

Analysis of orphan diseases with a KNIME workflow using Open PHACTS, with the potential of drug repurposing

Jana Gurinova, Daniela Digles, Gerhard Ecker

University of Vienna, Department of Pharmaceutical Chemistry, Althanstraße 14, 1090 Vienna, Austria

Worldwide an estimated number of 400 million people are affected by orphan diseases [1]. An orphan disease is defined as affecting less than 1 in 2000 citizens [2]. Such a low prevalence coupled with the sheer number of orphan diseases, estimated to be about 5000-8000, is the main reason for the small number of marketing approvals, amounting to treatments for roughly 200 conditions in the US and only about 45 in the European Union [1]. Drug repurposing therefore embodies an attractive option of reaching many patients with treatments that have already been deemed safe. This work aims at providing an overview of linked targets and drugs for orphan diseases, with a special focus on involved transporters.

The biggest European platform for orphan diseases is Orphanet [3], with comprehensive information for patients as well as for professionals. Orphanet also provides identifiers for the listed diseases as well as genes, which are perfectly suitable for data integration across different databases and thus also data enrichment with the aim of understanding and visualizing the role of diverse protein classes such as ion channels or transporters in orphan diseases. The Open PHACTS Discovery Platform has proven invaluable for exactly this kind of data integration given that the included databases such as UniProt, DisGeNET and DrugBank brought together enable the user to analyse the data based on various criteria as well as to draw previously unseen conclusions.

The workflow itself was established by using the KNIME Analytics Platform and starts with extracting data from Orphanet such as the name of the disease and its UMLS identifier. This key component is then connected to the Open PHACTS Discovery Platform through the use of Open PHACTS nodes available from <https://github.com/openphacts/OPS-Knime>, and yields further information such as which targets are involved in the disease plus their respective UniProt identifiers, and ultimately the drugs for these targets and their stages of development including the ones that are already approved.

As of now, approximately a third of the 9000 disorders including their sub-types listed in Orphanet are equipped with diverse identifiers, 2901 of which are UMLS identifiers used in the workflow for further analysis. Additionally, previously unknown links of diseases to drugs may be investigated for their potential regarding drug repurposing.

Acknowledgements

The work has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no. [115191], resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and in-kind contribution of EFPIA companies.

References

- [1] Kaplan W et al. Priority Medicines for Europe and the World Update 2013, *World Health Organization*, Geneva, 2013
- [2] EURORDIS. Rare Diseases: understanding this Public Health Priority. *EURORDIS*, Paris, 2005
- [3] Orphanet: an online database of rare diseases and orphan drugs. Copyright, INSERM 1997. Available at <http://www.orpha.net> (accessed December 2015)