Supplementary Methods

Preparation of a gold standard DDI dataset. The gold standard DDI dataset for DeepDDI was prepared using DrugBank[1] (version 5.1.1) to cover 222,127 DDIs from 222,041 drug pairs (Supplementary Data 1), which represent 113 DDI types (Spplementary Data 2). The newly prepared dataset was split into training (80%) and test (20%) datasets for the model development; within the training dataset, 10% was used for validation.

Generation of structural similarity profiles. A structural similarity profile (SSP) for an input drug was generated in the same manner as previously described[2], but with a greater number of approved drugs than the previous version DeepDDI: from 2,159 approved drugs to 2,386 approved drugs. A list of the approved drugs was obtained from DrugBank[1] (version 5.1.1). These approved drugs served as so-called 'fixed comparison targets'. Briefly, a SSP was generated by obtaining Tanimoto coefficient between an input drug and 2,386 approved drugs in a pairwise manner. Subsequently, two SSPs of each drug in an input pair were reduced to a dimension of 50 using principal component analysis, and merged to generate a feature vector, namely 'a combined SSP' that has 100 dimensions. The combined SSP was finally processed by a multilayer perceptron (MLP) of DeepDDI to predict DDI types for the input drug pair.

Development of a deep learning model for predicting DDI types. The updated DeepDDI has 113 output neurons, each of which represents a specific DDI type. Definitions of 113 DDI types are available in the form of sentences in Supplementary Data 2. DeepDDI was designed to perform the multi-label classification where a drug pair can be assigned with multiple DDI types. To train a DNN in DeepDDI, weights of the DNNs were randomly initialized based on the uniform distribution. For activation functions, the rectified linear unit (ReLU) was used for hidden layers, while the sigmoid function was used for the output layer. The DNN was trained

1

up to 300 epochs with early stopping set at 10 epochs, using binary cross-entropy as a loss function and Adam method[3] for optimization. In addition, the Monte Carlo (MC) dropout[4] was used to measure the prediction uncertainty (i.e., standard deviation) by generating ten MC dropout samples during the model training and for the new predictions. The mean and standard deviation of the ten prediction results are also generated as an output of the DeepDDI.

Hyperparameters, including the number of hidden layers, the number of nodes per hidden layer, learning rate, and dropout rate, were optimized with respect to F1 score via Bayesian optimization[5] (Supplementary Table 1). By examining 50 different architectures for the DNN, three hidden layers, 1,024 nodes for each layer, learning rate of 0.0069, and dropout rate of 0.3 showed the most optimal performance for the DeepDDI. Threshold for an output neuron's activity (predictive mean - half a standard deviation) was set at 0.27; having this threshold resulted in the best F1 scores among the threshold values between 0 to 1 with an interval of 0.01.

Processing of the initial prediction results from the DNN. DDI types initially predicted by the DeepDDI were considered valid if they satisfy the following two criteria[6]; 1) a drug should exist in the gold standard DDI dataset, which has a similar structure as one of the two input drugs (Tanimoto coefficient > 0.7); and 2) these structurally similar drugs from DeepDDI and the dataset should be associated with the same DDI type.

Preparation of a dataset for prescription drugs from FDA Adverse Event Reporting System (FAERS). Of the 2,159,674 cases (prescriptions) reported in the FDA FAERS database from Q4 2012 to Q2 2018 (https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html), only those involving two to five drugs, all assigned with DrugBank IDs, were considered in this study. As a result, 2,248 prescription drugs from the 301,615 FAERS cases

were considered in this study. As a result of running the DeepDDI, 1,628 prescription drugs were predicted to have DDIs with nirmatrelvir and/or ritonavir (Supplementary Data 3).

Development environment and data availability. All the computations were implemented by using Python 3.6 in the Linux Ubuntu environment. A Python package Keras (version 2.0.6) (https://keras.io/) with the TensorFlow backend[7] (version 1.2.1) was used to develop the DNN of DeepDDI. Source code for the updated DeepDDI is available at https://bitbucket.org/kaistsystemsbiology/deepddi2.

Supplementary Tables

Supplementary Table 1. Hyperparameters examined for the DNN in the DeepDDI.

Hyperparameter	Tested values	Optimal value for the DDI type prediction
Number of hidden layers	3, 4, 5, 6, 7, 8	3
Number of nodes per layer	256, 512, 1024	1024
Dropout rate	0.3, 0.4, 0.5	0.3
Learning rate	log0.005 - log0.2	0.0069

Supplementary References

1. Wishart, D.S.*, et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074-D1082 (2018).
2. Ryu, J.Y., Kim, H.U. & Lee, S.Y. Deep learning improves prediction of drug-drug and drug-food interactions. *Proc. Natl. Acad. Sci. USA* **115**, E4304-E4311 (2018).
3. Kingma, D.P. & Ba, J. Adam: a method for stochastic optimization. *arXiv* (2014).
4. Gal, Y. & Ghahramani, Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. in *Int. Conf. Mach. Learn.* 1050-1059 (2016).
5. Snoek, J., Larochelle, H. & Adams, R.P. Practical bayesian optimization of machine learning algorithms. in *Adv. Neural Inf. Process. Syst.* 2951-2959 (2012).
6. Vilar, S.*, et al.* Similarity-based modeling in large-scale prediction of drug-drug interactions. *Nat. Protoc.* **9**, 2147-2163 (2014).
7. Abadi, M.*, et al.* TensorFlow: a system for large-scale machine learning. *Proceedings of Osdi'16: 12th Usenix Symposium on Operating Systems Design and Implementation*, 265-283 (2016).