|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DATABASE** | **DRUG LABEL** | | | | **GENE LABEL** | | | | | **CRITERIA** |
| **SMILES** | **IUPAC NAME** | **INCH** | **ATC CODES** | **BINDING AFFINITY** | **GENE SYMBOL** | **ENTREZ ID** | **ENSEMBL ID** | **UNIPROT ID** |
| BindingDB | √ |  | √ |  | √ |  |  |  | √ | (i) Binding affinity, including at least one of Ki, Kd, IC50 or EC50 ≤ 10 μM; (ii) the presence of UniProt ID representation of the protein; and (iii) the protein being tagged as "review". |
| ChEMBL | √ | √ | √ | √ | √ | √ |  | √ | √ |
| Gtopdb | √ | √ | √ |  | √ | √ |  | √ | √ |
| pubchem | √ | √ | √ | √ | √ | √ |  | √ | √ |
| TTD | √ |  | √ |  | √ | √ |  |  | √ |
| DGIdb | √ | √ |  |  |  | √ | √ |  |  | Fourteen of these experimentally validated database sources were screened for drug target interactions and eight were excluded. |
| Drugbank | √ | √ | √ | √ |  | √ |  | √ | √ | The drug target interaction relationships downloaded from the Drugbank database have been experimented with to remove interactions in which no clear target information is available. |
| PharmGKB | √ | √ | √ | √ |  | √ |  | √ |  | In the PharmGKB, there are 3 types of drug-target relationships: "associated", "not associated" and "ambiguous", and only the data marked as "associated" are selected. |
| CancerDR | √ |  |  |  |  | √ |  | √ |  | CancerDR database is mainly used for the study of drugs and sensitive cell lines, but lacks the information of the corresponding target effects |
| STITCH | √ |  |  |  |  | √ |  |  |  | In STITCH and SuperPred, the majority of drug target associations are based on computational prediction, but not confirmed by biological experiments, which do not meet the high confident purpose of this study |
| SuperPred | √ |  |  | √ |  | √ |  |  |  |
| SIDER | √ |  |  | √ |  |  |  |  |  | SIDER database is used for the study of drug sider effects but lacks target information |
| KEGG | √ | √ | √ | √ |  | √ |  | √ | √ | In KEGG,the drug-target information are simply collected from ChEMBL, Drugbank and Pubchem, without selection or integration |

