

Supplementary Information: Table S1 Summary Statistics of Drugboxes and Chemboxes

# of Infoboxes	Drugbox Pages	Chembox Pages
1	8284	12237
2	41	89
3	6	9
4	0	1
5	1	1
Total	8332	12337

# of DSSTox Mappings	Drugboxes	Chemboxes
0	874	201
1	6433	10445
2	974	1443
3	95	277
4	12	57
5	1	20
6	0	6
7	0	2
Total	8389	12451

“Naïve” Missing Identifiers	Drugboxes	Chemboxes
No DTXSID	3526	3174
No CASRN	617	510
No SMILES	1533	665
No InChIKey	1882	1425
No SMILES & No InChIKey	1479	617

“True” Missing Identifiers	Drugboxes	Chemboxes
No CASRN	123	153
No SMILES	272	284

No InChIKey	419	695
No SMILES & No InChIKey	201	219

Disagreement Type	Drugboxes	Chemboxes
CASRNs Disagree	103	156
InChIKey Disagrees w/ DSSTox Structure	699	1415
SMILES Disagrees w/ DSSTox Structure	893	1613
WP InChIKey & Smiles Disagree w/ Each Other	311	453