

Supplementary Material – Publication II

S1. Supplementary files description

Explanation of the .xlsx files provided along with the manuscript.

- **S1.Literature_cases.xlsx**
DR cases annotated in the tables of the three selected publications.
 - Xue et al.
 - Jarada et al.
 - Li and Jones.
- **S2.UMLS_triples.xlsx**
Separately normalization of the DR cases by CUIs' extraction:
 - Xue et al.
 - Jarada et al.
 - Li and Jones.Selection of the unique triples:
 - Unique Triples
- **S3.DISNET_triples.xlsx**
DISNET mapping result triples:
 - DISNET VocabulariesFinal triples with genes, symptoms, drug target genes and categories information:
 - DISNET Final Triples
- **S4.Genes_analysis.xlsx**
Summary of the number of common genes for each DR case:
 - SummaryGenes associated with each DR disease and genes in common of each DR triple:
 - D-D GenesDrug targets information (protein target, gene that encodes it and type of action executed by the drug):
 - Drug Targets InfoDR triples that share the gene that encodes the drug target. GDAs included.
 - Target Genes Data
- **S5.Symptoms_analysis.xlsx**
Summary of the number of common symptoms for each DR case:
 - SummarySymptoms associated with each DR disease and symptoms in common of each DR triple:
 - D-D Symptoms
- **S6.Categories_analysis.xlsx**
MeSH-PA categories for DR drugs:
 - DrugsICD-10-CM categories for DR diseases:
 - Diseases
- **S7.DR_validation_summary.xlsx**
DR cases and sources employed:

- DR Cases
- Publications

DR triples employed:

- DR Triples

Summary of the number of genes, symptoms and drug target genes shared by each DR case:

- DR Cases Numbers

S2. Statistical significance of the differences between DR and non-DR via Welch T-Test

Table 1. Comparison of DR diseases vs DISNET diseases GDA score distributions via Welch T-Test. Each set's GDA scores' mean \pm their standard deviation are indicated. The last column includes the p-value resulted from the test.

	"DR Disease – Shared Target Gene" score	"DISNET Disease – Gene" score	p-value
OD	0.3198701 \pm 0.240605	0.1399051 \pm 0.1317480	5.77e ⁻⁰⁷⁹
ND	0.2155844 \pm 2105412		0.02306

Table 2. Comparison of actual DR triples vs random DISNET triples GDA score distributions via Welch T-Test. GDA scores' mean \pm their standard deviation are indicated for each type of association in random and DR triples. The last column includes the p-value resulted from the test.

	DR triples GDA score	Random triples GDA score	p-value
OD – Shared Target Gene	0.3198701 \pm 0.240605	0,164262 \pm 0.156046	2.45e ⁻⁰⁷
ND – Shared Target Gene	0.2155844 \pm 2105412	0,1598439 \pm 0.145139	0.02306

Table 3. Comparison of DR diseases vs DISNET diseases distances via Welch T-Test. Mean of the symptom-based distances \pm their standard deviation are indicated for DR and DISNET disease pairs. The last column includes the p-value resulted from the test.

DR OD – ND	DISNET D – D	p-value
0.964919034 \pm 0.04468276	0.97146033 \pm 0.04377608	0.03971