Supplementary Information: Analysis of Uncertainty of Neural Fingerprint-based Models

Christian W. Feldmann[1,+], Jochen Sieg[1,+], and Miriam Mathea[1,+,*]

[1]BASF SE, Ludwigshafen, Germany
[*]miriam.mathea@basf.com
[+]these authors contributed equally to this work

Mai 2024

Table S1. Performance of all endpoints. The Brier score and balanced accuracy of all endpoints for Chemprop, Morgan FP + RF, and Neural FP + RF are shown using the random split. Morgan FP corresponds to Morgan count fingerprints.

endpoint	model	Balanced accuracy	Brier score
APR_HepG2_CellLoss_72h_dn	Chemprop	0.70 ± 0.02	0.22 ± 0.02
APR_HepG2_CellLoss_72h_dn	Morgan FP + RF	0.68 ± 0.04	0.20 ± 0.02
APR_HepG2_CellLoss_72h_dn	Neural FP + RF	0.71 ± 0.03	0.19 ± 0.01
ATG_NRF2_ARE_CIS_up	Chemprop	0.73 ± 0.01	0.19 ± 0.01
ATG_NRF2_ARE_CIS_up	Morgan FP + RF	0.70 ± 0.01	0.18 ± 0.00
ATG_NRF2_ARE_CIS_up	Neural FP + RF	0.71 ± 0.01	0.18 ± 0.00
ATG_PXRE_CIS_up	Chemprop	0.74 ± 0.01	0.19 ± 0.01
ATG_PXRE_CIS_up	Morgan FP + RF	0.73 ± 0.01	0.19 ± 0.01
ATG_PXRE_CIS_up	Neural FP + RF	0.74 ± 0.01	0.18 ± 0.00
BSK_3C_HLADR_down	Chemprop	0.69 ± 0.04	0.21 ± 0.02
BSK_3C_HLADR_down	Morgan FP + RF	0.65 ± 0.03	0.19 ± 0.01
BSK_3C_HLADR_down	Neural FP + RF	0.67 ± 0.03	0.19 ± 0.01
BSK_3C_Proliferation_down	Chemprop	0.70 ± 0.01	0.21 ± 0.01
BSK_3C_Proliferation_down	Morgan FP + RF	0.68 ± 0.02	0.19 ± 0.01
BSK_3C_Proliferation_down	Neural FP + RF	0.72 ± 0.03	0.19 ± 0.01
BSK_3C_SRB_down	Chemprop	0.68 ± 0.02	0.20 ± 0.02
BSK_3C_SRB_down	Morgan FP + RF	0.61 ± 0.02	0.18 ± 0.01
BSK_3C_SRB_down	Neural FP + RF	0.66 ± 0.04	0.18 ± 0.01
BSK_3C_Vis_down	Chemprop	0.66 ± 0.03	0.19 ± 0.02
BSK_3C_Vis_down	Morgan FP + RF	0.60 ± 0.03	0.18 ± 0.01
BSK_3C_Vis_down	Neural FP + RF	0.63 ± 0.03	0.17 ± 0.01
BSK_4H_Eotaxin3_down	Chemprop	0.64 ± 0.04	0.20 ± 0.01
BSK_4H_Eotaxin3_down	Morgan FP + RF	0.56 ± 0.02	0.19 ± 0.01
BSK_4H_Eotaxin3_down	Neural FP + RF	0.62 ± 0.03	0.18 ± 0.00
BSK_CASM3C_Proliferation_down	Chemprop	0.65 ± 0.02	0.21 ± 0.01
BSK_CASM3C_Proliferation_down	Morgan FP + RF	0.58 ± 0.02	0.18 ± 0.01
BSK_CASM3C_Proliferation_down	Neural FP + RF	0.61 ± 0.01	0.18 ± 0.00
BSK_LPS_VCAM1_down	Chemprop	0.68 ± 0.02	0.19 ± 0.01
BSK_LPS_VCAM1_down	Morgan FP + RF	0.61 ± 0.02	0.17 ± 0.00
BSK_LPS_VCAM1_down	Neural FP + RF	0.63 ± 0.01	0.17 ± 0.00
BSK_SAg_CD38_down	Chemprop	0.68 ± 0.03	0.19 ± 0.02
BSK_SAg_CD38_down	Morgan FP + RF	0.60 ± 0.03	0.18 ± 0.01
BSK_SAg_CD38_down	Neural FP + RF	0.64 ± 0.02	0.17 ± 0.01
BSK_SAg_CD40_down	Chemprop	0.67 ± 0.04	0.19 ± 0.02
BSK_SAg_CD40_down	Morgan FP + RF	0.60 ± 0.02	0.18 ± 0.01
BSK_SAg_CD40_down	Neural FP + RF	0.63 ± 0.03	0.17 ± 0.01
BSK_SAg_Proliferation_down	Chemprop	0.69 ± 0.01	0.21 ± 0.01
BSK_SAg_Proliferation_down	Morgan FP + RF	0.65 ± 0.02	0.19 ± 0.01
BSK_SAg_Proliferation_down	Neural FP + RF	0.67 ± 0.04	0.19 ± 0.02
BSK_hDFCGF_CollagenIII_down	Chemprop	0.72 ± 0.01	0.18 ± 0.01
BSK_hDFCGF_CollagenIII_down	Morgan FP + RF	0.59 ± 0.01	0.17 ± 0.01
BSK_hDFCGF_CollagenIII_down	Neural FP + RF	0.63 ± 0.03	0.17 ± 0.01
BSK_hDFCGF_Proliferation_down	Chemprop	0.70 ± 0.02	0.22 ± 0.02
BSK_hDFCGF_Proliferation_down	Morgan FP + RF	0.67 ± 0.03	0.20 ± 0.01
BSK_hDFCGF_Proliferation_down	Neural FP + RF	0.69 ± 0.02	0.20 ± 0.00
CEETOX_H295R_11DCORT_dn	Chemprop	0.58 ± 0.02	0.23 ± 0.01
CEETOX_H295R_11DCORT_dn	Morgan FP + RF	0.58 ± 0.02	0.22 ± 0.01
CEETOX_H295R_11DCORT_dn	Neural FP + RF	0.64 ± 0.05	0.21 ± 0.01
CEETOX_H295R_ANDR_dn	Chemprop	0.60 ± 0.05	0.25 ± 0.03
CEETOX_H295R_ANDR_dn	Morgan FP + RF	0.58 ± 0.03	0.22 ± 0.02
CEETOX_H295R_ANDR_dn	Neural FP + RF	0.62 ± 0.07	0.21 ± 0.03
TOX21_ARE_BLA_agonist_ratio	Chemprop	0.67 ± 0.01	0.15 ± 0.01
TOX21_ARE_BLA_agonist_ratio	Morgan FP + RF	0.59 ± 0.01	0.14 ± 0.00
TOX21_ARE_BLA_agonist_ratio	Neural FP + RF	0.60 ± 0.01	0.14 ± 0.00
TOX21_TR_LUC_GH3_Antagonist	Chemprop	0.75 ± 0.01	0.12 ± 0.00
TOX21_TR_LUC_GH3_Antagonist	Morgan FP + RF	0.67 ± 0.01	0.11 ± 0.00
TOX21_TR_LUC_GH3_Antagonist	Neural FP + RF	0.69 ± 0.01	0.11 ± 0.00

Table S2. Comparison of calibrated and uncalibrated RF models using the neural fingerprint. Shown are the balanced accuracy, Brier score, and log loss of the models on the 19 ToxCast endpoint data sets over a **cluster** split.

endpoint	metric	Calibrated RF	Neural FP RF
APR_HepG2_CellLoss_72h_dn	BA	0.69 ± 0.06	0.69 ± 0.06
APR_HepG2_CellLoss_72h_dn	Brier score	0.20 ± 0.01	0.20 ± 0.02
APR_HepG2_CellLoss_72h_dn	Log loss	0.59 ± 0.03	0.57 ± 0.04
ATG_NRF2_ARE_CIS_up	BA	0.68 ± 0.02	0.69 ± 0.02
ATG_NRF2_ARE_CIS_up	Brier score	0.19 ± 0.01	0.19 ± 0.02
ATG_NRF2_ARE_CIS_up	Log loss	0.56 ± 0.03	0.56 ± 0.04
ATG_PXRE_CIS_up	BA	0.71 ± 0.02	0.71 ± 0.03
ATG_PXRE_CIS_up	Brier score	0.19 ± 0.01	0.19 ± 0.01
ATG_PXRE_CIS_up	Log loss	0.58 ± 0.02	0.57 ± 0.02
BSK_3C_HLADR_down	BA	0.65 ± 0.03	0.63 ± 0.03
BSK_3C_HLADR_down	Brier score	0.21 ± 0.01	0.21 ± 0.01
BSK_3C_HLADR_down	Log loss	0.60 ± 0.02	0.60 ± 0.02
BSK_3C_Proliferation_down	BA	0.64 ± 0.02	0.64 ± 0.02
BSK_3C_Proliferation_down	Brier score	0.21 ± 0.01	0.21 ± 0.01
BSK_3C_Proliferation_down	Log loss	0.61 ± 0.03	0.61 ± 0.03
BSK_3C_SRB_down	BA	0.61 ± 0.05	0.63 ± 0.05
BSK_3C_SRB_down	Brier score	0.19 ± 0.02	0.19 ± 0.02
BSK_3C_SRB_down	Log loss	0.56 ± 0.04	0.55 ± 0.04
BSK_3C_Vis_down	BA	0.57 ± 0.02	0.59 ± 0.04
BSK_3C_Vis_down	Brier score	0.18 ± 0.01	0.18 ± 0.01
BSK_3C_Vis_down	Log loss	0.54 ± 0.03	0.54 ± 0.03
BSK_4H_Eotaxin3_down	BA	0.56 ± 0.03	0.57 ± 0.03
BSK_4H_Eotaxin3_down	Brier score	0.19 ± 0.01	0.19 ± 0.01
BSK_4H_Eotaxin3_down	Log loss	0.56 ± 0.01	0.55 ± 0.02
BSK_CASM3C_Proliferation_down	BA	0.54 ± 0.02	0.58 ± 0.02
BSK_CASM3C_Proliferation_down	Brier score	0.20 ± 0.02	0.20 ± 0.02
BSK_CASM3C_Proliferation_down	Log loss	0.58 ± 0.04	0.58 ± 0.04
BSK_LPS_VCAM1_down	BA	0.55 ± 0.02	0.55 ± 0.02
BSK_LPS_VCAM1_down	Brier score	0.19 ± 0.01	0.19 ± 0.01
BSK_LPS_VCAM1_down	Log loss	0.56 ± 0.02	0.55 ± 0.02
BSK_SAg_CD38_down	BA	0.58 ± 0.03	0.59 ± 0.04
BSK_SAg_CD38_down	Brier score	0.19 ± 0.00	0.19 ± 0.00
BSK_SAg_CD38_down	Log loss	0.55 ± 0.01	0.56 ± 0.02
BSK_SAg_CD40_down	BA	0.56 ± 0.04	0.58 ± 0.04
BSK_SAg_CD40_down	Brier score	0.19 ± 0.01	0.19 ± 0.01
BSK_SAg_CD40_down	Log loss	0.56 ± 0.03	0.55 ± 0.02
BSK_SAg_Proliferation_down	BA	0.60 ± 0.03	0.63 ± 0.02
BSK_SAg_Proliferation_down	Brier score	0.21 ± 0.01	0.21 ± 0.01
BSK_SAg_Proliferation_down	Log loss	0.61 ± 0.02	0.60 ± 0.01
BSK_hDFCGF_CollagenIII_down	BA	0.54 ± 0.02	0.59 ± 0.02
BSK_hDFCGF_CollagenIII_down	Brier score	0.19 ± 0.02	0.19 ± 0.02
BSK_hDFCGF_CollagenIII_down	Log loss	0.56 ± 0.04	0.55 ± 0.04
BSK_hDFCGF_Proliferation_down	BA	0.65 ± 0.03	0.68 ± 0.04
BSK_hDFCGF_Proliferation_down	Brier score	0.21 ± 0.01	0.21 ± 0.02
BSK_hDFCGF_Proliferation_down	Log loss	0.61 ± 0.03	0.60 ± 0.04
CEETOX_H295R_11DCORT_dn	BA	0.57 ± 0.02	0.59 ± 0.02
CEETOX_H295R_11DCORT_dn	Brier score	0.22 ± 0.02	0.22 ± 0.02
CEETOX_H295R_11DCORT_dn	Log loss	0.63 ± 0.04	0.64 ± 0.06
CEETOX_H295R_ANDR_dn	BA	0.54 ± 0.04	0.61 ± 0.03
CEETOX_H295R_ANDR_dn	Brier score	0.22 ± 0.01	0.22 ± 0.01
CEETOX_H295R_ANDR_dn	Log loss	0.63 ± 0.03	0.62 ± 0.02
TOX21_ARE_BLA_agonist_ratio	BA	0.50 ± 0.00	0.57 ± 0.01
TOX21_ARE_BLA_agonist_ratio	Brier score	0.16 ± 0.02	0.15 ± 0.01
TOX21_ARE_BLA_agonist_ratio	Log loss	0.48 ± 0.04	0.46 ± 0.03
TOX21_TR_LUC_GH3_Antagonist	BA	0.63 ± 0.03	0.62 ± 0.03
TOX21_TR_LUC_GH3_Antagonist	Brier score	0.13 ± 0.01	0.13 ± 0.01
TOX21_TR_LUC_GH3_Antagonist	Log loss	0.41 ± 0.03	0.41 ± 0.02

Table S3. Comparison of calibrated and uncalibrated RF models using the neural fingerprint. The balanced accuracy, Brier score, and log loss of the models on the 19 ToxCast endpoint data sets over a **random** split are shown.

endpoint	metric	Calibrated RF	Neural FP RF
APR_HepG2_CellLoss_72h_dn	BA	0.71 ± 0.03	0.71 ± 0.03
APR_HepG2_CellLoss_72h_dn	Brier score	0.19 ± 0.01	0.19 ± 0.01
APR_HepG2_CellLoss_72h_dn	Log loss	0.57 ± 0.03	0.56 ± 0.03
ATG_NRF2_ARE_CIS_up	BA	0.71 ± 0.01	0.71 ± 0.01
ATG_NRF2_ARE_CIS_up	Brier score	0.18 ± 0.00	0.18 ± 0.00
ATG_NRF2_ARE_CIS_up	Log loss	0.53 ± 0.01	0.53 ± 0.01
ATG_PXRE_CIS_up	BA	0.74 ± 0.01	0.74 ± 0.01
ATG_PXRE_CIS_up	Brier score	0.18 ± 0.00	0.18 ± 0.00
ATG_PXRE_CIS_up	Log loss	0.55 ± 0.01	0.54 ± 0.01
BSK_3C_HLADR_down	BA	0.68 ± 0.03	0.67 ± 0.03
BSK_3C_HLADR_down	Brier score	0.19 ± 0.01	0.19 ± 0.01
BSK_3C_HLADR_down	Log loss	0.56 ± 0.02	0.56 ± 0.01
BSK_3C_Proliferation_down	BA	0.71 ± 0.04	0.72 ± 0.03
BSK_3C_Proliferation_down	Brier score	0.19 ± 0.01	0.19 ± 0.01
BSK_3C_Proliferation_down	Log loss	0.56 ± 0.03	0.56 ± 0.03
BSK_3C_SRB_down	BA	0.63 ± 0.05	0.66 ± 0.04
BSK_3C_SRB_down	Brier score	0.18 ± 0.01	0.18 ± 0.01
BSK_3C_SRB_down	Log loss	0.54 ± 0.03	0.53 ± 0.03
BSK_3C_Vis_down	BA	0.62 ± 0.03	0.63 ± 0.03
BSK_3C_Vis_down	Brier score	0.17 ± 0.01	0.17 ± 0.01
BSK_3C_Vis_down	Log loss	0.51 ± 0.02	0.51 ± 0.02
BSK_4H_Eotaxin3_down	BA	0.59 ± 0.01	0.62 ± 0.03
BSK_4H_Eotaxin3_down	Brier score	0.18 ± 0.00	0.18 ± 0.00
BSK_4H_Eotaxin3_down	Log loss	0.54 ± 0.01	0.53 ± 0.01
BSK_CASM3C_Proliferation_down	BA	0.59 ± 0.01	0.61 ± 0.01
BSK_CASM3C_Proliferation_down	Brier score	0.18 ± 0.00	0.18 ± 0.00
BSK_CASM3C_Proliferation_down	Log loss	0.53 ± 0.01	0.53 ± 0.01
BSK_LPS_VCAM1_down	BA	0.61 ± 0.01	0.63 ± 0.01
BSK_LPS_VCAM1_down	Brier score	0.17 ± 0.00	0.17 ± 0.00
BSK_LPS_VCAM1_down	Log loss	0.52 ± 0.01	0.52 ± 0.01
BSK_SAg_CD38_down	BA	0.64 ± 0.02	0.64 ± 0.02
BSK_SAg_CD38_down	Brier score	0.17 ± 0.01	0.17 ± 0.01
BSK_SAg_CD38_down	Log loss	0.51 ± 0.02	0.51 ± 0.02
BSK_SAg_CD40_down	BA	0.63 ± 0.04	0.63 ± 0.03
BSK_SAg_CD40_down	Brier score	0.17 ± 0.01	0.17 ± 0.01
BSK_SAg_CD40_down	Log loss	0.52 ± 0.03	0.51 ± 0.03
BSK_SAg_Proliferation_down	BA	0.65 ± 0.02	0.67 ± 0.04
BSK_SAg_Proliferation_down	Brier score	0.20 ± 0.02	0.19 ± 0.02
BSK_SAg_Proliferation_down	Log loss	0.57 ± 0.04	0.57 ± 0.04
BSK_hDFCGF_CollagenIII_down	BA	0.59 ± 0.03	0.63 ± 0.03
BSK_hDFCGF_CollagenIII_down	Brier score	0.17 ± 0.01	0.17 ± 0.01
BSK_hDFCGF_CollagenIII_down	Log loss	0.51 ± 0.02	0.50 ± 0.02
BSK_hDFCGF_Proliferation_down	BA	0.68 ± 0.02	0.69 ± 0.02
BSK_hDFCGF_Proliferation_down	Brier score	0.20 ± 0.01	0.20 ± 0.00
BSK_hDFCGF_Proliferation_down	Log loss	0.58 ± 0.01	0.57 ± 0.01
CEETOX_H295R_11DCORT_dn	BA	0.57 ± 0.02	0.64 ± 0.05
CEETOX_H295R_11DCORT_dn	Brier score	0.22 ± 0.00	0.21 ± 0.01
CEETOX_H295R_11DCORT_dn	Log loss	0.62 ± 0.01	0.60 ± 0.03
CEETOX_H295R_ANDR_dn	BA	0.55 ± 0.03	0.62 ± 0.07
CEETOX_H295R_ANDR_dn	Brier score	0.22 ± 0.01	0.21 ± 0.03
CEETOX_H295R_ANDR_dn	Log loss	0.62 ± 0.03	0.61 ± 0.06
TOX21_ARE_BLA_agonist_ratio	BA	0.52 ± 0.01	0.60 ± 0.01
TOX21_ARE_BLA_agonist_ratio	Brier score	0.15 ± 0.00	0.14 ± 0.00
TOX21_ARE_BLA_agonist_ratio	Log loss	0.46 ± 0.01	0.44 ± 0.01
TOX21_TR_LUC_GH3_Antagonist	BA	0.69 ± 0.01	0.69 ± 0.01
TOX21_TR_LUC_GH3_Antagonist	Brier score	0.11 ± 0.00	0.11 ± 0.00
TOX21_TR_LUC_GH3_Antagonist	Log loss	0.36 ± 0.01	0.36 ± 0.00

Table S4. Performance of models on test sets generated with agglomerative clustering and random splitting. Precision and recall values are reported as mean ± standard deviation. Morgan FP is the Morgan count fingerprint.

model	Agglomerative clustering		Random	
	Precision	Recall	Precision	Recall
Morgan FP + KNN	0.56 ± 0.12	0.24 ± 0.12	0.61 ± 0.09	0.30 ± 0.11
Neural FP + KNN	0.52 ± 0.10	0.38 ± 0.15	0.56 ± 0.08	0.41 ± 0.14
Morgan FP + RF	0.58 ± 0.15	0.24 ± 0.19	0.66 ± 0.08	0.37 ± 0.15
Neural FP + RF	0.57 ± 0.10	0.41 ± 0.18	0.63 ± 0.07	0.47 ± 0.15
Morgan FP + SVC	0.56 ± 0.10	0.49 ± 0.13	0.61 ± 0.08	0.52 ± 0.10
Neural FP + SVC	0.59 ± 0.09	0.44 ± 0.15	0.64 ± 0.08	0.46 ± 0.14
Chemprop	0.57 ± 0.09	0.51 ± 0.13	0.63 ± 0.07	0.52 ± 0.10
Cal. Chemprop	0.53 ± 0.26	0.32 ± 0.23	0.58 ± 0.29	0.33 ± 0.24

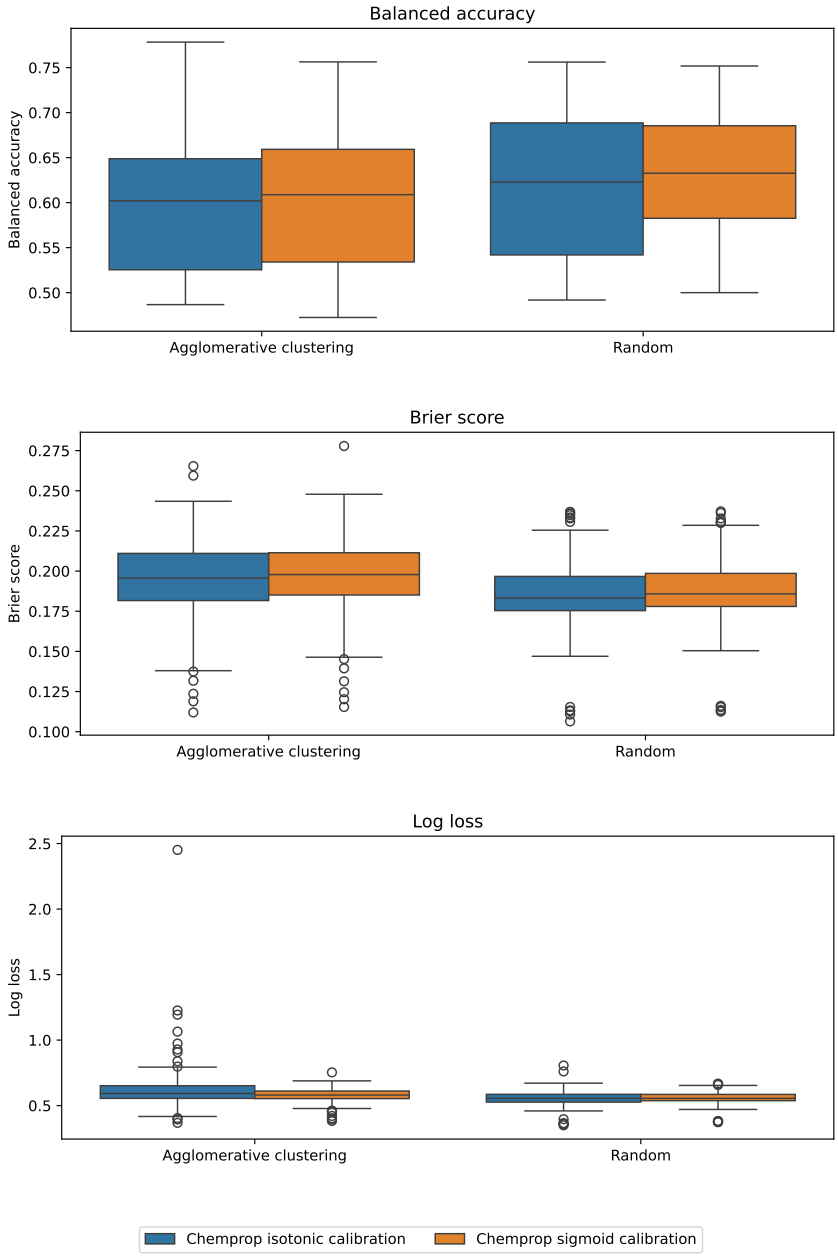

Fig. S1. Performance comparison of predictions and uncertainty estimates of Chemprop calibrated with isotonic regression and sigmoid calibration. The boxplots show the prediction performance (balanced accuracy, higher values are better) and uncertainty estimates (Brier score and Log loss, lower values are better) of all folds in the five-fold cross validation on the ToxCast subset (19 endpoints).

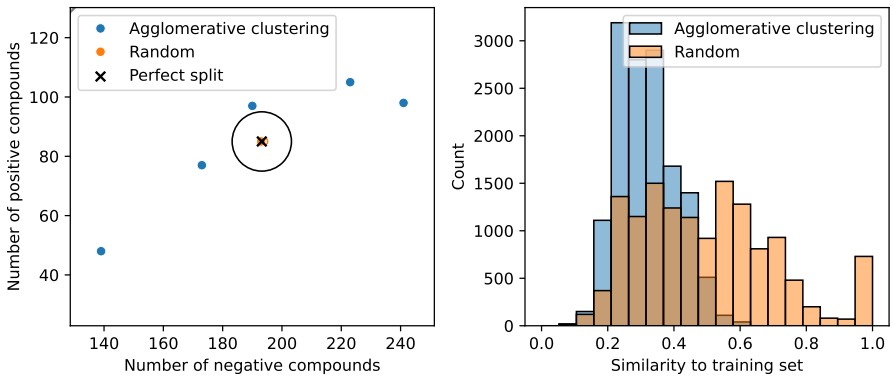

Fig. S2. Train/Test data distributions for BSK_SAg_CD40_down endpoint. **Left:** The number of positive and negative labeled compounds in the folds obtained from clustering-based splitting (blue) and random splitting (orange). Dividing the total number of compounds of each class by five yields the optimal composition of a fold (black x), which is closely matched by the random split. The circle indicates a hypothetical fold with a radius of 10 compounds. Due to the clustering's constraints, only one fold ended up close to this area, while the remaining four folds differed in size and composition. **Right:** The similarity to the nearest training set neighbor for each test set compound across all five-folds. Similarity values of less than 0.3 to 0.4 indicate only a limited relation between test and training sets. On the other hand, folds obtained from the random split were significantly more similar, where some test compounds had similarities to the training compounds close to 1.

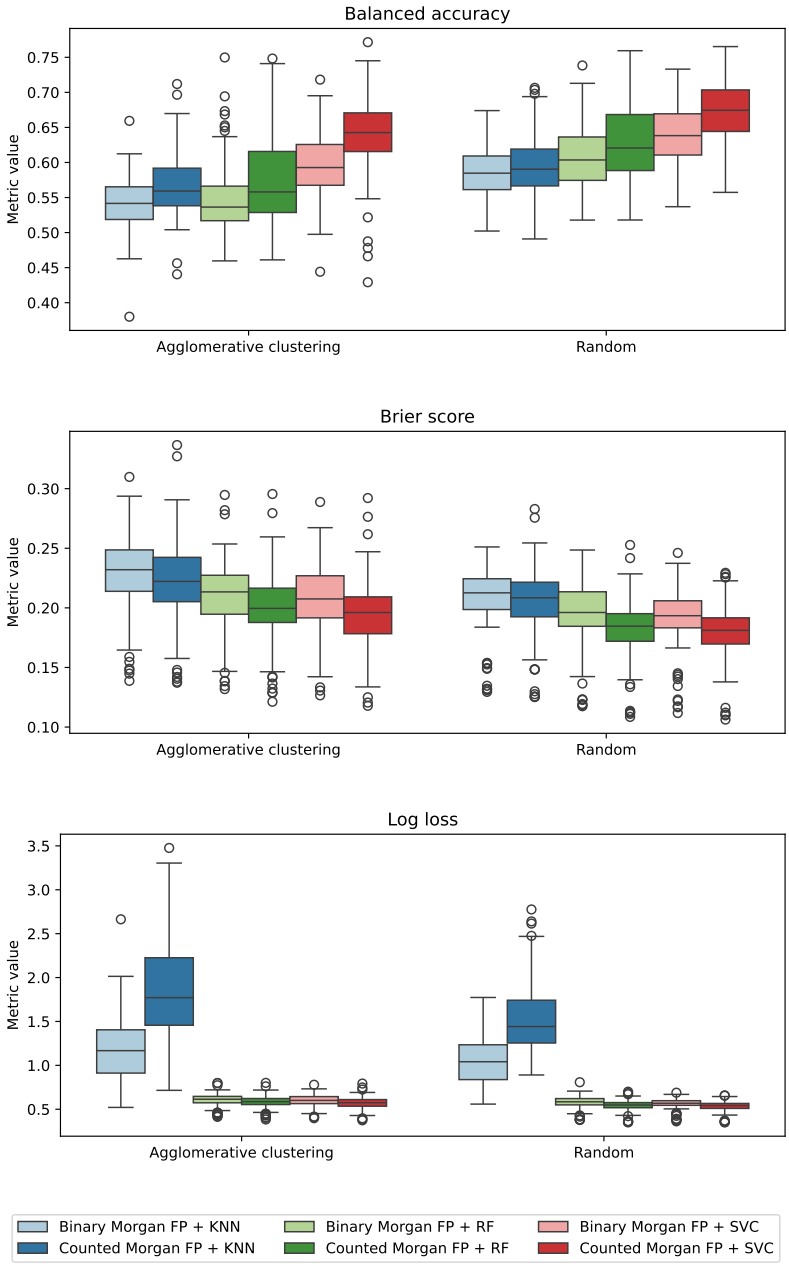

Fig. S3. Performance of Morgan binary fingerprints and Morgan count fingerprints. Performance of predictions (balanced accuracy, higher values are better) and uncertainty estimates (Brier score and Log loss, lower values are better) of all folds in the five-fold cross validation on the ToxCast subset (19 endpoints).

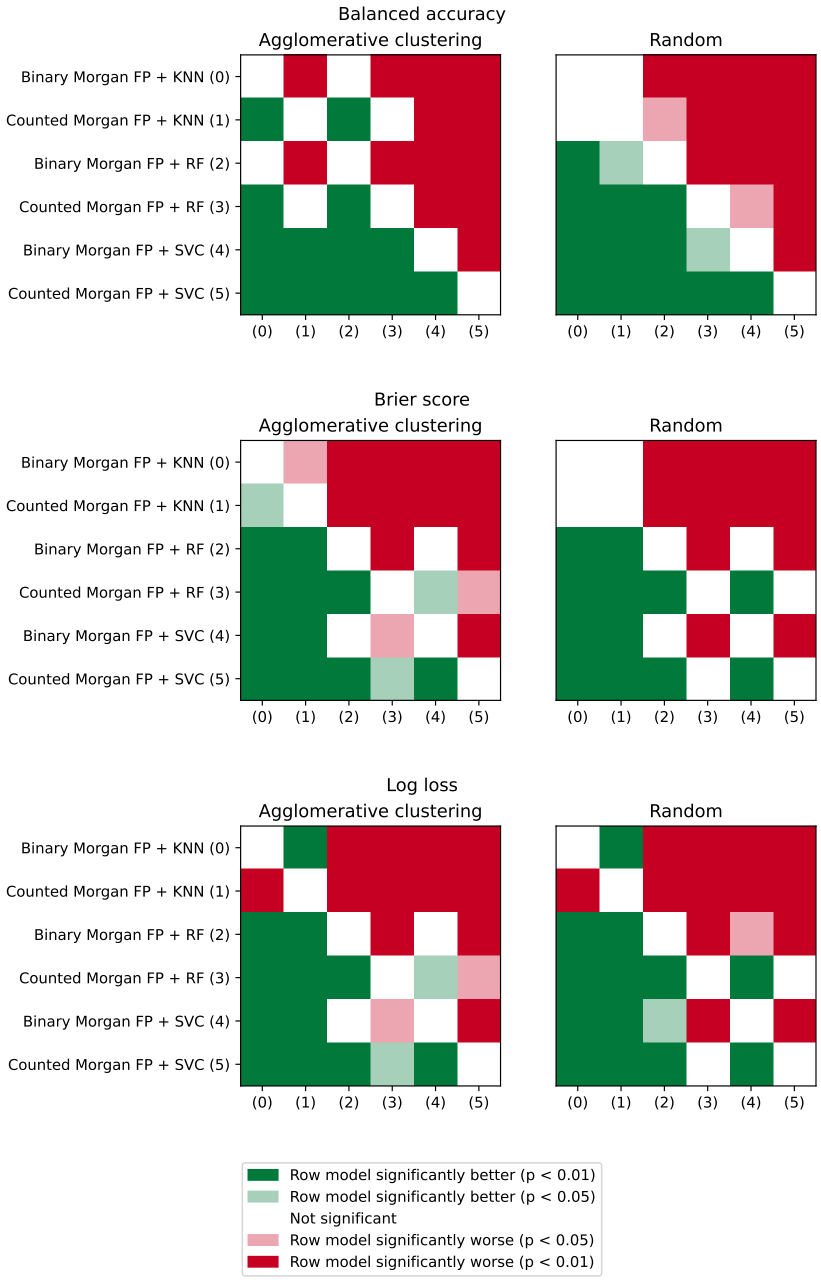

Fig. S4. Significance test comparison of Morgan binary fingerprints and Morgan count fingerprints in the five-fold cross validation on the ToxCast subset (19 endpoints) using one-sided Mann Whitney U test.

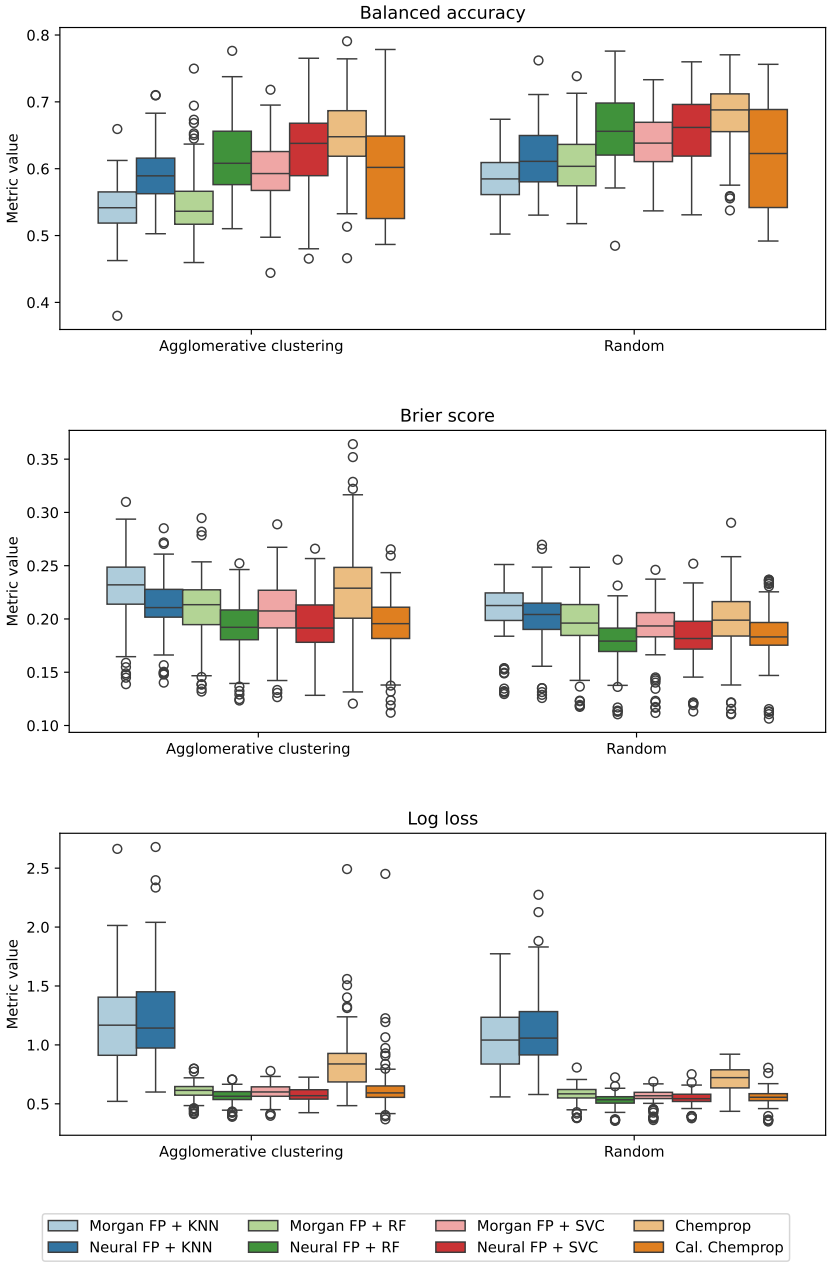

Fig. S5. Performance of Morgan binary fingerprints and neural fingerprints from Chemprop. Performance of predictions (balanced accuracy, higher values are better) and uncertainty estimates (Brier score and Log loss, lower values are better) of all folds in the five-fold cross validation on the ToxCast subset (19 endpoints).

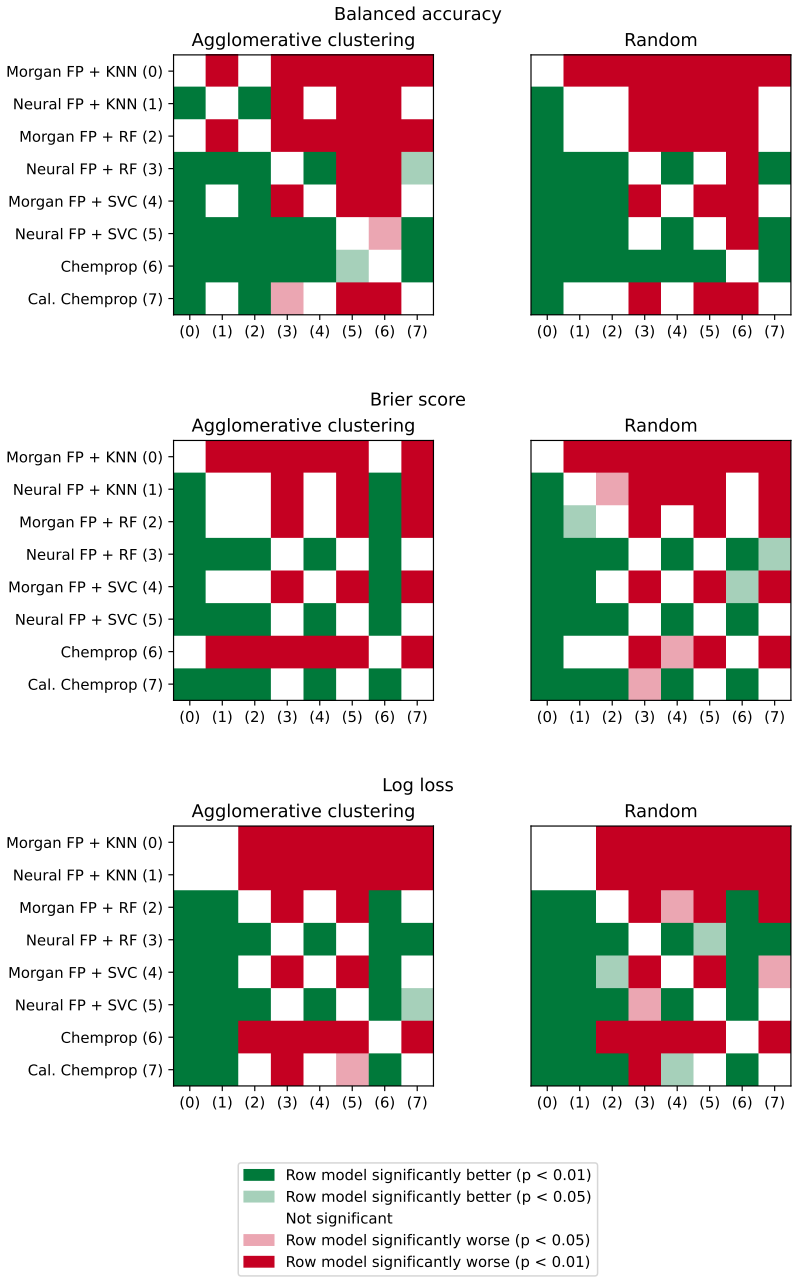

Fig. S6. Significance test comparison of Morgan binary fingerprints and neural fingerprints from Chemprop in the five-fold cross validation on the ToxCast subset (19 endpoints) using one-sided Mann Whitney U test.

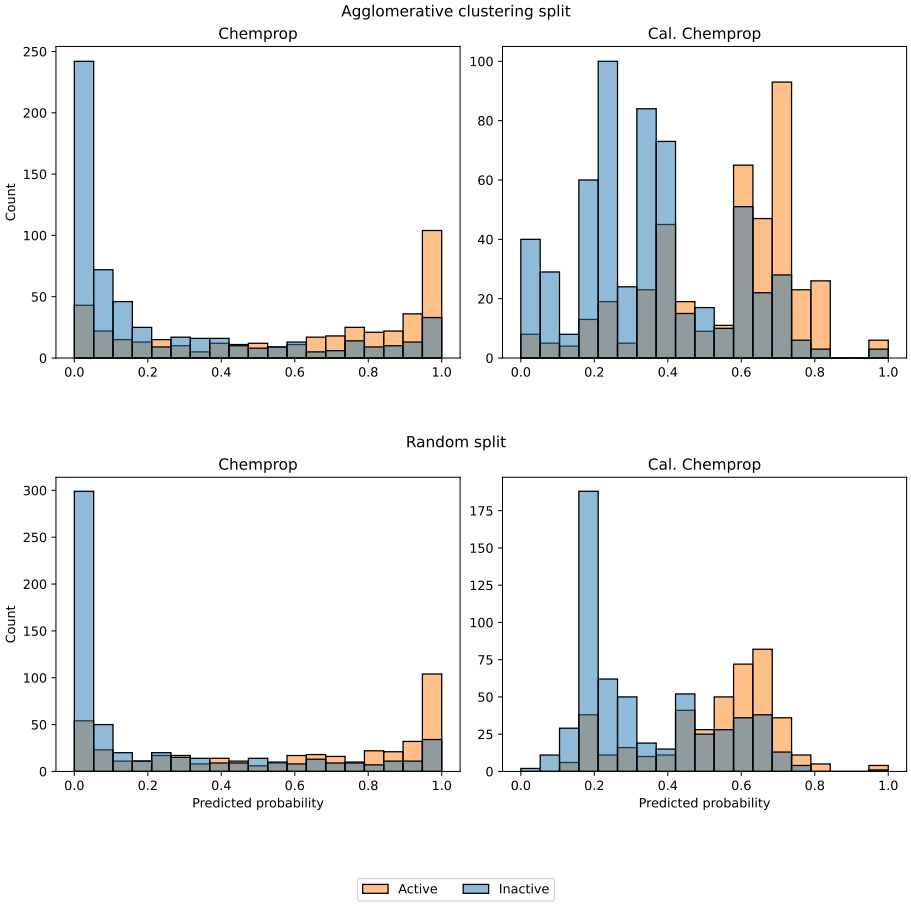

Fig. S7. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the APR_HepG2_CellLoss_72h_dn ToxCast endpoint.

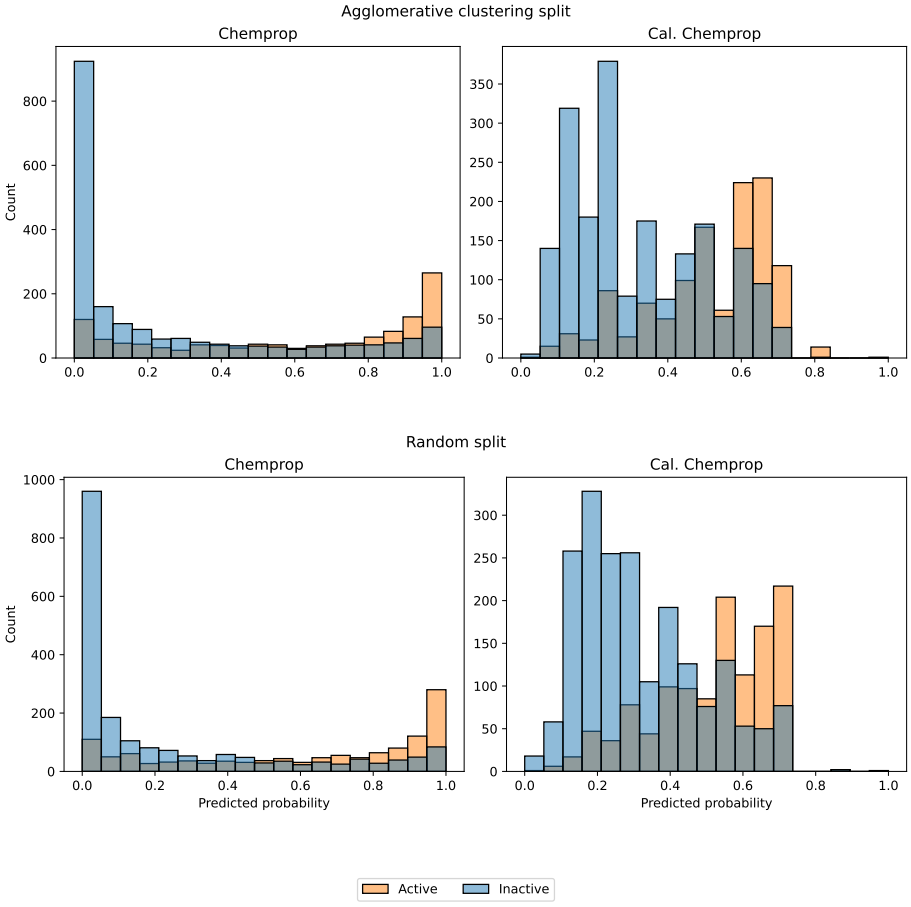

Fig. S8. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the ATG_NRF2_ARE_CIS_up ToxCast endpoint.

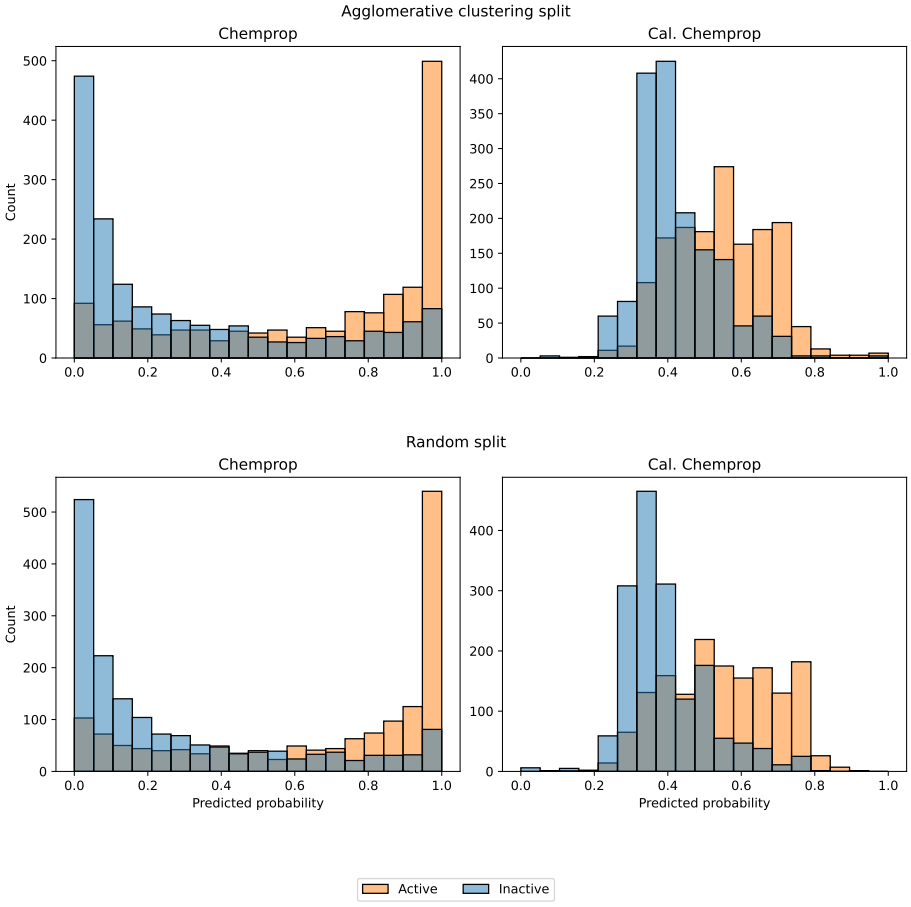

Fig. S9. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the ATG_PXRE_CIS_up ToxCast endpoint.

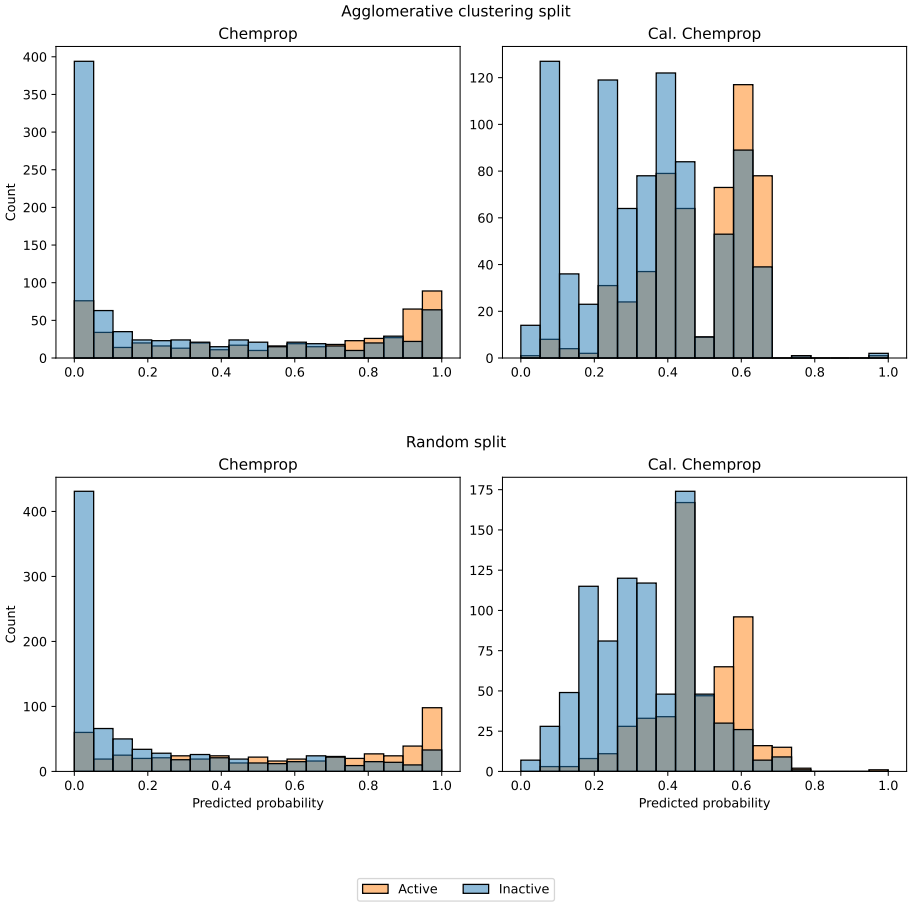

Fig. S10. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_3C_HLADR_down ToxCast endpoint.

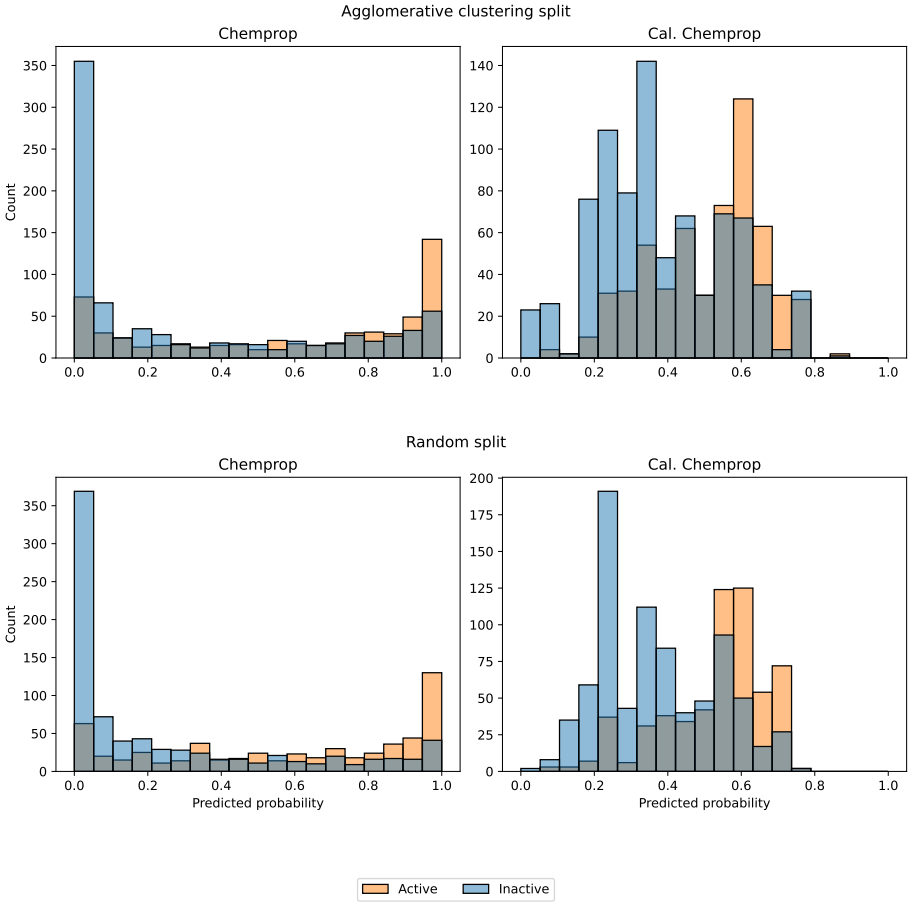

Fig. S11. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_3C_Proliferation_down ToxCast endpoint.

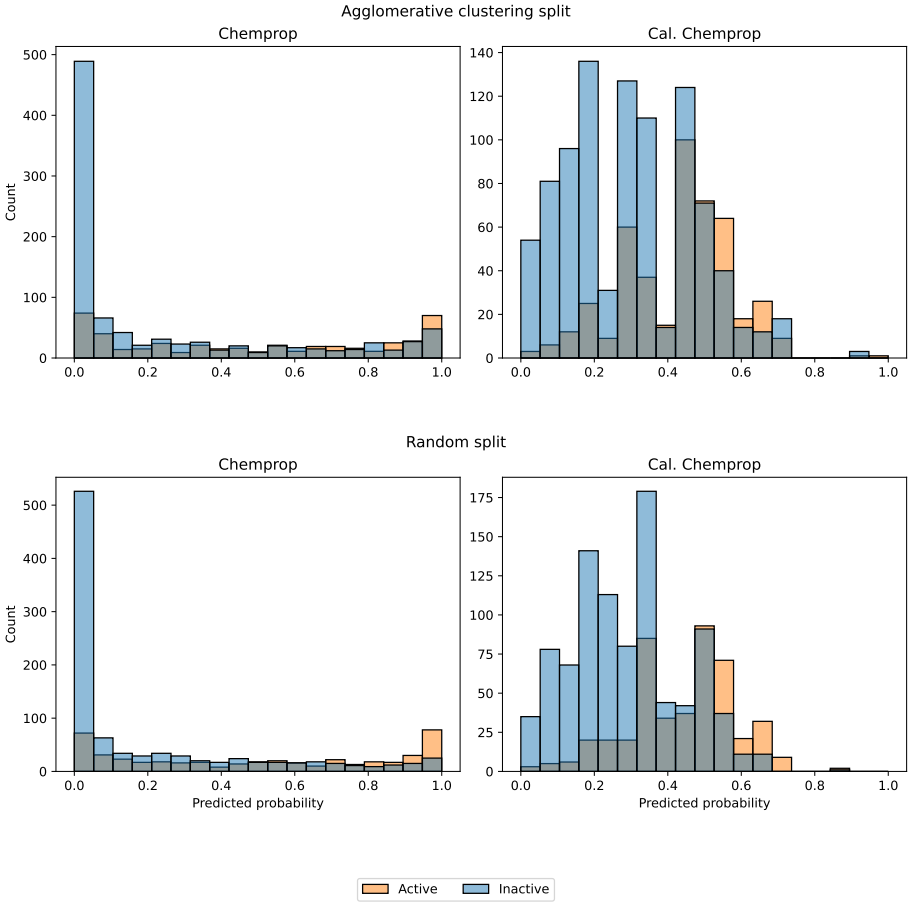

Fig. S12. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_3C_SRB_down ToxCast endpoint.

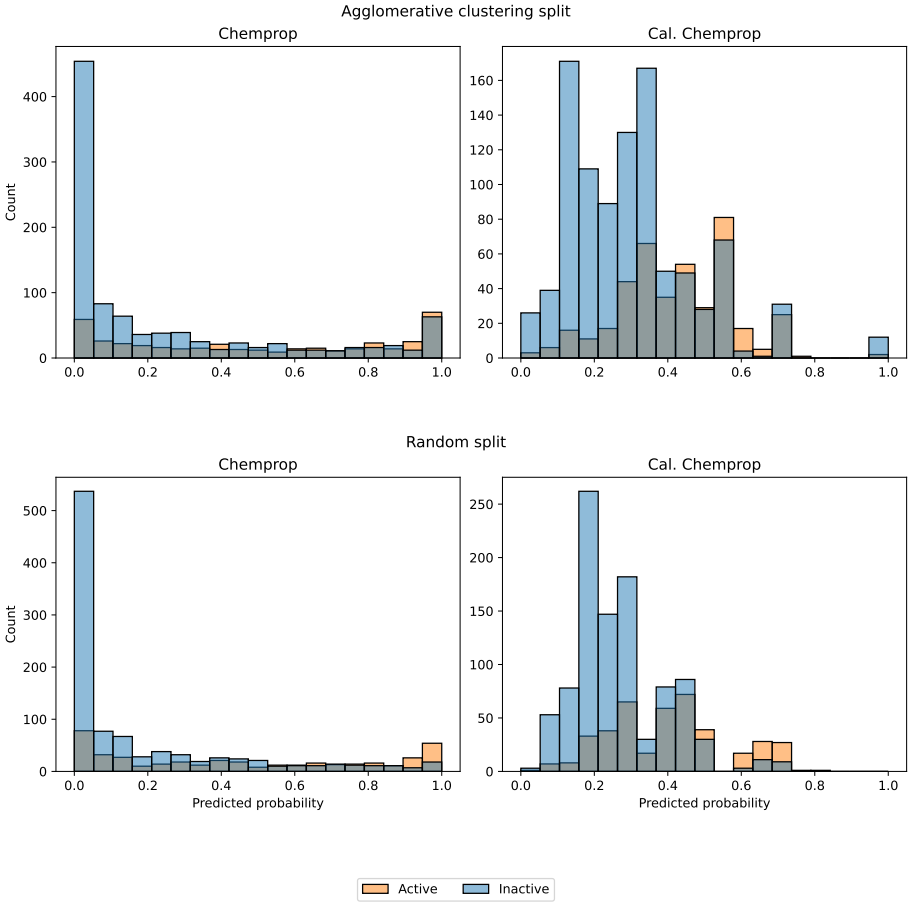

Fig. S13. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_3C_Vis_down ToxCast endpoint.

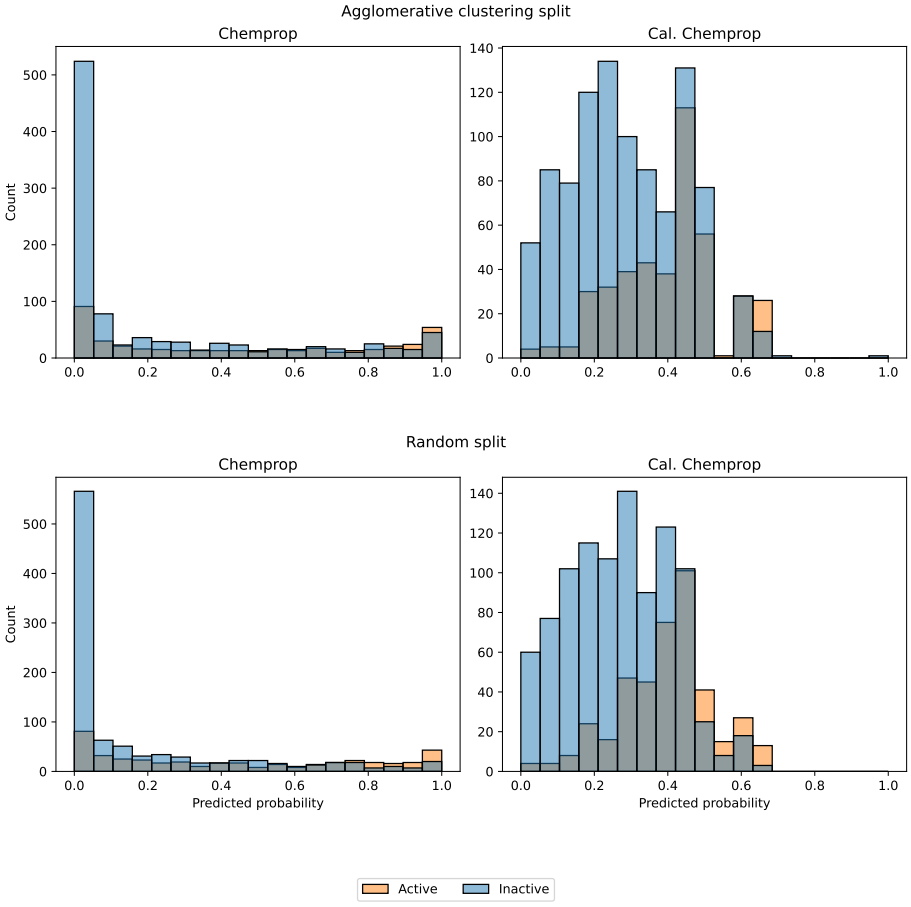

Fig. S14. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_4H_Eotaxin3_down ToxCast endpoint.

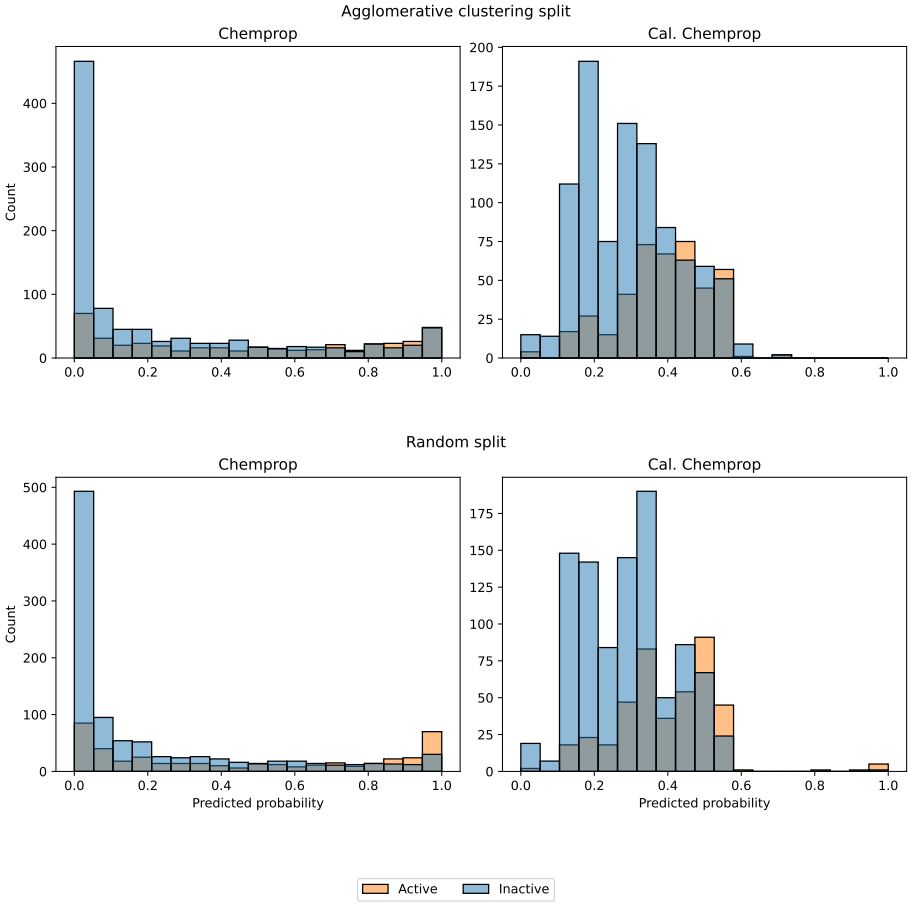

Fig. S15. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_CASM3C_Proliferation_down ToxCast endpoint.

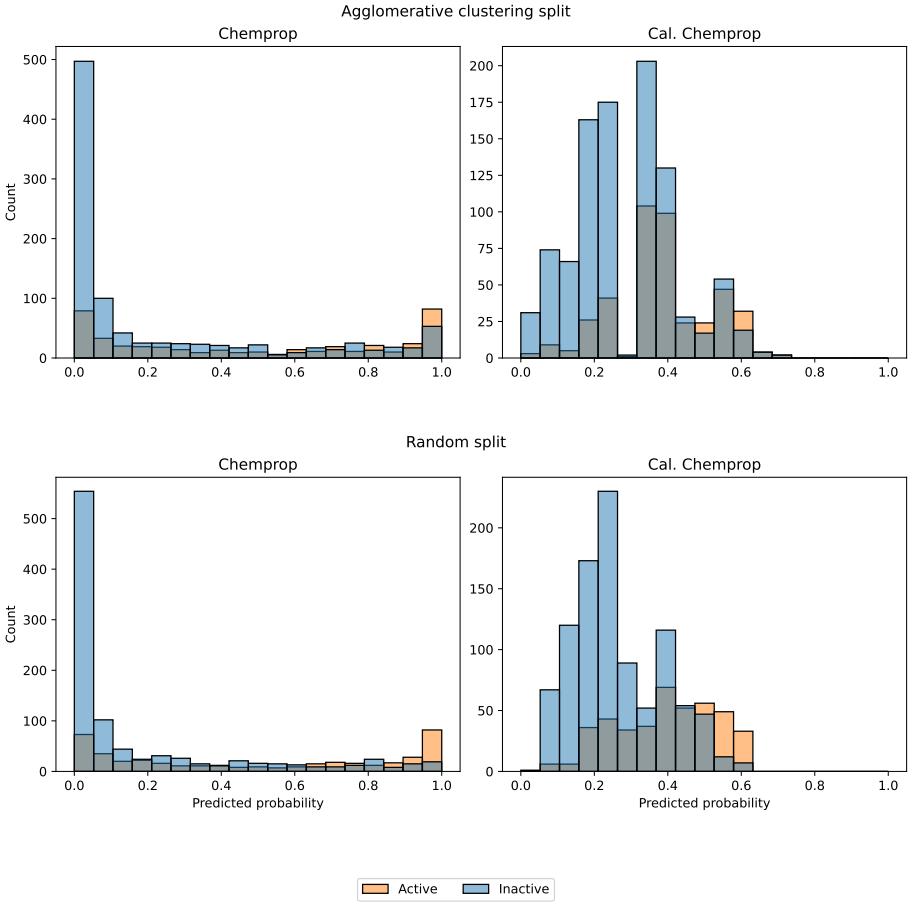

Fig. S16. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_LPS_VCAM1_down ToxCast endpoint.

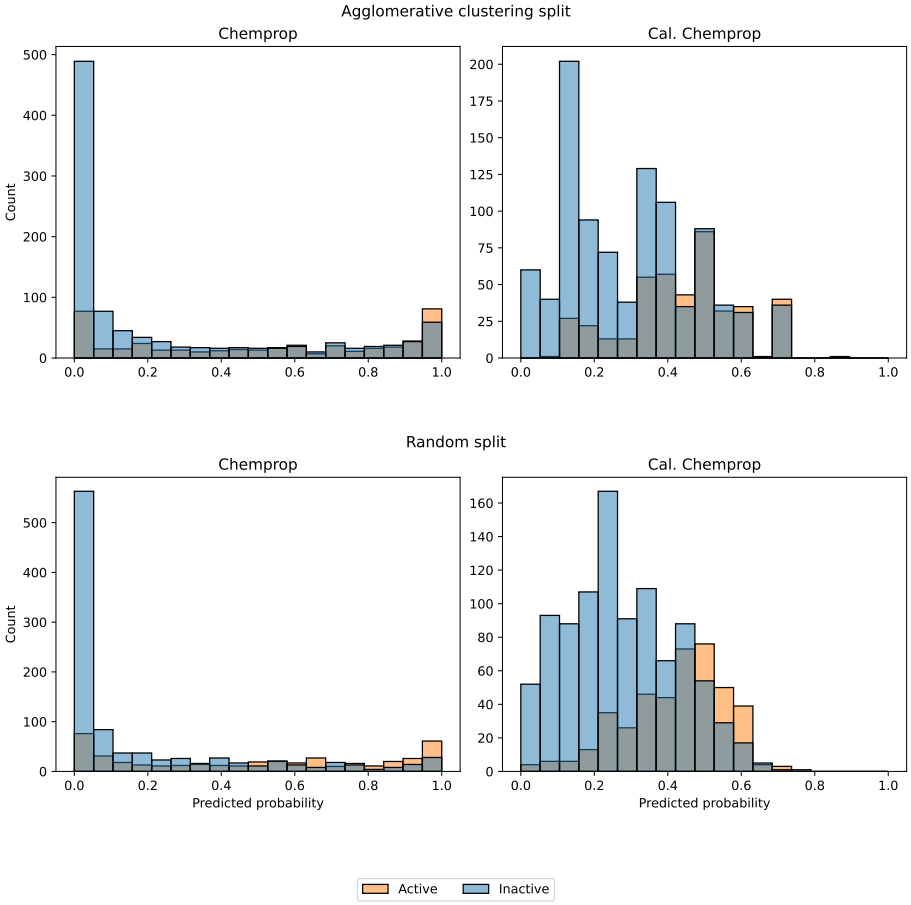

Fig. S17. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_SAg_CD38_down ToxCast endpoint.

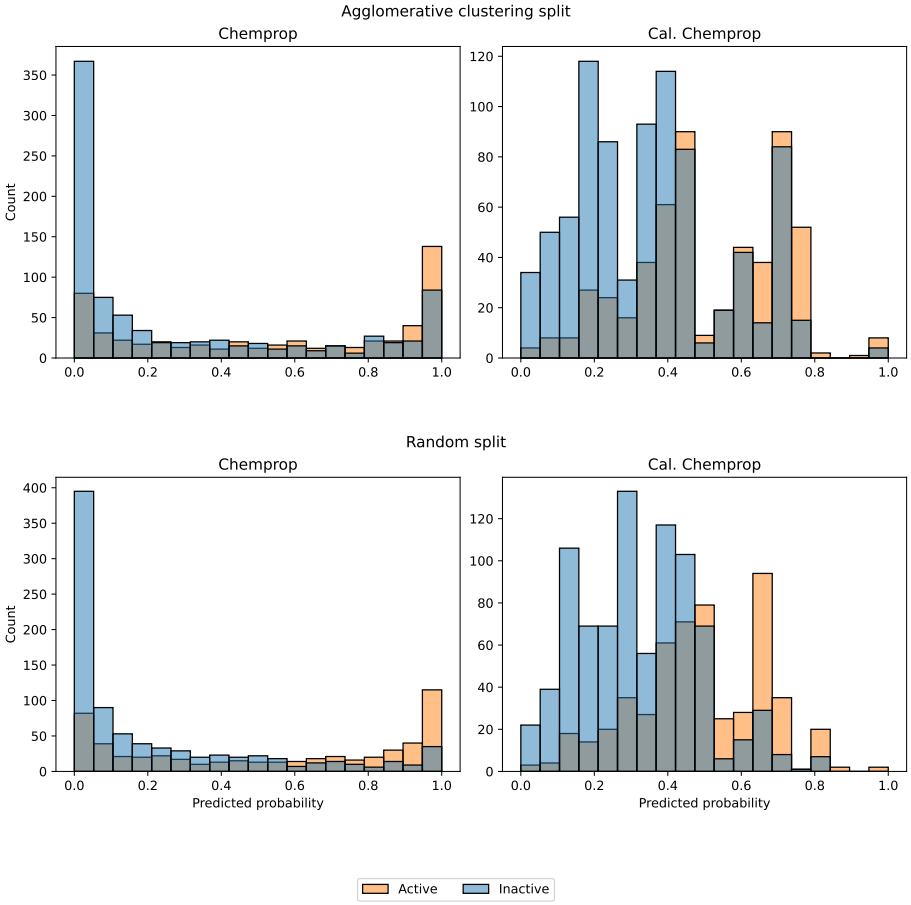

Fig. S18. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_SAg_Proliferation_down ToxCast endpoint.

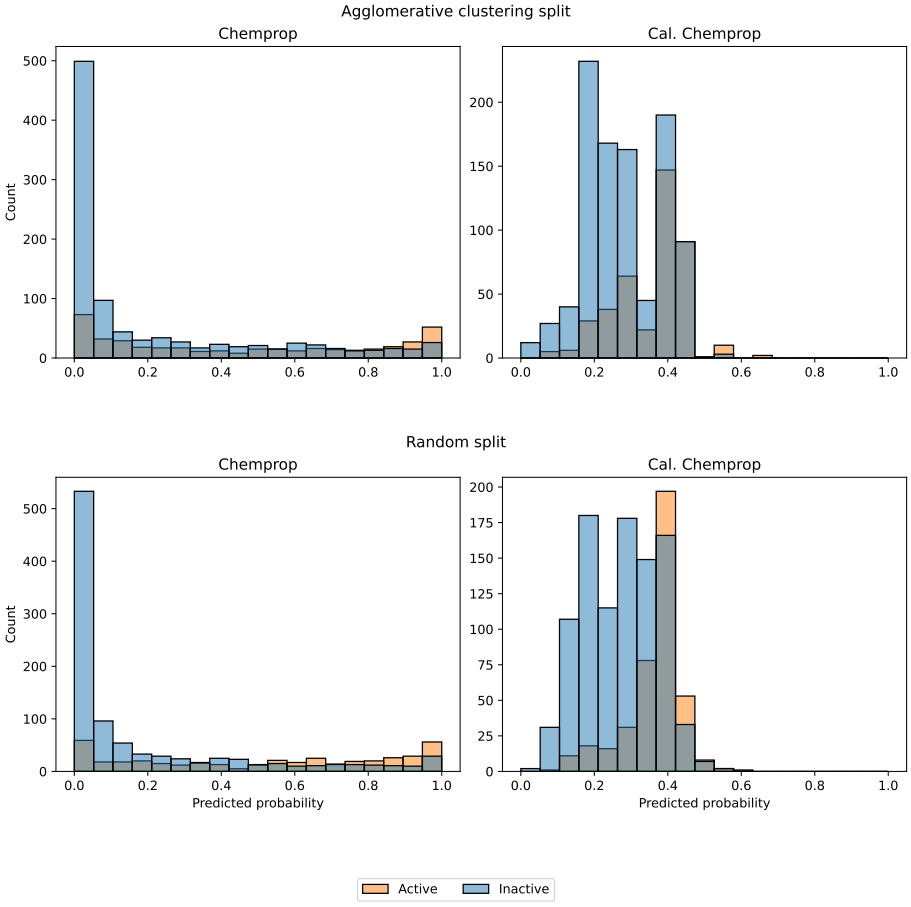

Fig. S19. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_hDFCGF_CollagenIII_down ToxCast endpoint.

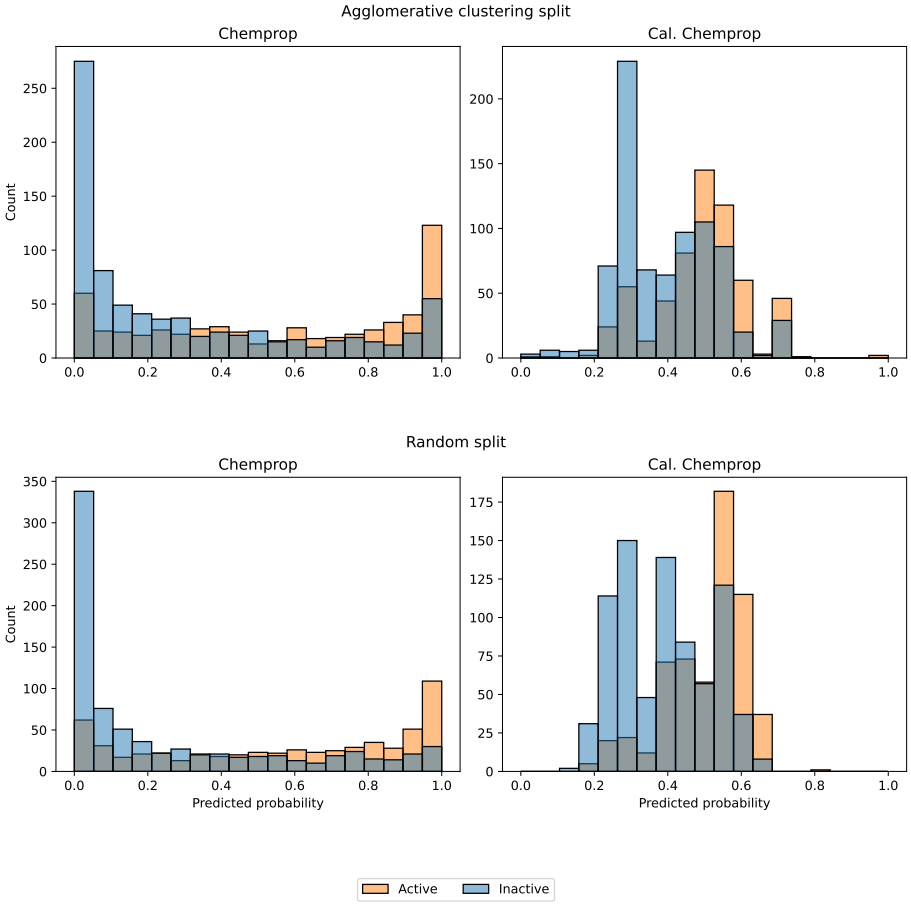

Fig. S20. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the BSK_hDFCGF_Proliferation_down ToxCast endpoint.

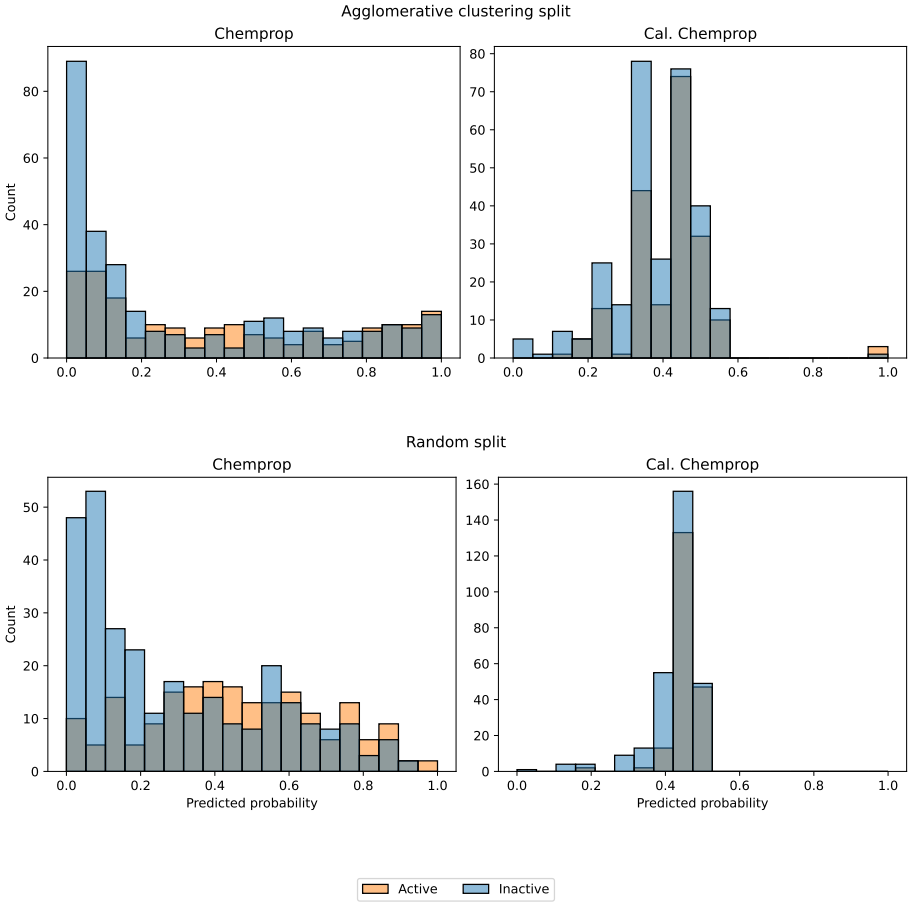

Fig. S21. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the CEETOX_H295R_11DCORT_dn ToxCast endpoint.

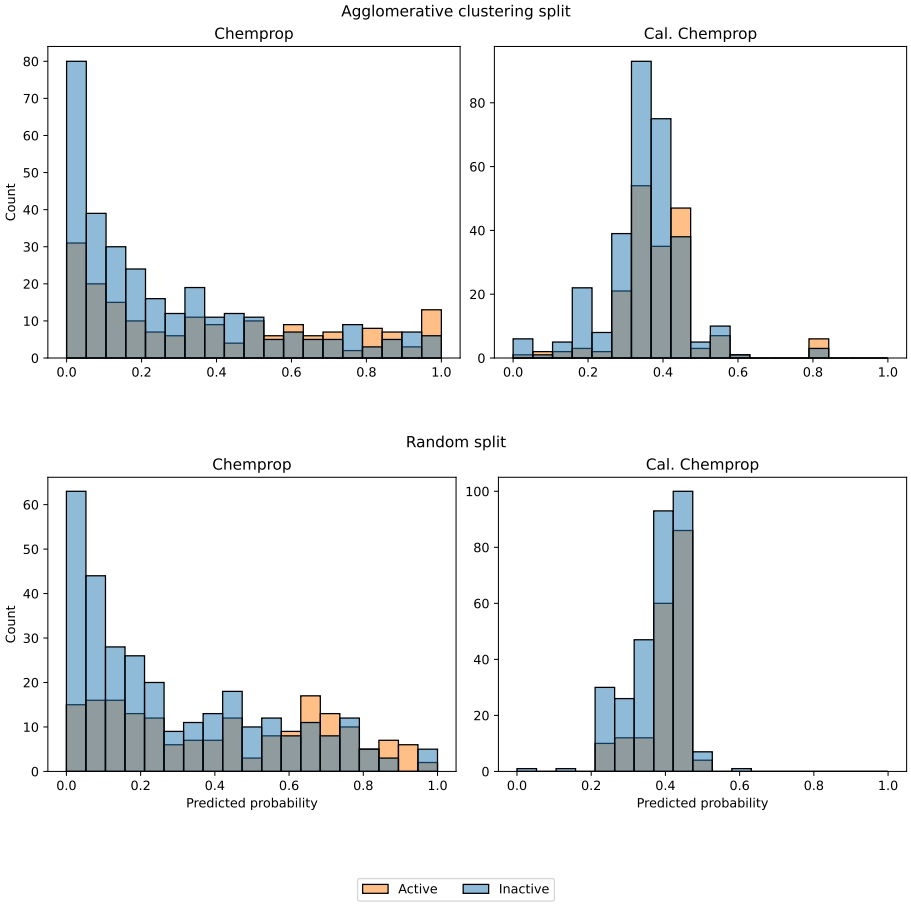

Fig. S22. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the CEETOX_H295R_ANDR_dn ToxCast endpoint.

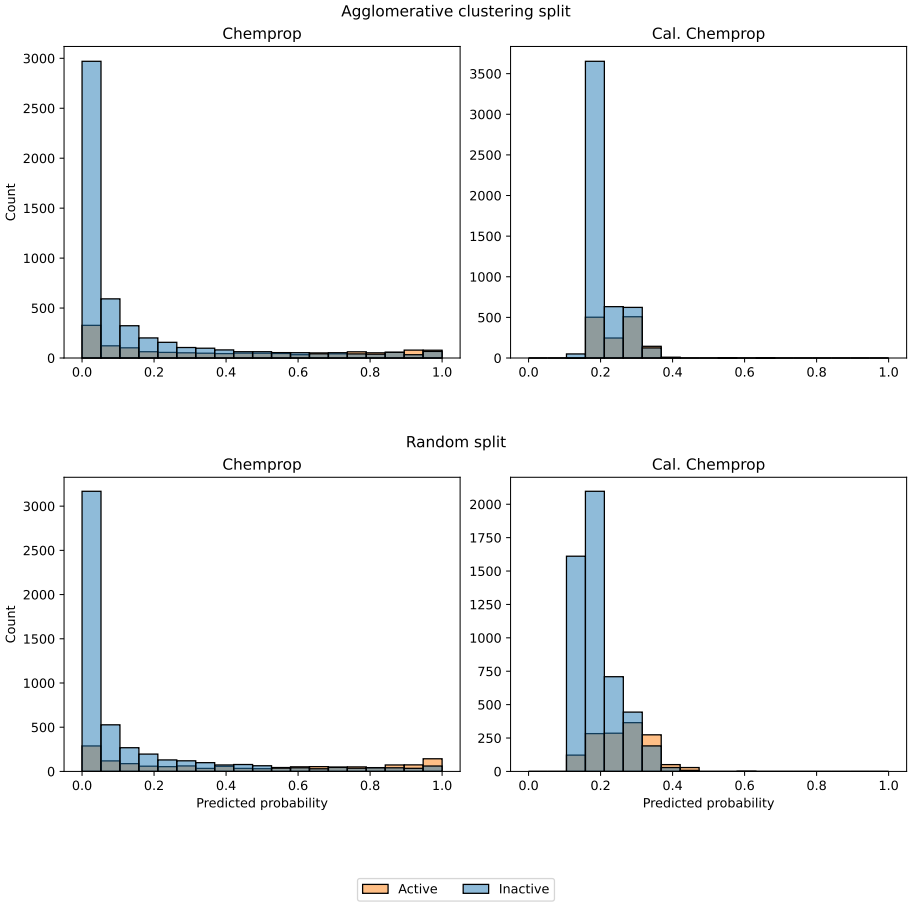

Fig. S23. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the TOX21_ARE_BLA_agonist_ratio ToxCast endpoint.

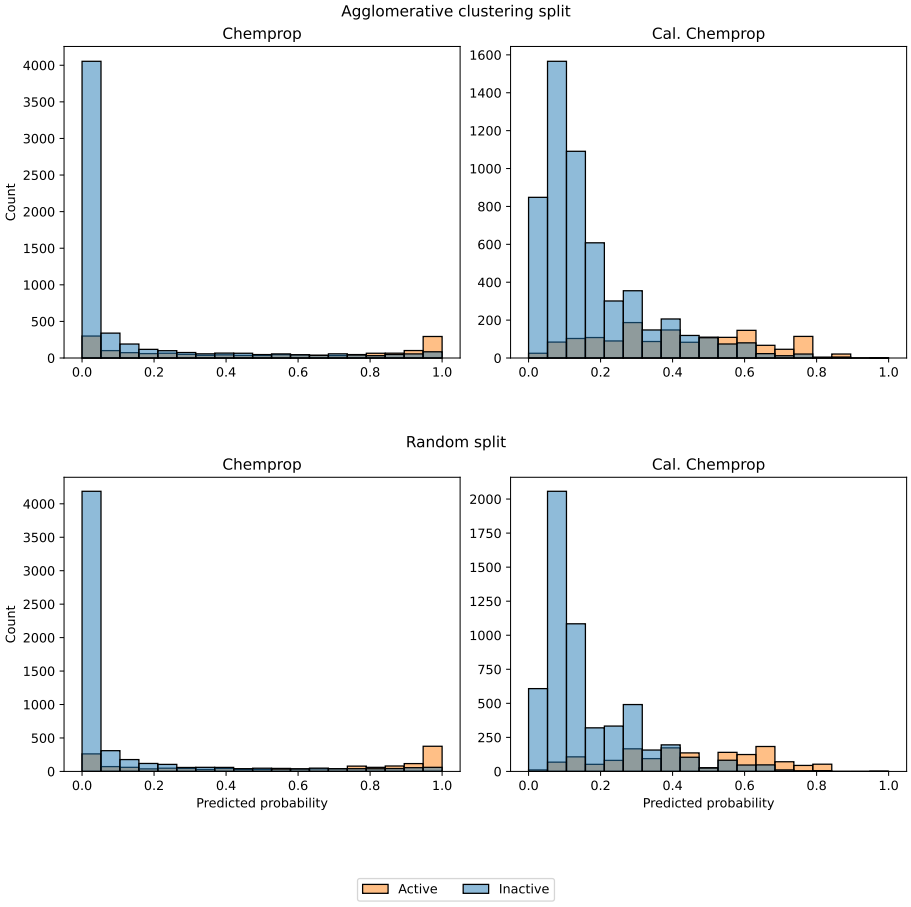

Fig. S24. Distributions of predicted probabilities of Chemprop and calibrated Chemprop on the TOX21_TR_LUC_GH3_Antagonist ToxCast endpoint.

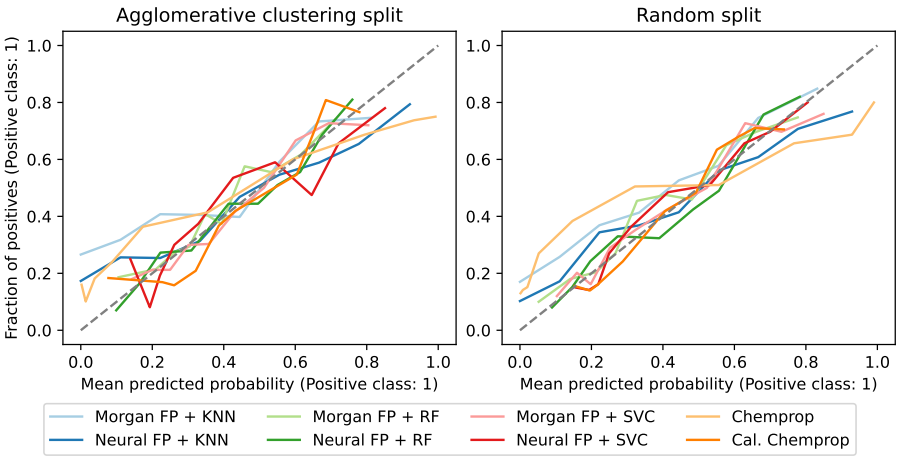

Fig. S25. Calibration curves for endpoint APR_HepG2_CellLoss_72h_dn.

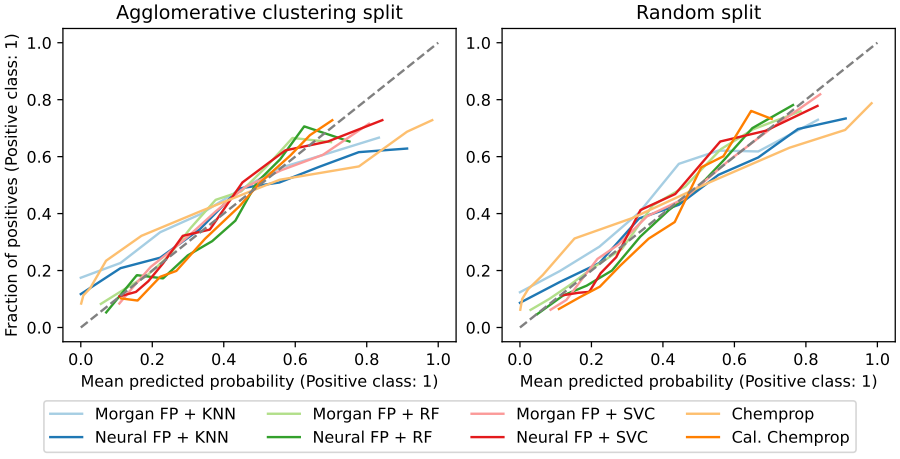

Fig. S26. Calibration curves for endpoint ATG_NRF2_ARE_CIS_up.

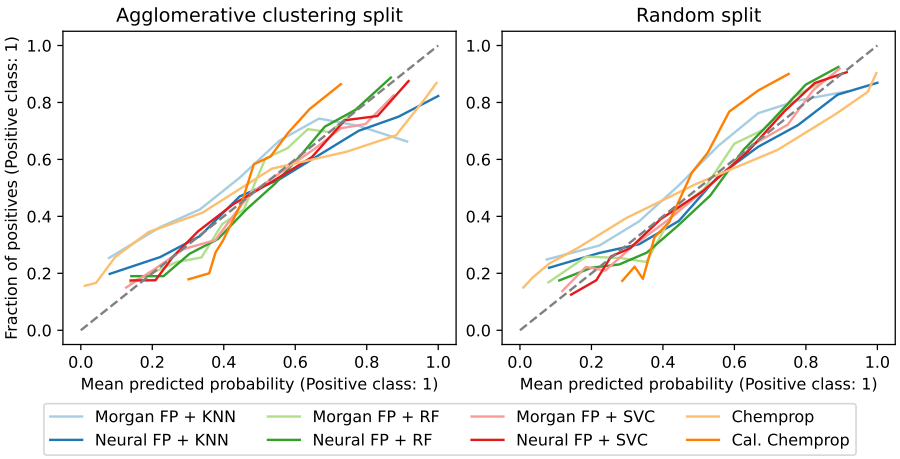

Fig. S27. Calibration curves for endpoint ATG_PXRE_CIS_up.

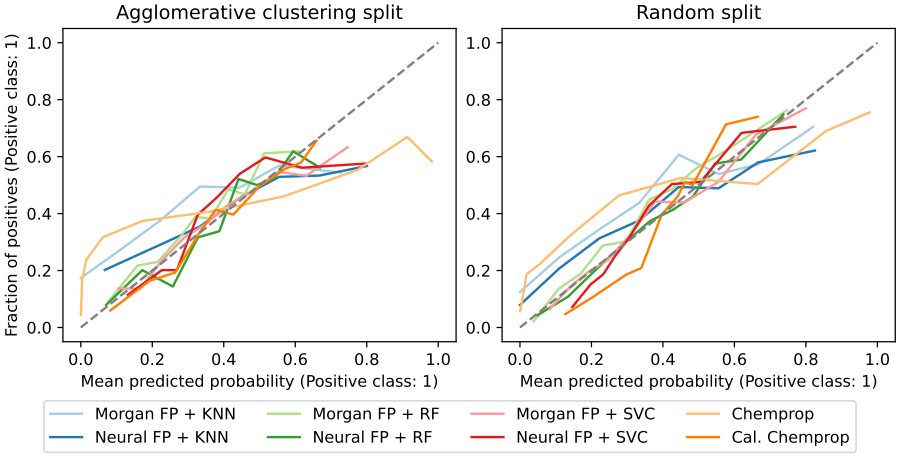

Fig. S28. Calibration curves for endpoint BSK_3C_HLADR_down.

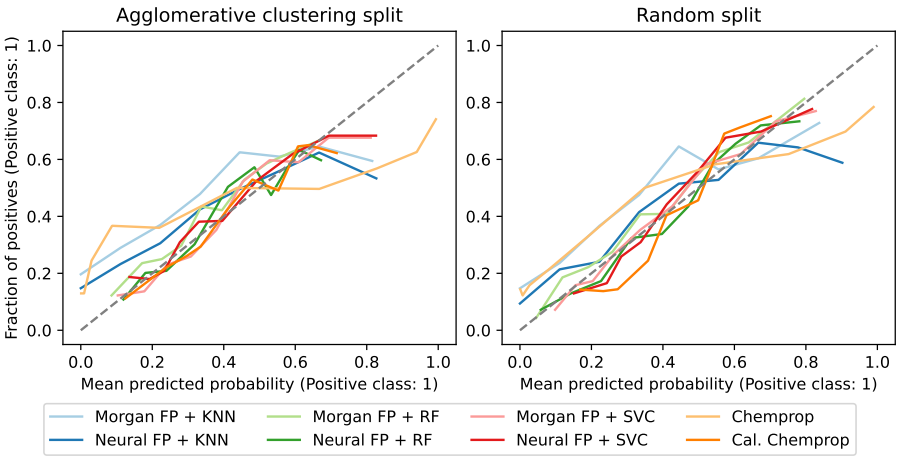

Fig. S29. Calibration curves for endpoint BSK_3C_Proliferation_down.

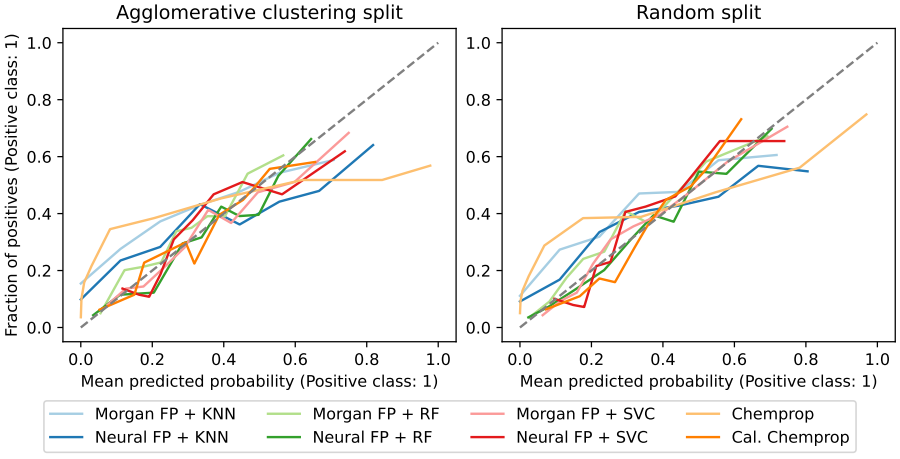

Fig. S30. Calibration curves for endpoint BSK_3C_SRB_down.

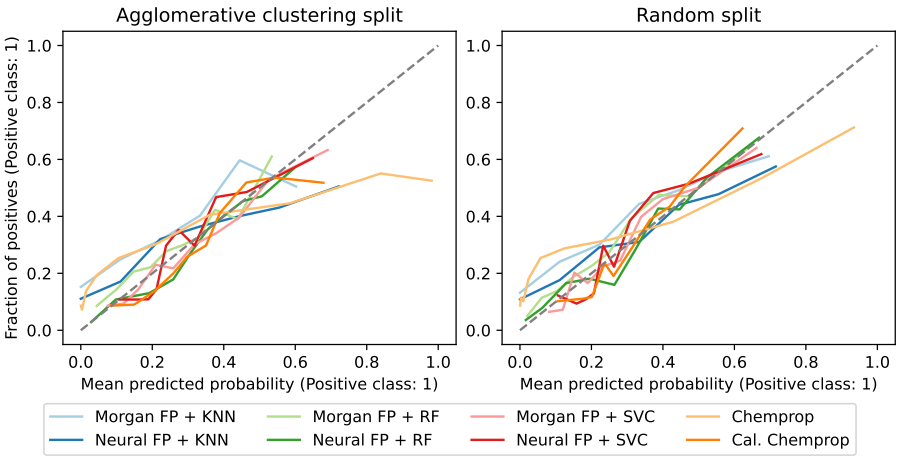

Fig. S31. Calibration curves for endpoint BSK_3C_Vis_down.

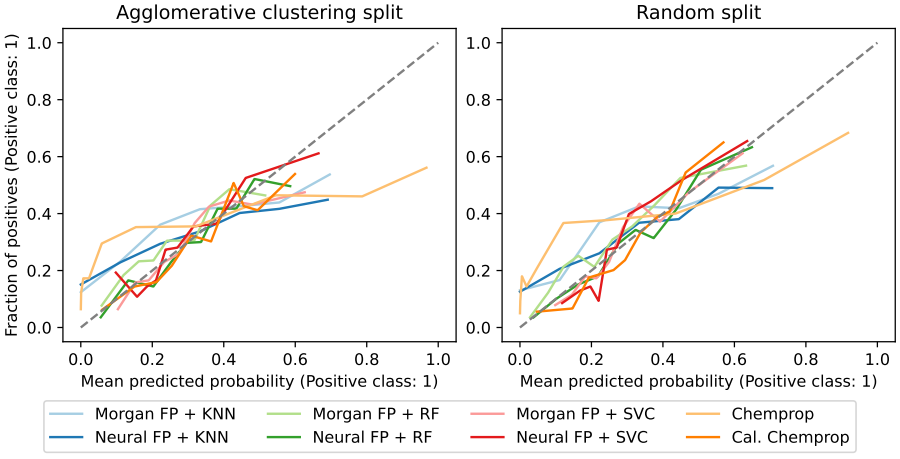

Fig. S32. Calibration curves for endpoint BSK_4H_Eotaxin3_down.

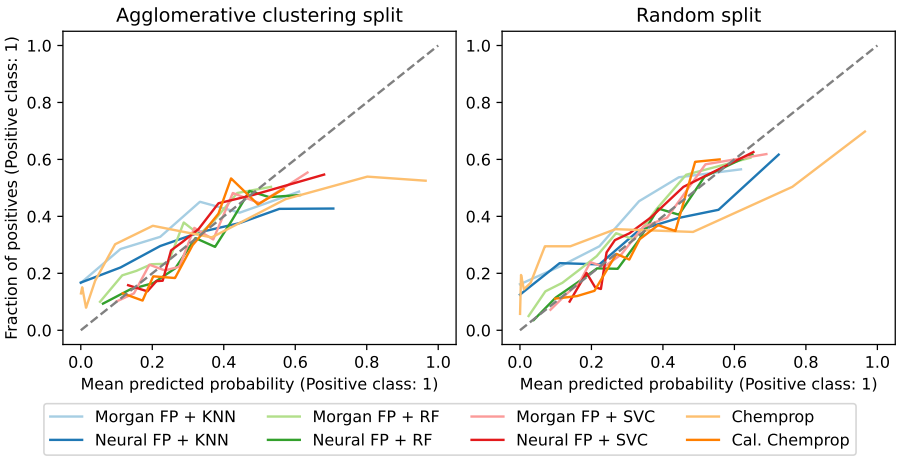

Fig. S33. Calibration curves for endpoint BSK_CASM3C_Proliferation_down.

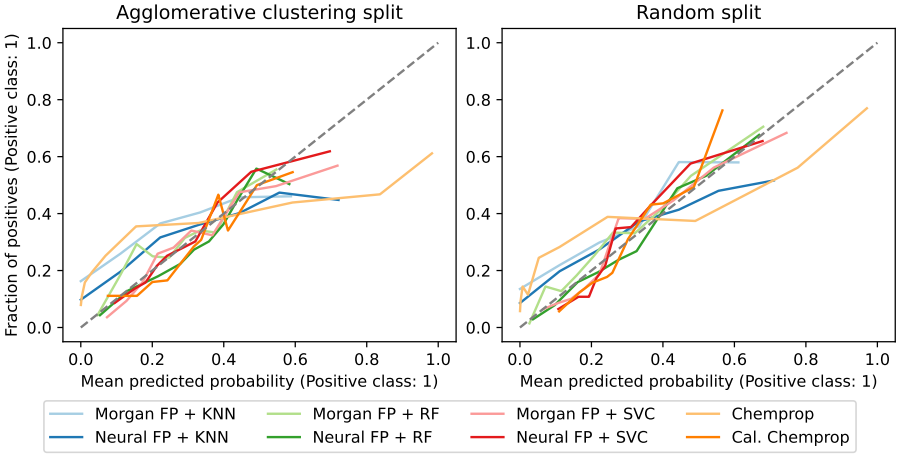

Fig. S34. Calibration curves for endpoint BSK_LPS_VCAM1_down.

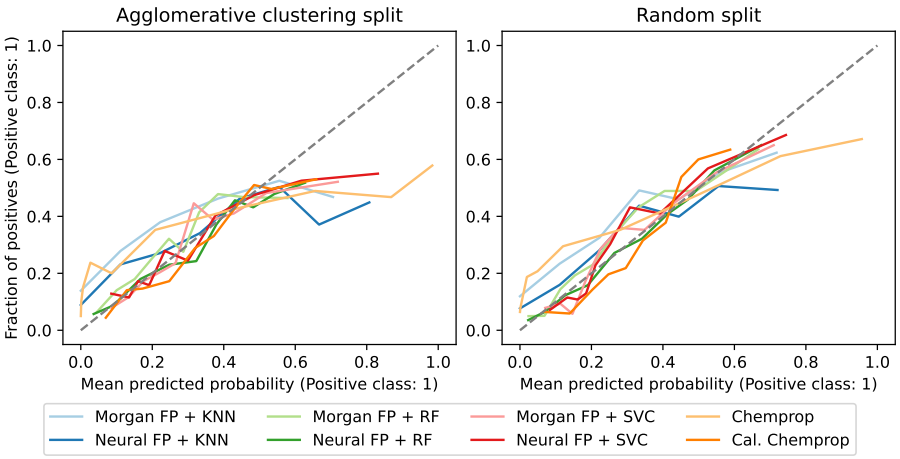

Fig. S35. Calibration curves for endpoint BSK_SAg_CD38_down.

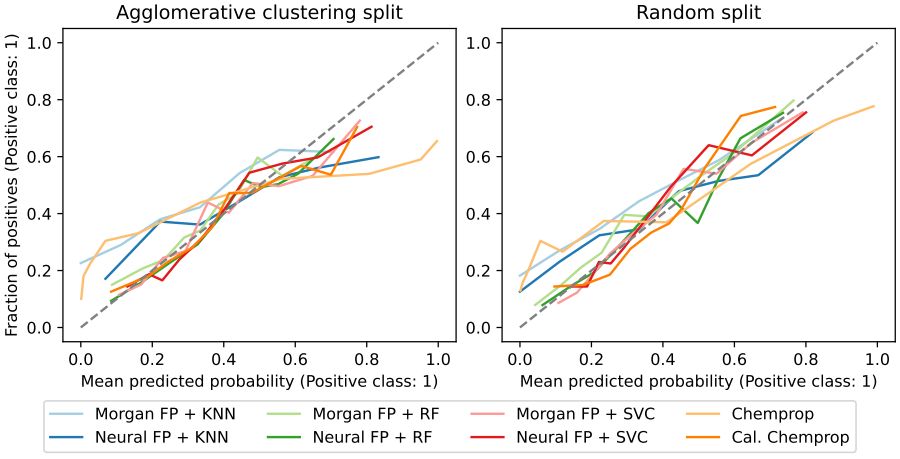

Fig. S36. Calibration curves for endpoint BSK_SAg_Proliferation_down.

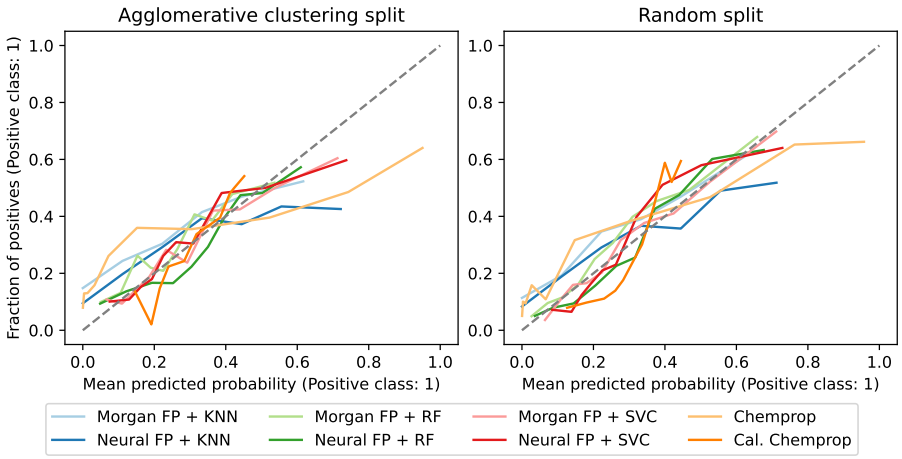

Fig. S37. Calibration curves for endpoint BSK_hDFCGF_CollagenIII_down.

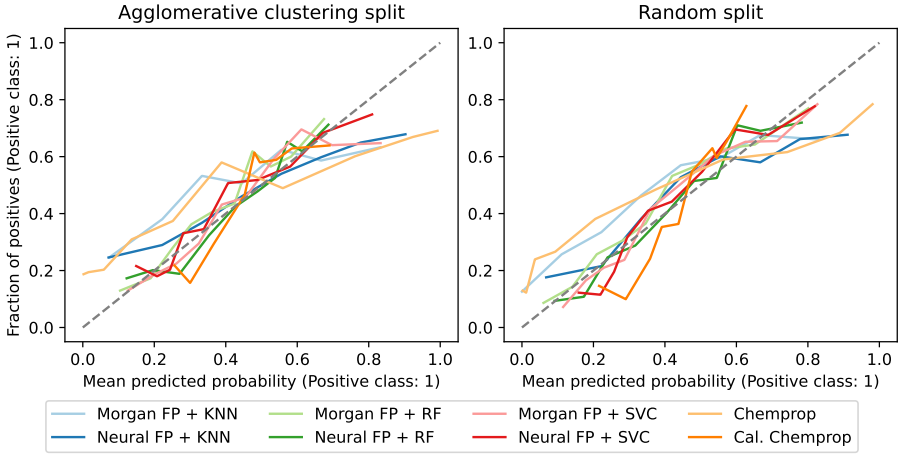

Fig. S38. Calibration curves for endpoint BSK_hDFCGF_Proliferation_down.

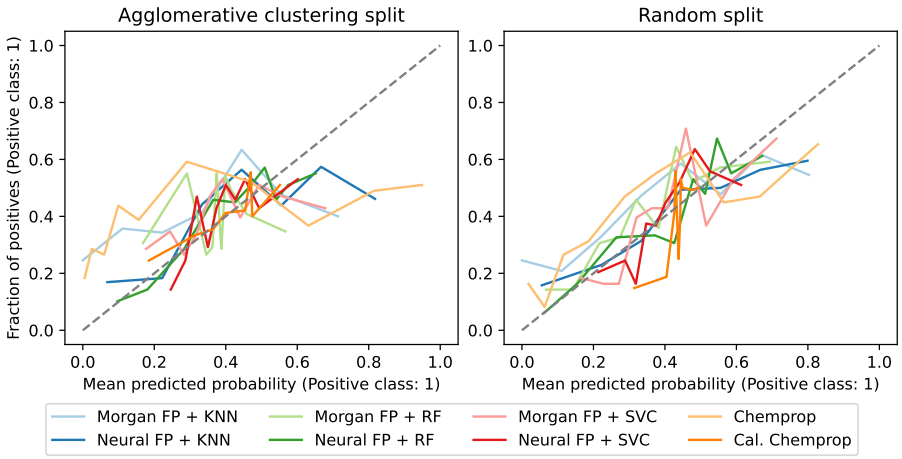

Fig. S39. Calibration curves for endpoint CEETOX_H295R_11DCORT_dn.

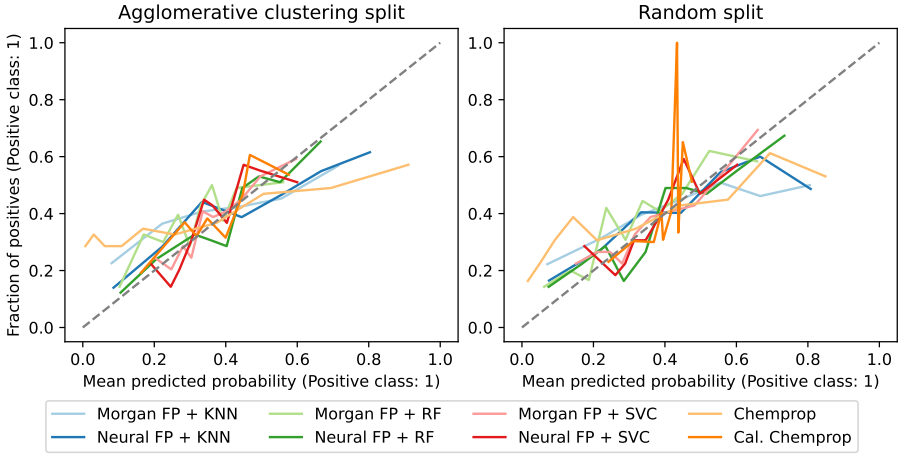

Fig. S40. Calibration curves for endpoint CEETOX_H295R_ANDR_dn.

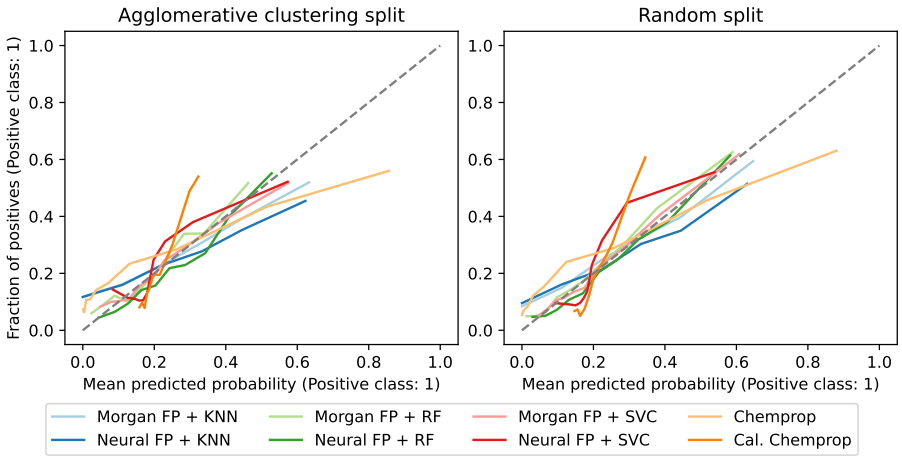

Fig. S41. Calibration curves for endpoint TOX21_ARE_BLA_agonist_ratio.

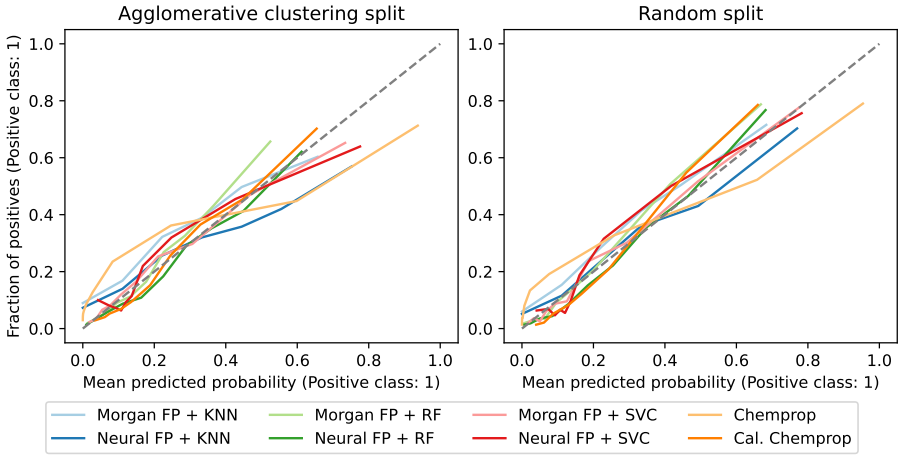

Fig. S42. Calibration curves for endpoint TOX21_TR_LUC_GH3_Antagonist.