# Supplementary Information: The SIDER database of drugs and side effects

## Supplementary Table 1: UMLS Semantic Types included in SIDER

|  |  |
| --- | --- |
| Semantic Type  (subtypes are also included) | Description |
| A1.2.2 | Anatomical Abnormality |
| A2.2 | Finding |
| B2.2.1.1 | Physiologic Function |
| B2.2.1.2 | Pathologic Function |
| B2.3 | Injury or Poisoning |

## Supplementary Table 2: Regular expressions for excluding sentences from text-mining

|  |
| --- |
| **ADR and Indications section** |
| neither |
| risk (of|factor) |
| not (been )?(associated|observed|reported|recommended) |
| (associated|observed|reported) with other |
| is suspected |
| disease[- ]related |
| from causes other |
| from other causes |
| no evidence |
| not (accompanied|indicated|expected|suspected) |
| no( significant) difference |
| (single|case) report |
| **Indications section** |
| (were|is|may be|not) ([^ ]+ )? (associated|observed|reported|impaired|demonstrated|shown) |
| (may|might|can|could) occur |
| has caused |
| causes |
| increased incidence |
| see (contraind|overdos|abuse|dosage|warnings|precautions|interactions|symptoms and treatment of overdos|(description of)?clinical studies|posology) |

## Supplementary Table 3: Parts of words used for the detection of indications by Natural Language Processing

|  |  |
| --- | --- |
| **Whitelist** | **Blacklist** |
| ameliorat | not\* |
| associated | caused |
| control |  |
| decrease rate |  |
| effective |  |
| heal |  |
| improve |  |
| indicated |  |
| maintain |  |
| maintenance |  |
| management |  |
| prevent |  |
| prophylaxis |  |
| reduce |  |
| reduction |  |
| relief |  |
| relieve |  |
| shown patients |  |
| suffering |  |
| symptoms |  |
| therapy |  |
| treat |  |
| use\* |  |
| used |  |
| useful |  |

\* match only complete word

## Supplementary Table 4: Detection of pre-existing conditions of patients using Natural Language Processing

|  |  |
| --- | --- |
| **Governor** | **Relation** |
| children | prep\_with |
| history | prep\_of |
| patient | prep\_with |
| patients | nn |
| patients | num |
| patients | prep\_with |
| studies | nn |
| studies | num |
| studies | prep\_in |
| studies | prep\_of |
| study | nn |
| study | nn |
| treated | prep\_for |
| treatment | prep\_of |
| trial | nn |
| trial | num |
| trial | prep\_for |
| trial | prep\_in |
| trial | prep\_of |
| trials | nn |
| trials | num |
| trials | prep\_for |
| trials | prep\_in |
| trials | prep\_of |