in Virtual Screening and Chemical Biology. *Methods Mol. Biol.* **2011**, *672*, 175−196.
(17) Nigsch, F.; Bender, A.; Jenkins, J. L.; Mitchell, J. B. O. Ligand-Target Prediction Using Winnow and Naive Bayesian Algorithms and the Implications of Overall Performance Statistics. *J. Chem. Inf. Model.* **2008**, *48*, 2313−2325.
(18) Kangas, J. D.; Naik, A. W.; Murphy, R. F. Efficient Discovery of Responses of Proteins to Compounds Using Active Learning. *BMC Bioinf.* **2014**, *15*, 143.
(19) Naik, A. W.; Kangas, J. D.; Langmead, C. J.; Murphy, R. F. Efficient Modeling and Active Learning Discovery of Biological Responses. *PLoS One* **2013**, *8*, e83996.
(20) Pauwels, E.; Lajaunie, C.; Vert, J.-P. A Bayesian Active Learning Strategy for Sequential Experimental Design in Systems Biology. *BMC Syst. Biol.* **2014**, *8*, 1−11.
(21) Petrone, P. M.; Wassermann, A. M.; Lounkine, E.; Kutchukian, P.; Simms, B.; Jenkins, J.; Selzer, P.; Glick, M. Biodiversity of Small Molecules − a New Perspective in Screening Set Selection. *Drug Discovery Today* **2013**, *18*, 674−680.
(22) Koutsoukas, A.; Lowe, R.; KalantarMotamedi, Y.; Mussa, H. Y.; Klaffke, W.; Mitchell, J. B. O.; Glen, R. C.; Bender, A. In Silico Target Predictions: Defining a Benchmarking Dataset and Comparison of Performance of the Multiclass Naïve Bayes and Parzen-Rosenblatt Window. *J. Chem. Inf. Model.* **2013**, *53*, 1957−1966.
(23) Lounkine, E.; Kutchukian, P. S.; Glick, M. Chemometric Applications of Naïve Bayesian Models in Drug Discovery. In *Chemoinformatics for Drug Discovery*; Bajorath, J., Ed.; John Wiley & Sons, Inc.: 2013; pp 131−148.
(24) Davies, J. W.; Glick, M.; Jenkins, J. L. Streamlining Lead Discovery by Aligning in Silico and High-Throughput Screening. *Curr. Opin. Chem. Biol.* **2006**, *10*, 343−351.
(25) Nidhi; Glick, M.; Davies, J. W.; Jenkins, J. L. Prediction of Biological Targets for Compounds Using Multiple-Category Bayesian Models Trained on Chemogenomics Databases. *J. Chem. Inf. Model.* **2006**, *46*, 1124−1133.
(26) Wang, S.; Li, Y.; Xu, L.; Li, D.; Hou, T. Recent Developments in Computational Prediction of hERG Blockage. *Curr. Top. Med. Chem.* **2013**, *13*, 1317−1326.
(27) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742−754.
(28) PipelinePilot, Version 8.5, Accelrys: San Diego. http://accelrys.com/products/pipeline-pilot/ (accessed Apr 21, 2015).
(29) Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open Source Scientific Tools for Python. http://www.scipy.org/ (accessed Apr 21, 2015).

DOI: 10.1021/acs.jcim.5b00054
J. Chem. Inf. Model. 2015, 55, 956−962