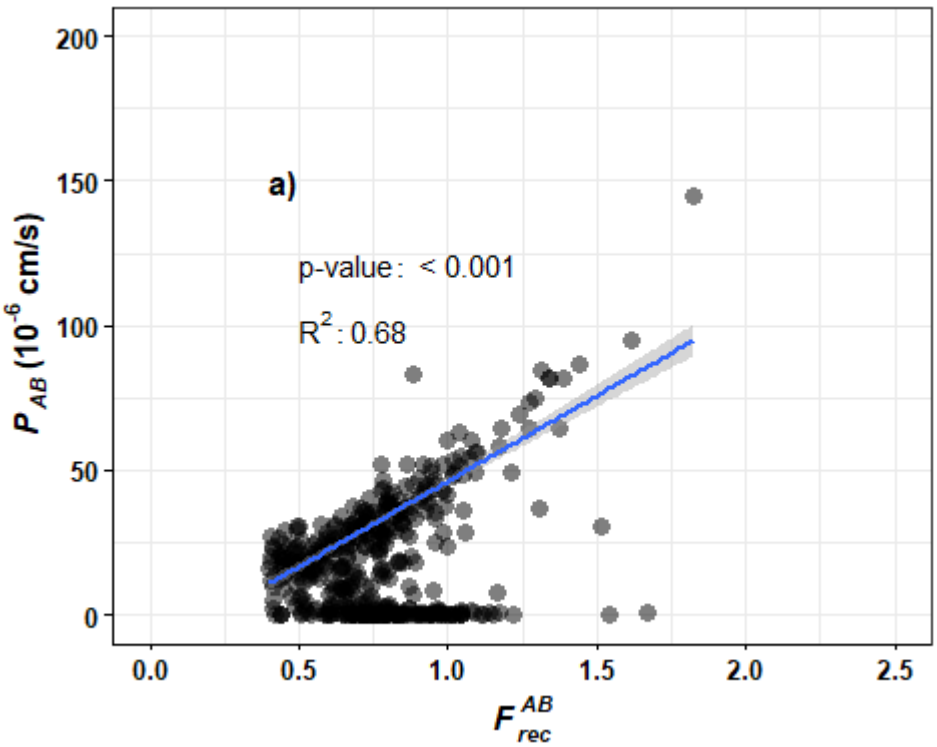
Honda et al.:

Impact of Gut Permeability on Estimation of Oral Bioavailability for Chemicals in Commerce and the Environment

Supplementary Data

1 Caco-2 data analysis

Fig. S1: Caco-2 permeability measurements (P_{app}^{AB}) plotted against recovery



doi:10.14573/altex.2403271s1

Fig. S2: Caco-2 permeability measurements (P_{app}^{BA}) plotted against recovery

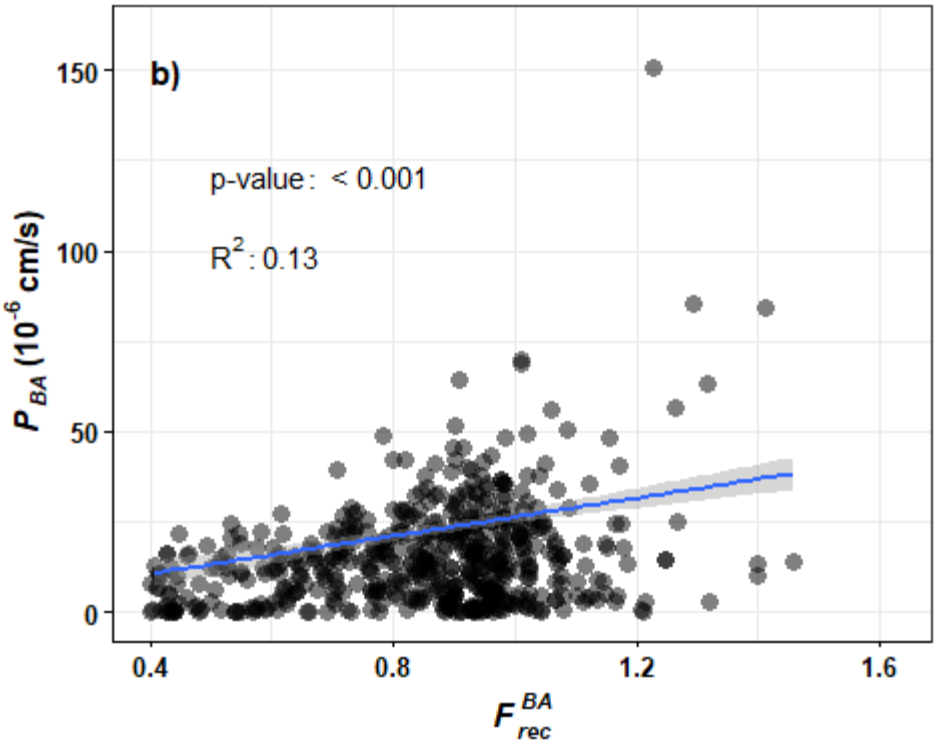


Fig. S3: Histograms P_{app}^{BA} (10^{-6} cm/s) for (A) warfarin – the high permeability control chemical, (B) talinolol – the P-pg active efflux transporter control chemical, and (C) ranitidine – the low permeability control chemical

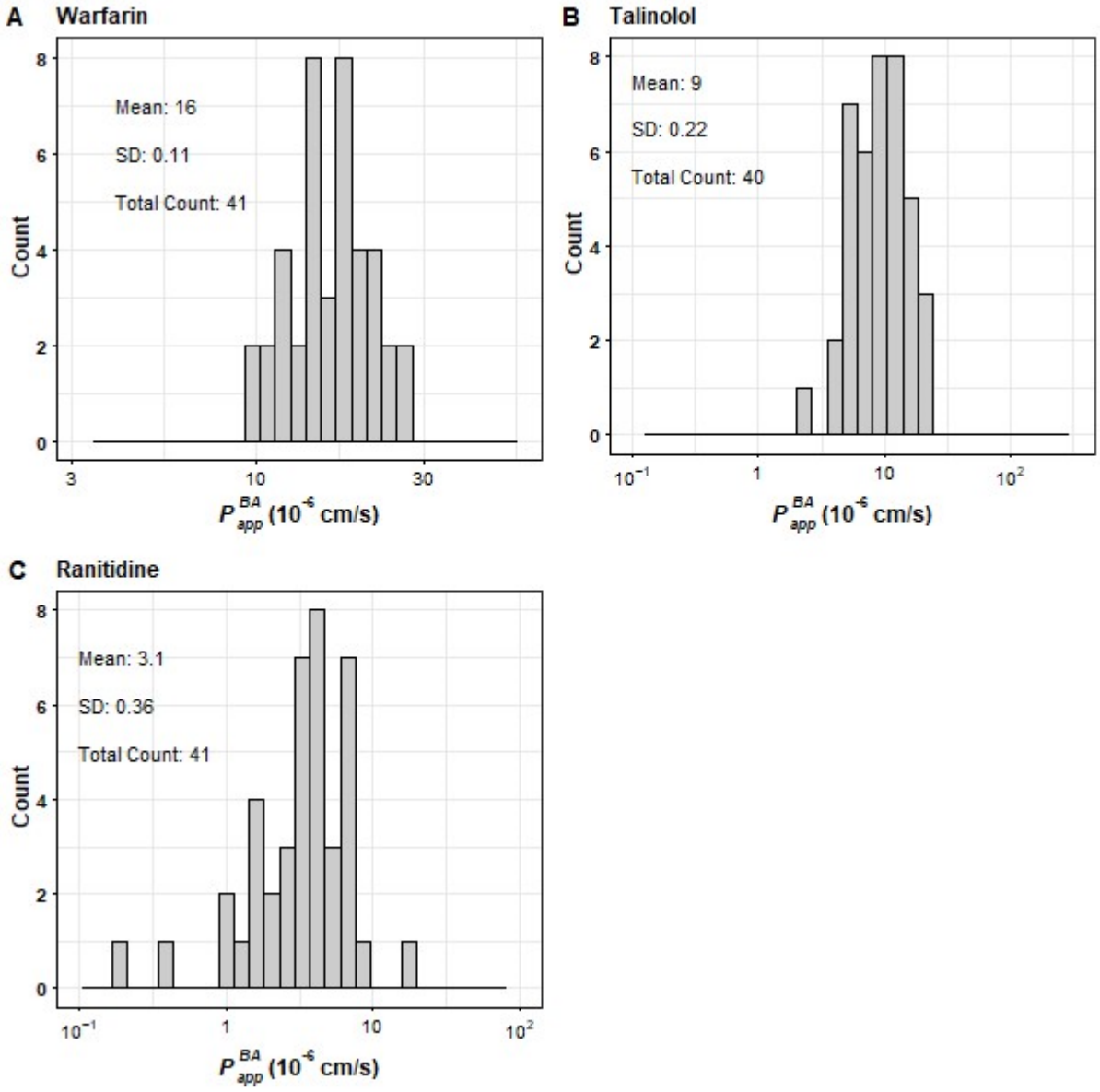


Fig. S4: Histograms of efflux ratio for (A) warfarin – the high permeability control chemical, (B) P_{app}^{AB} for talinolol – the P-pg efflux transporter control chemical, and (C) ranitidine – the low permeability control chemical
Efflux ratio histogram for talinolol is shown in the main manuscript (Fig. ~~1b~~1B).

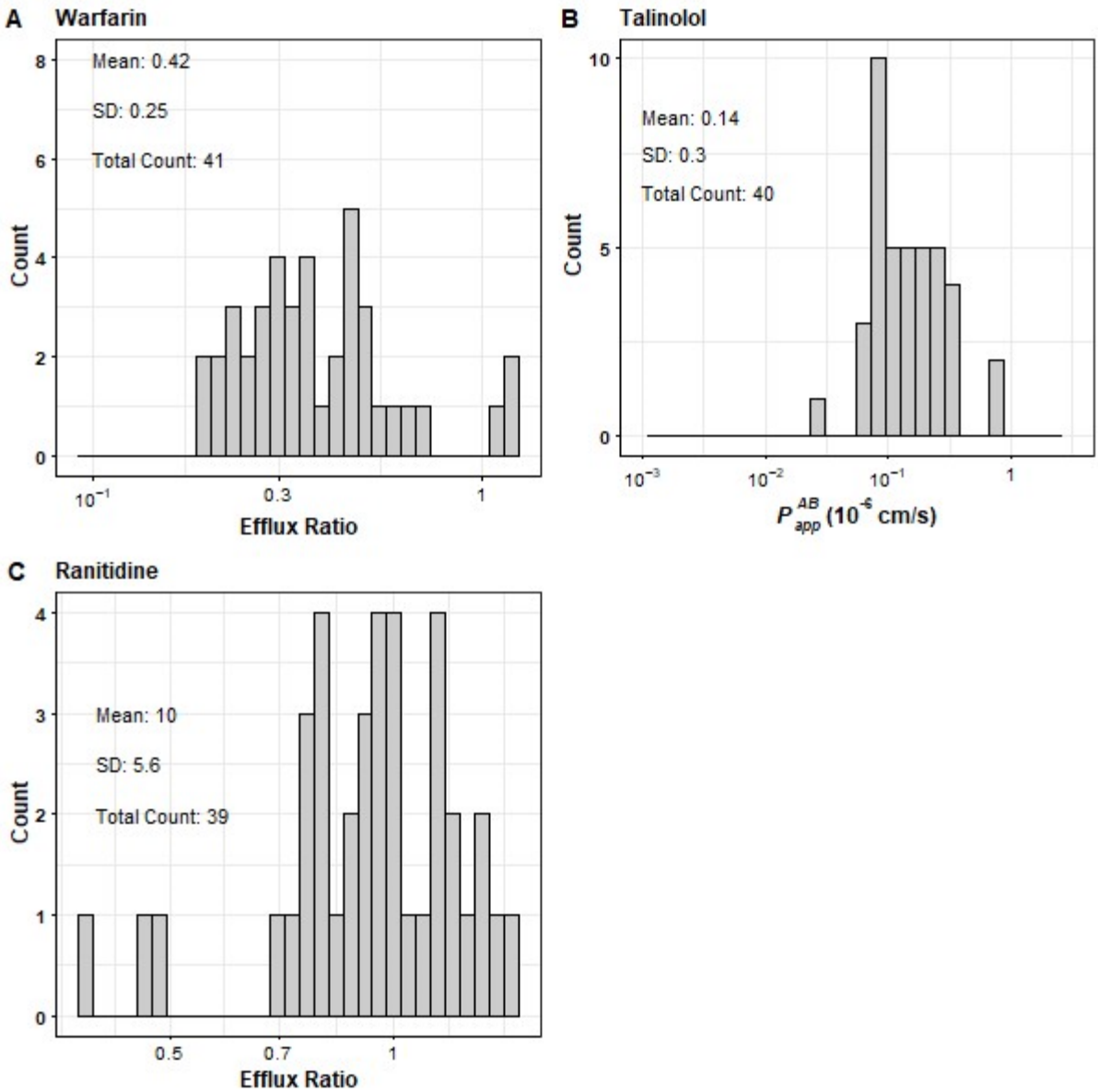


Fig. S5: Histograms of fractional recovery (F_{rec}^{AB}) for (A) warfarin – the high permeability control chemical, (B) talinolol – the P-pg active efflux transporter control chemical, and (C) ranitidine – the low permeability control chemical

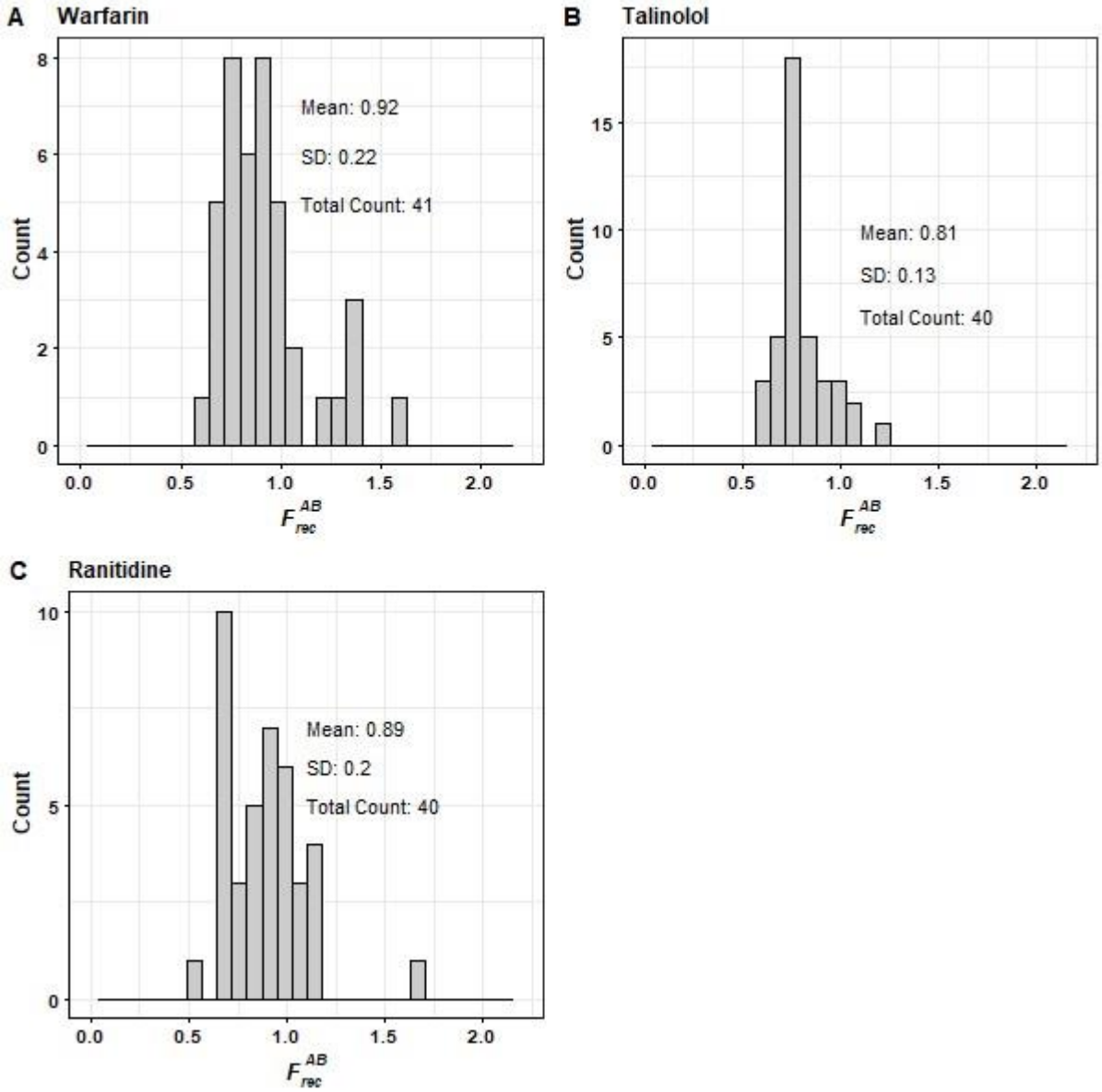


Fig. S6: Histograms of F_{rec}^{BA} for (A) warfarin – the high permeability control chemical, (B) talinolol – the P-pg active efflux transporter control chemical, and (C) ranitidine – the low permeability control chemical

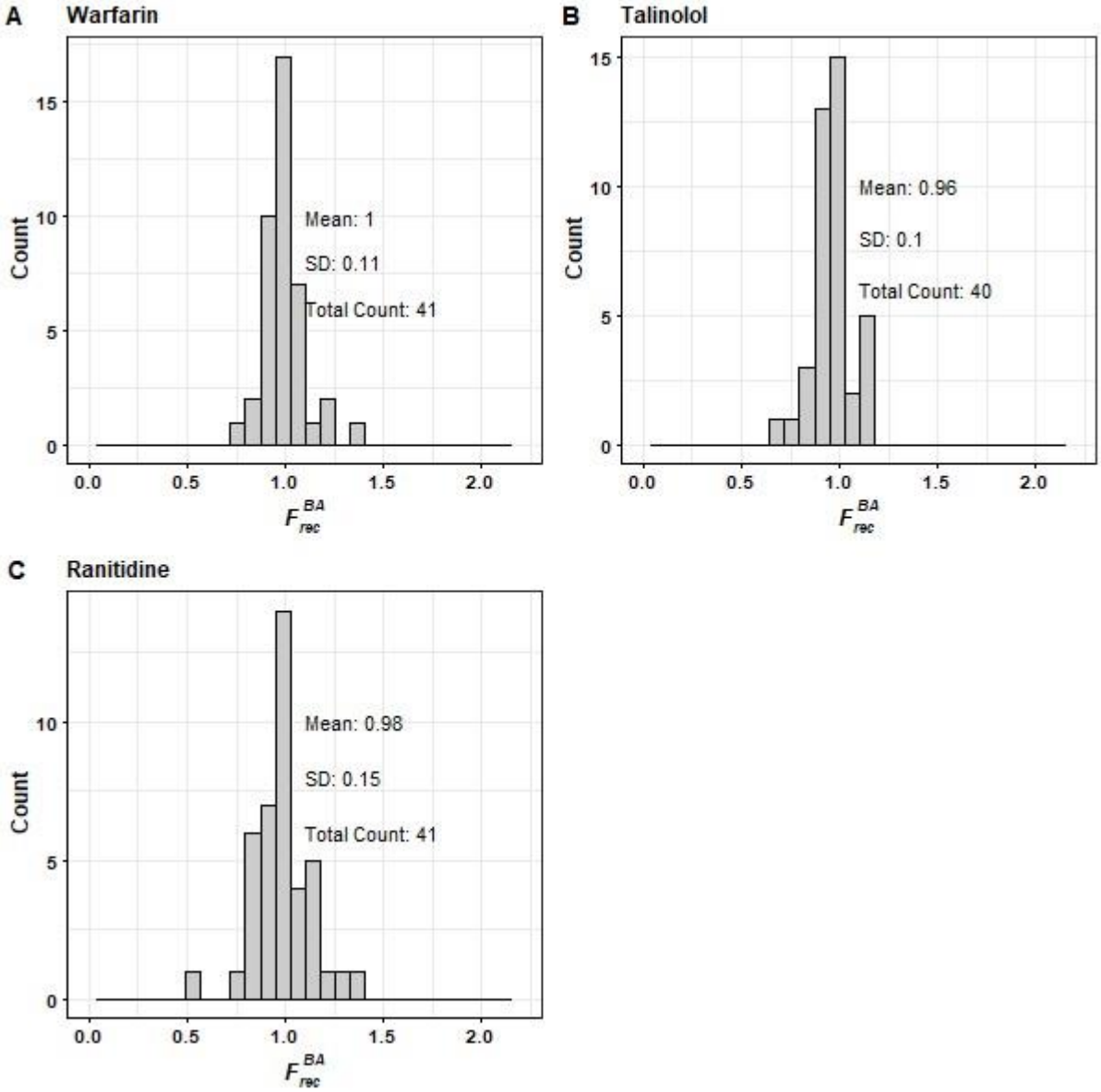


Fig S7: Histograms of fractional recovery for (A) the apical to basolateral and (B) the basolateral to apical experiments for test chemicals
The dashed vertical line corresponds to a value of one.

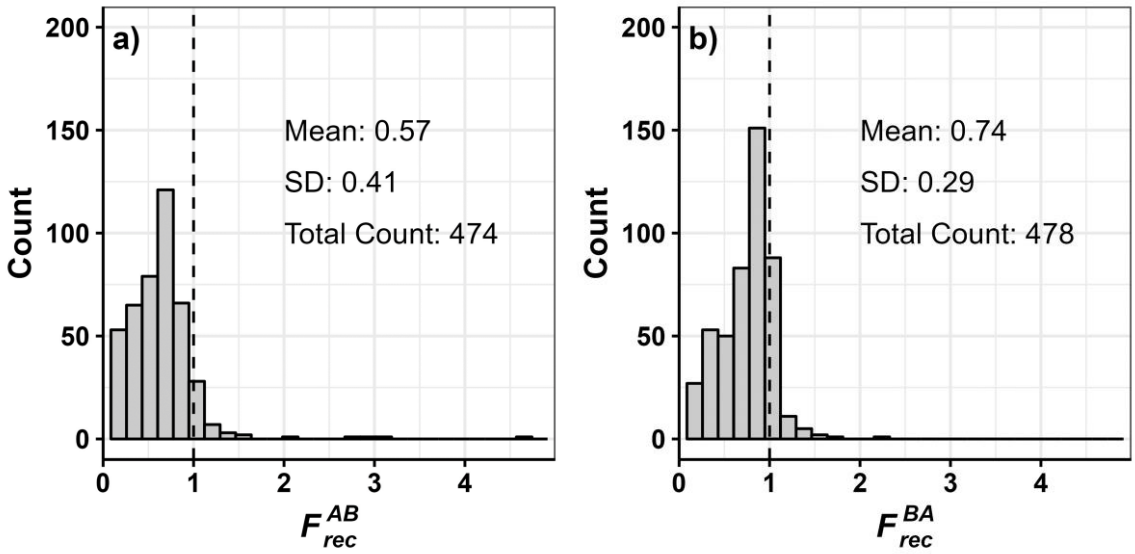
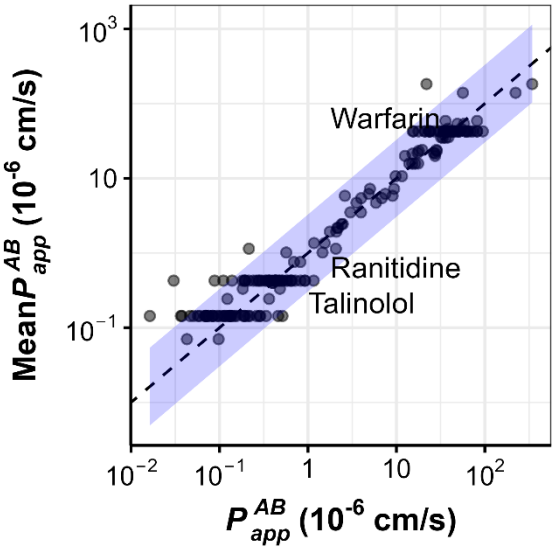


Fig. S8: Plot showing variability in repeated measurements
Individual values are plotted against chemical-specific means. The three horizontal bands correspond to the control chemicals: warfarin (high permeability), talinolol (P-pg active efflux transporter), and ranitidine (low permeability). The shaded area is the 95% confidence interval on the regression of the measured P_{app}^{AB} values on the per-chemical mean P_{app}^{AB} value.



2 QSPR model

Fig. S9: Distribution of measured Caco-2 permeabilities in the training and test sets

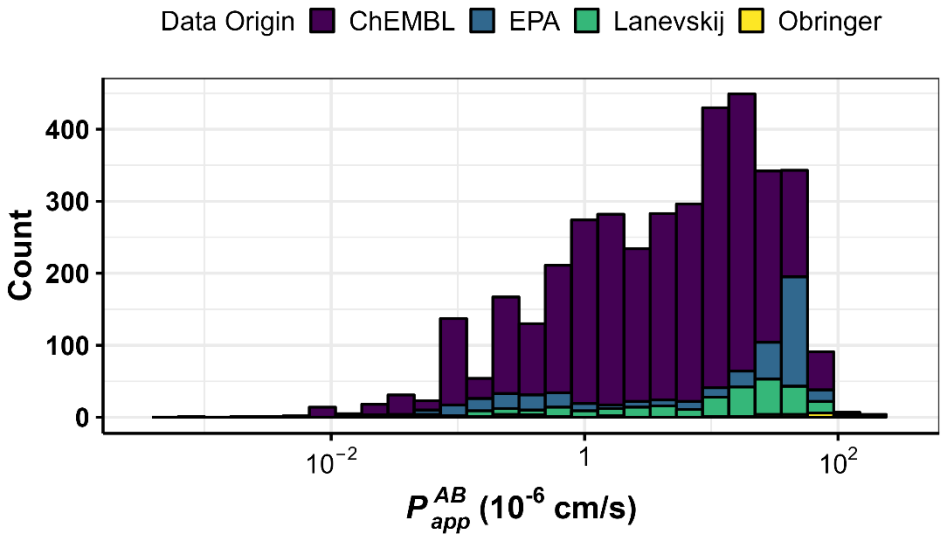


Fig. S10: Effect of recursive feature elimination
Note decline in model accuracy as the number of descriptors included was decreased from an initial set of 73.

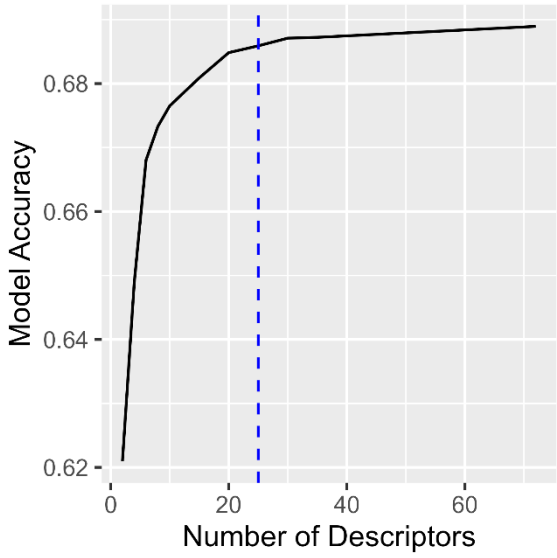


Fig. S11: Parameters included in final QSPR model and their relative importance

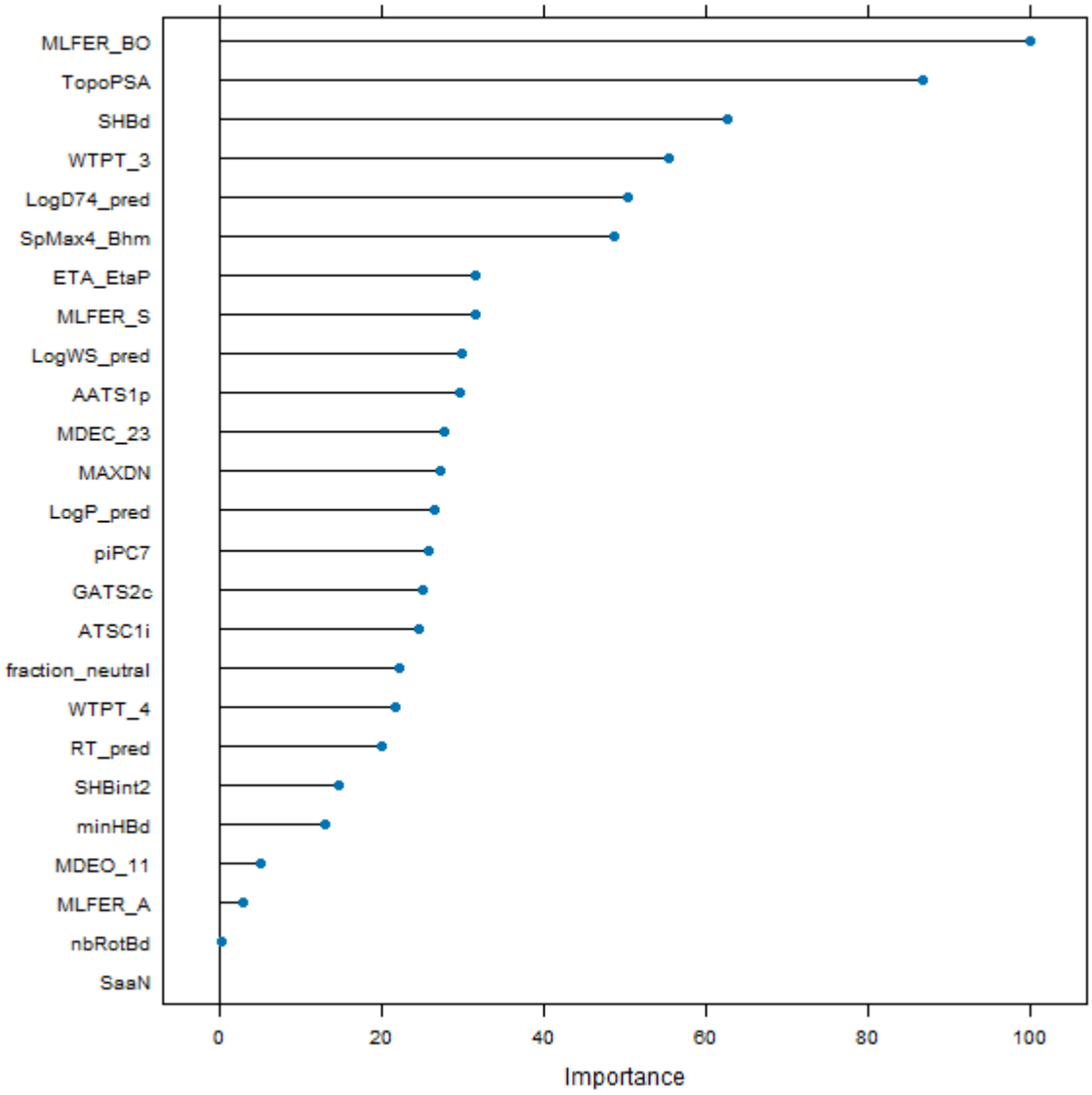


Fig. S12: Comparison of QSPR predictions to $\log_{10} P_{app}^{AB}$ for all chemicals with measured data
The dashed line is $y = x$.

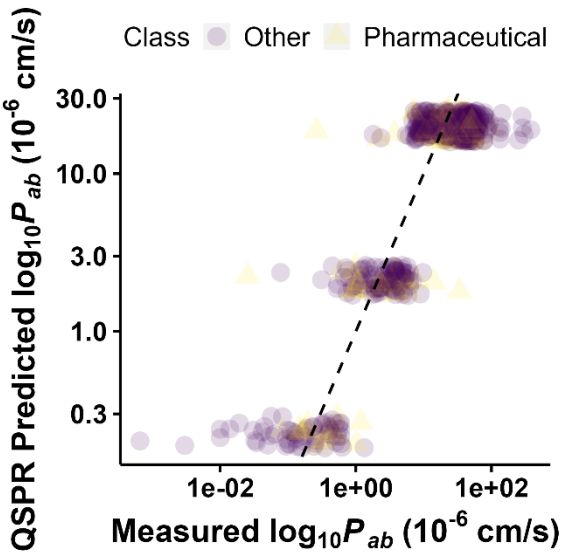
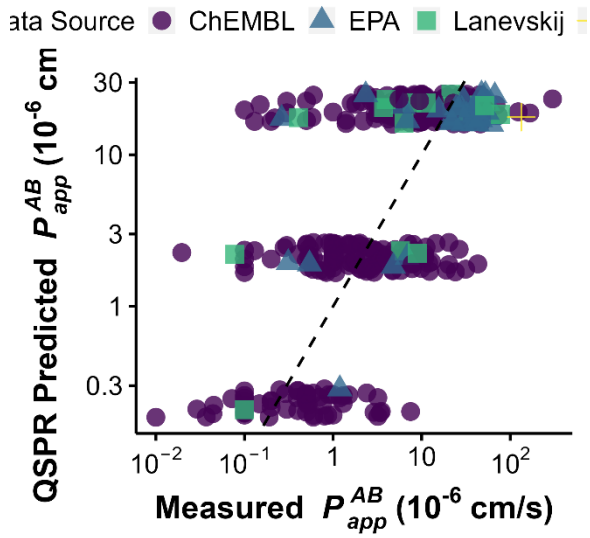


Fig. S13: Comparison of QSPR predictions to $\log_{10} P_{app}^{AB}$ for test set chemicals only (those that were withheld from training the model)
The dashed line is $y = x$.



3 Prediction of bioavailability and associated parameters

Fig. S14: Caco-2-predicted $P_{eff,h}$ vs measured $P_{eff,h}$ (left) compared to QSPR-predicted $P_{eff,h}$ vs measured $P_{eff,h}$ (right)
Data are from Dalhgren et al. (2015). The dashed line is $y = x$.

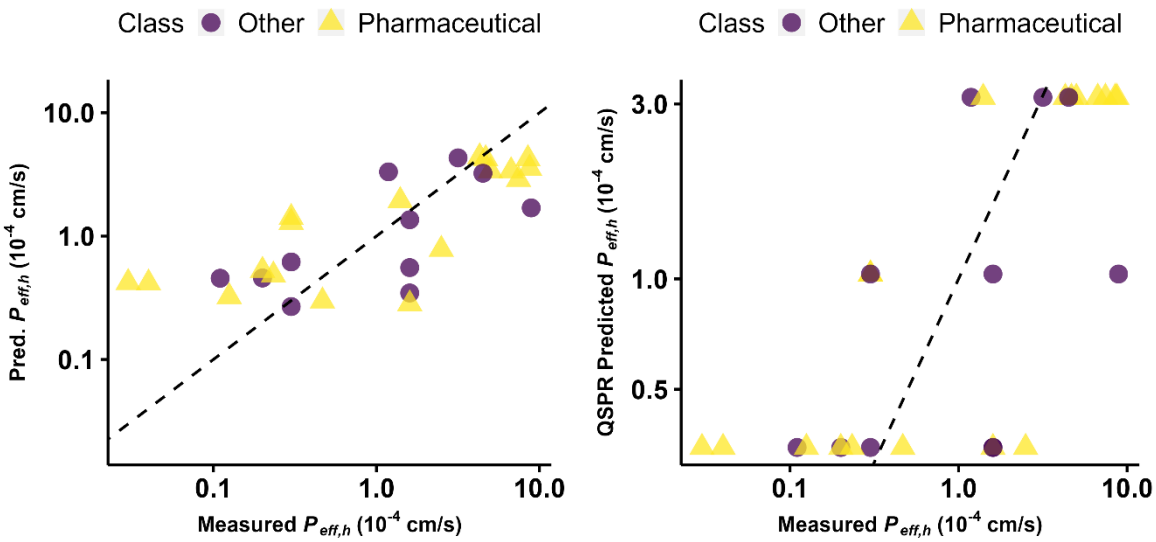


Fig. S15: Caco-2-predicted $F_{abs,h}$ vs measured $F_{abs,h}$ (left) compared to QSPR-predicted $F_{abs,h}$ vs measured $F_{abs,h}$ (right)
Data are from Varma et al. (2010). The dashed line is $y = x$.

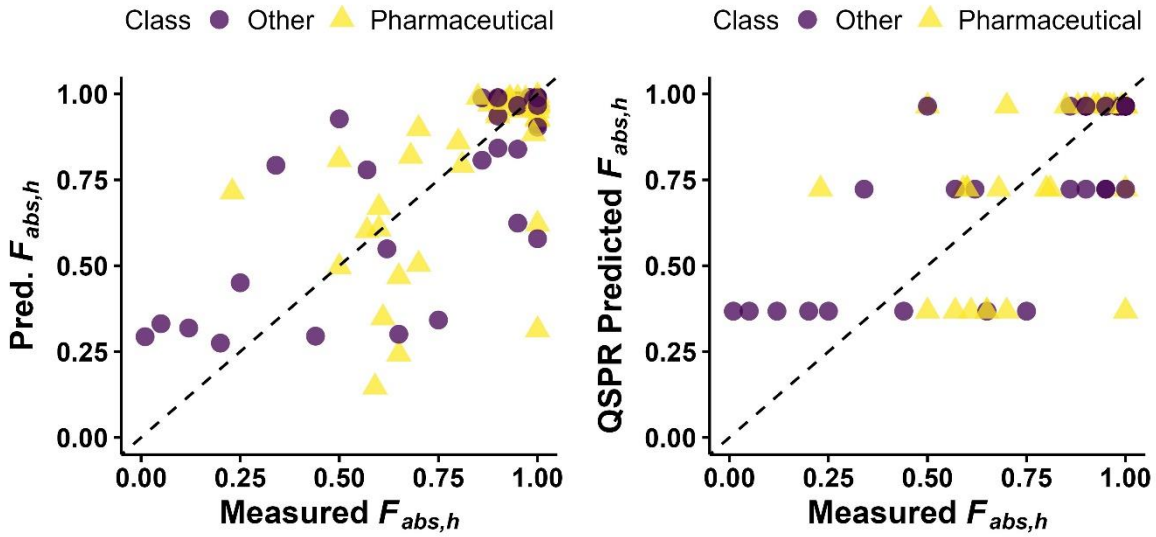


Fig. S16: Predicted $F_{gut,h}$ vs measured (top left, Q-gut model) compared to QSPR-predicted $F_{gut,h}$ vs measured $F_{gut,h}$ (top right) $F_{gut,h}$ predicted by ADMet Predictor (V. 10.0) vs measured $F_{gut,h}$ (bottom)
Data are from Varma et al. (2010). The dashed line is $y = x$.

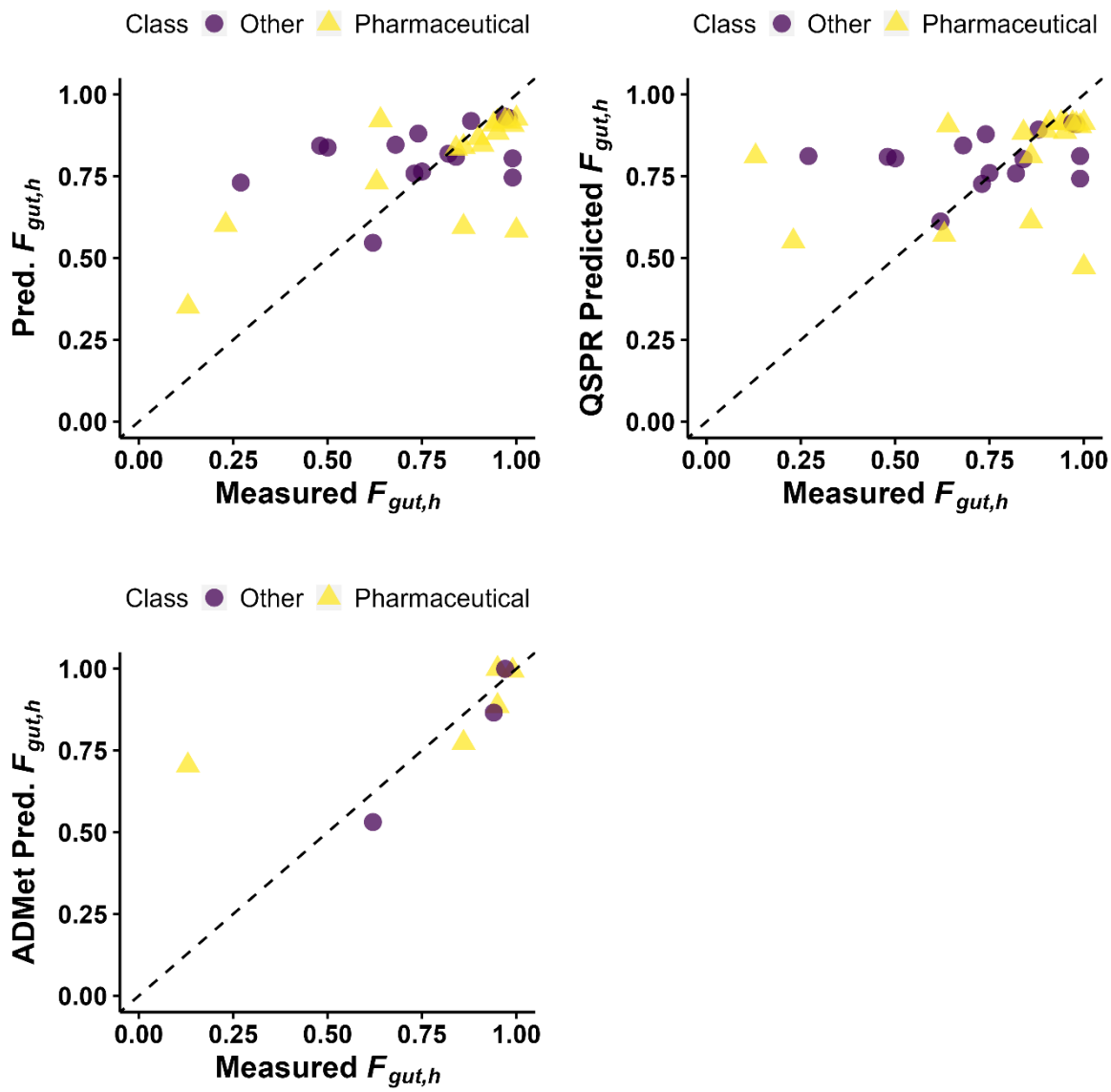


Fig. S17: Predicted $F_{hep,h}$ vs measured $F_{hep,h}$ (original *httk*, left) compared to QSPR-predicted $F_{hep,h}$ vs measured $F_{hep,h}$ (right)
Data are from Varma et al. (2010). The dashed line is $y = x$.

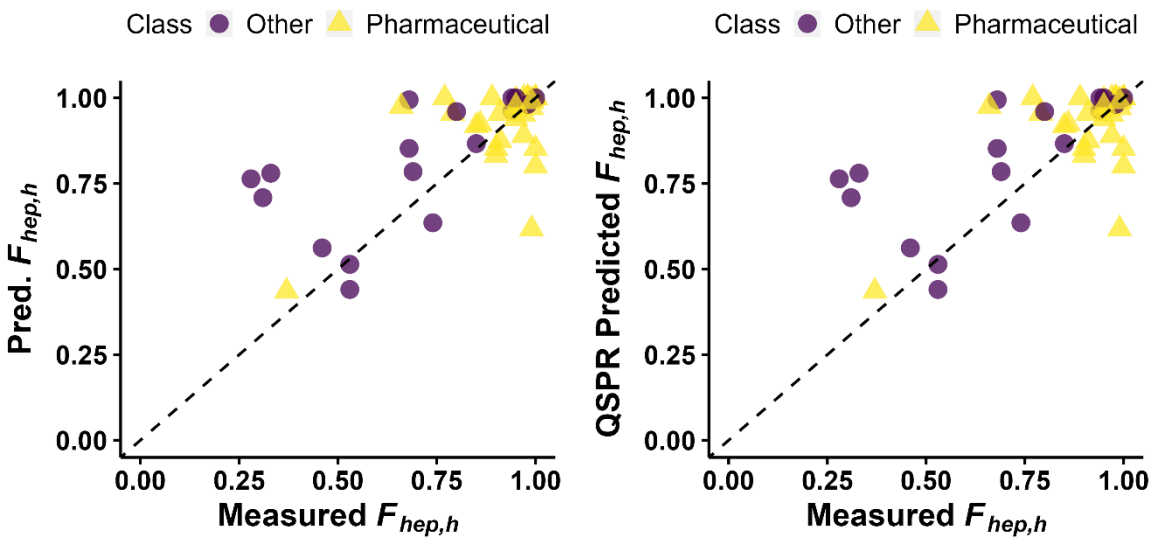


Fig. S18: Predicted vs measured $F_{bio,h}$ (top left) compared to QSPR-predicted $F_{bio,h}$ vs measured $F_{bio,h}$ (top right)
$F_{bio,h}$ predicted by ADMet Predictor (V. 10.0) vs measured $F_{bio,h}$ (bottom left). Bottom right panel shows change from original *httk* (considering only $F_{hep,h}$) to new approach incorporating improved estimates of $F_{abs,h}$ and $F_{gut,h}$. Data are from Kim et al. (2014). The dashed line is $y = x$.

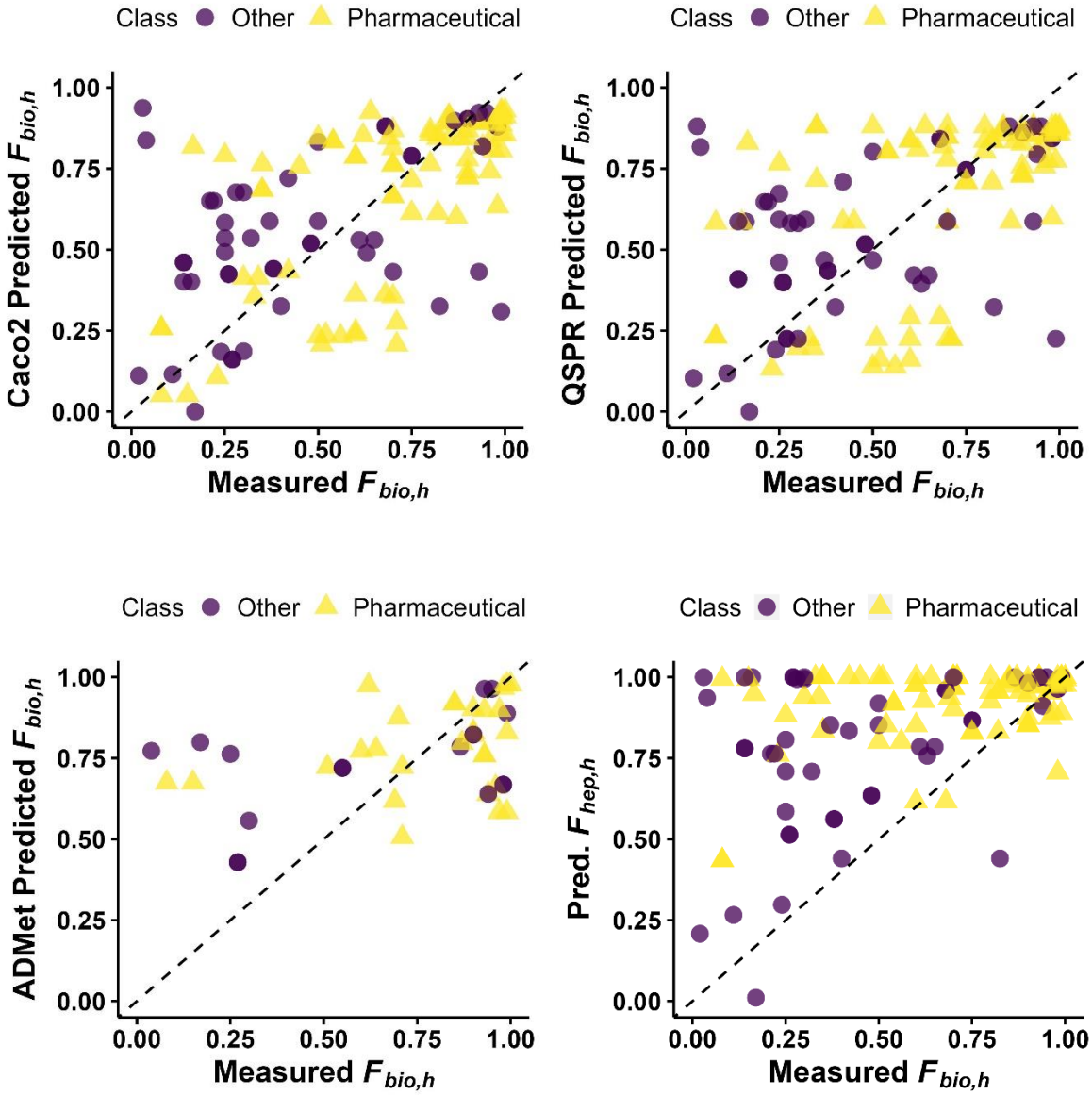
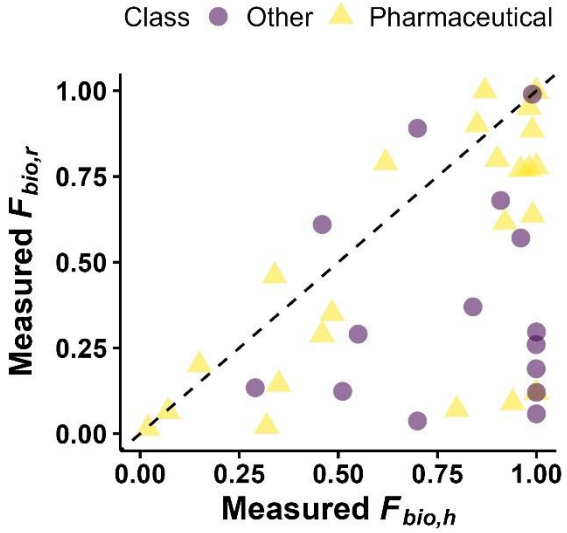


Fig. S19: Correlation between F_{bio} measured *in vivo* in humans ($F_{bio,h}$) and F_{bio} measured *in vivo* in rats ($F_{bio,r}$)
Data are from Musther et al. (2014) and Wambaugh et al. (2018). The dashed line is $y = x$.



4 BER

Fig. S20: Histogram of F_{abs} and BER
F_{abs} is less than 50% for only 6 chemicals where BER can be estimated (Tab. 4 in main manuscript).

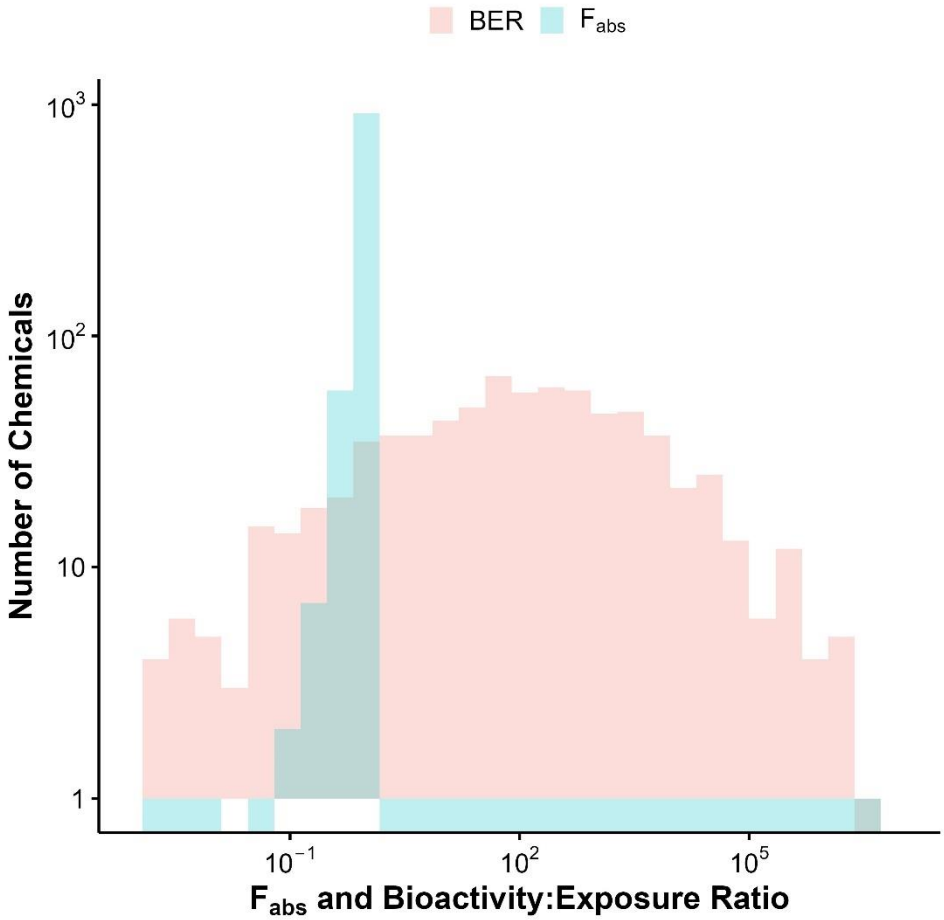
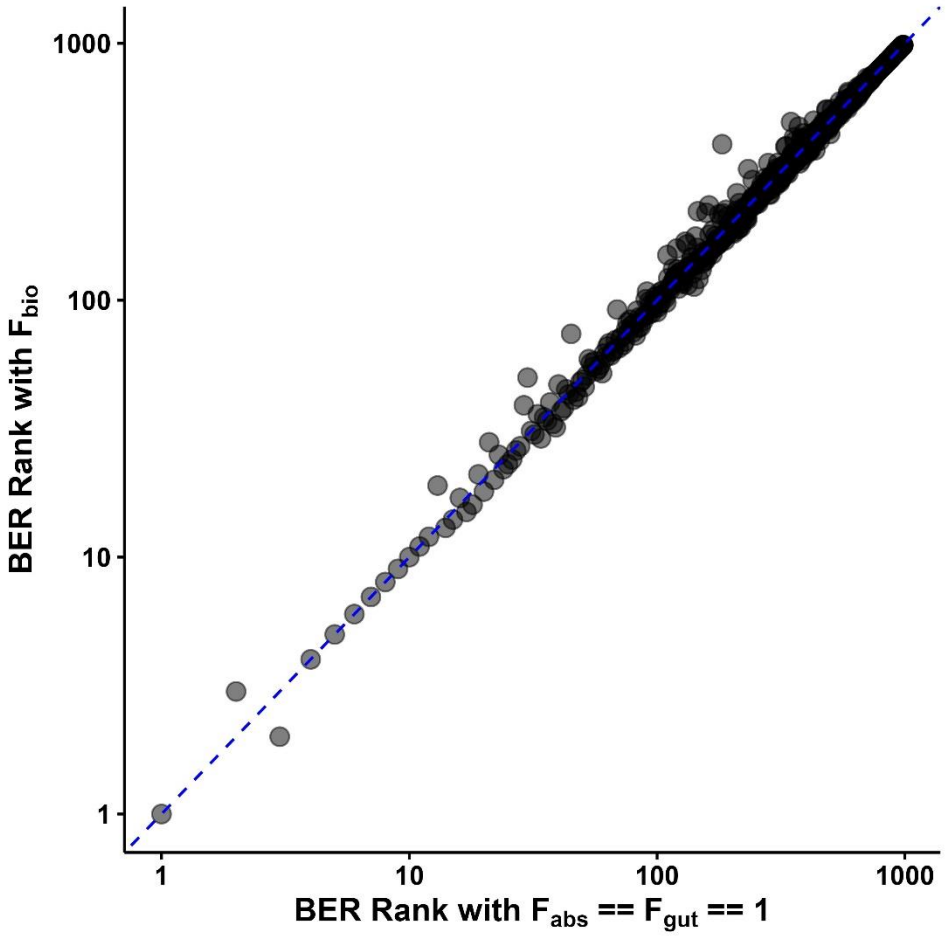


Fig. S21: Rank-rank BER plot assuming 100% absorption versus F_{bio}



Reference

Kim, M. T., Sedykh, A., Chakravarti, S. K. et al. (2014). Critical evaluation of human oral bioavailability for pharmaceutical drugs by using various cheminformatics approaches. *Pharm Res 31*, 1002-1014. doi:10.1007/s11095-013-1222-1