University of Nebraska, Duke University UCD-SNOMEDCT Graph Application

Example Graph Set 1

Tom Balmat, Duke University Research Computing

Following are example graphs generated using the visNetwork package of R in a web-enabled Shiny application. Source data are anonymized urea cycle disorder (UCD) participant observations with associated SNOMEDCT encoded clinical diagnoses and RxNorm encoded prescriptions. The data were supplied by the University of Nebraska Medical Center and the Duke University School of Nursing. Observations reside in a Neo4j graph database and were queried using Neo4j's Cypher query language along with user supplied values from prompts presented by the Shiny app. The intent is to illustrate methods of exploring relationships between select groups of participants, SNOMEDCT concepts, and prescriptions, using interactive graph visualization.

Notes and observations:

- Vertex (node) and edge size corresponds to observation frequency (number of observations). The size of connected vertices must be considered when assessing significance of relative relationships. For instance, in figure 1, due to its heavier connecting edge, the relationship of UCDProx to Attention Deficit Hyperactivity Disorder may appear more significant than that of UCDProx to Mood Disorder. However, the heavier edge is due to there being more observations involved in the first relationship than in the second. On the one hand, this is informative, while on the other, differences in proportion UCD-distal and UCD-proximal observations are approximately equal for both Attention Deficit and Mood disorders. This is also informative. Suggestions on representing proportions in addition to frequencies will be appreciated.

- Concept and Prescription nodes are constructed from distinct participant relationships. For instance, if participant A has multiple diagnoses of concept B, the B node includes A once only.

- Although absent in graphs appearing here, the app renders labels containing numbers of observations when nodes and edges are hovered over. Labels are modifiable and can contain anything that can be queried or computed regarding associated participant, diagnosis, and prescription nodes and relationships.

- Several graphs exhibit a confounding of node variable levels. For instance, figure 2 reveals a set of concepts that is associated strictly with males, while a disjoint set is associated strictly with females. It is possible that conditions are sex specific or, possibly, in the data, all participants exhibiting a given condition are of a single sex. Care must be taken when generalizing relationships involving confounded entities. In this example, it is not clear whether a relationship involving sex, concept, or both are indicated.

- Figure 4 shows one participant with Age of Onset below twelve days, but associated with Adult Attention Deficit Hyperactivity Disorder. It was verified that Neonatal Onset Age for this participant is less than 28 days. Is there a (clinically) meaningful explanation for this?

- visNetwork attempts to place nodes such that those with high mass (many observations) appear in a region along with connected nodes. Higher mass causes greater gravitational-like attraction. Once a graph is rendered, nodes can be repositioned using selectable methods that affect attraction (causing connected nodes to be "dragged" along). Attraction can also be disabled, so that nodes can be repositioned without affecting others. Each graph presented had nodes manually repositioned after initial rendering. Nodes can also be programmatically repositioned and it may be useful to implement

alternative algorithms, such as for collecting nodes by type (Concept nodes in one region, Participant nodes in another, Prescription nodes in a third region).

- Features are implemented to subset a graph by either highlighting a node's nearest neighborhood (figure 5) or by truncating the graph to the neighborhood of a selected node (figure 6). In a single-node subset, edge size indicates proportion of observations in each relationship (edge). Multiple node subsets are also possible, based on user specified node or edge filters.

- Figures 7 through 12 include nodes for prescriptions. Prescriptions are connected, by user selection, to either Concept nodes (figures 7 and 8) or Participant property nodes (figures 9 and 10). With three node categories and many prescriptions, graphs become busy, making the subset feature convenient for detailed exploration of relationships (figures 11 and 12).

- Note the increase in the number of prescription nodes when subsuming is requested (figure 8 compared to figure 7). This occurs because, for instance, both Phenylbutyrate (type "ingredient") and Glycerol Phenylbutyrate (type "precise ingredient") subsume Glycerol Phenylbutyrate 1100 MG/ML Oral Solution [Ravicti], causing two subsuming nodes to appear for this prescription in figure 8. A means of resolving single subsuming prescriptions is needed.

1 Graphs relating mental disorder SNOMEDCT concepts with participant features
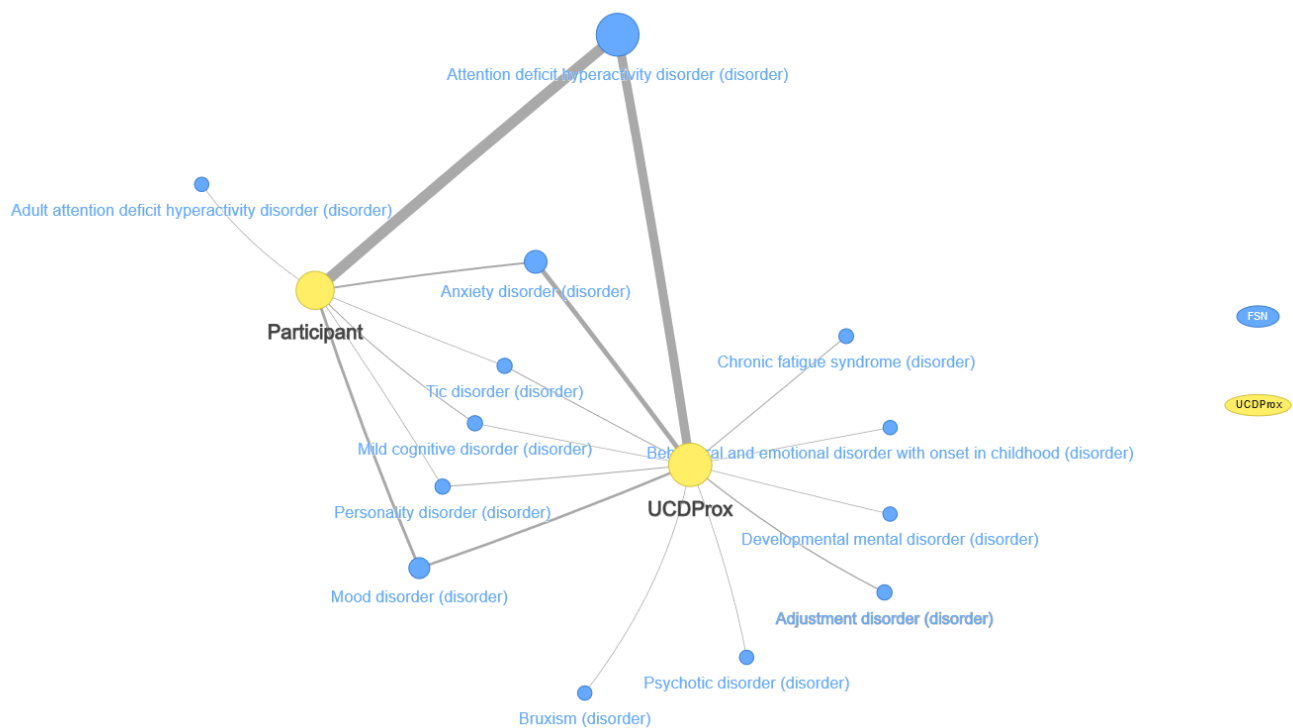

Figure 1: Graph with Mental Disorder concepts (blue, Fully Specified Name) and Participant UCD_Proximal state (yellow) as nodes.
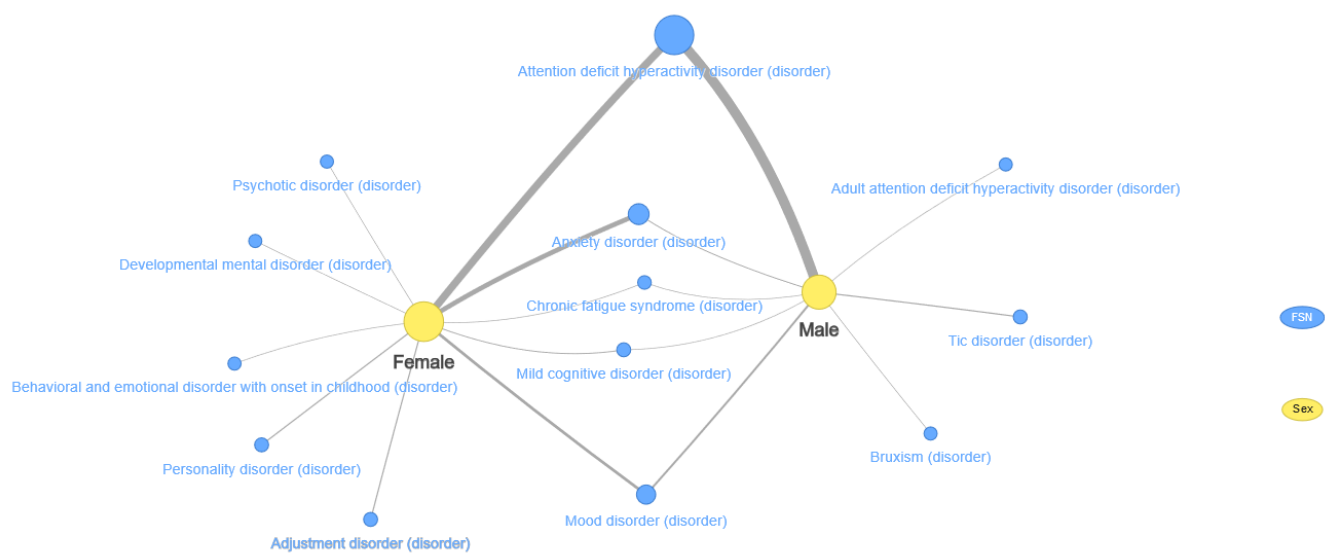

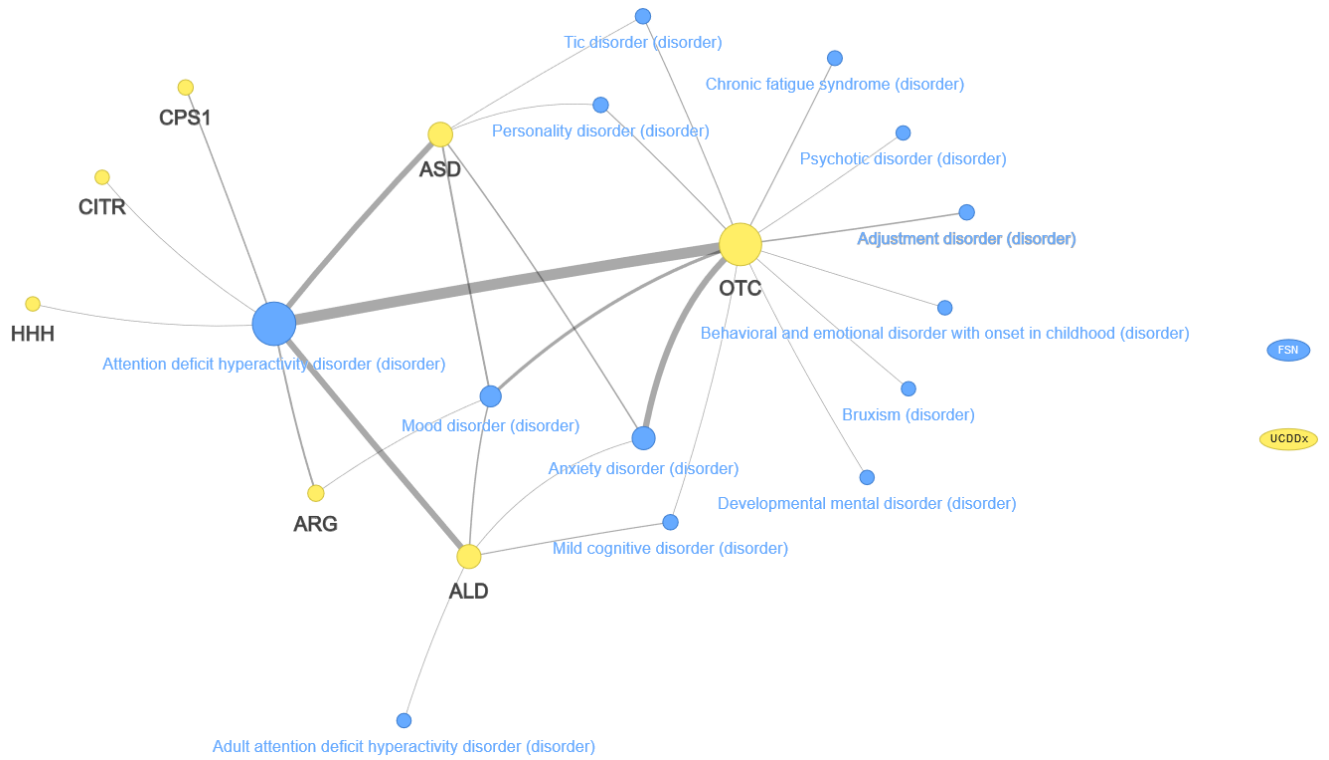Figure 2: Graph with Mental Disorder concepts (blue) and Participant Sex (yellow) as nodes.

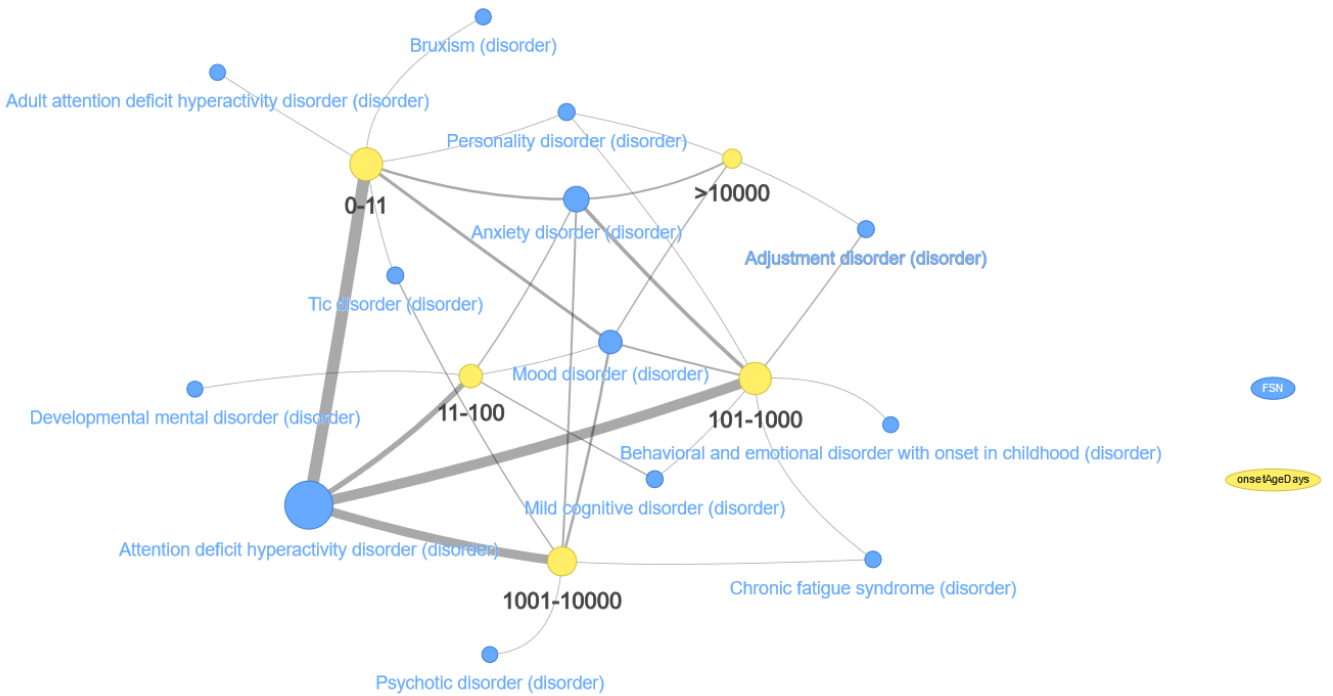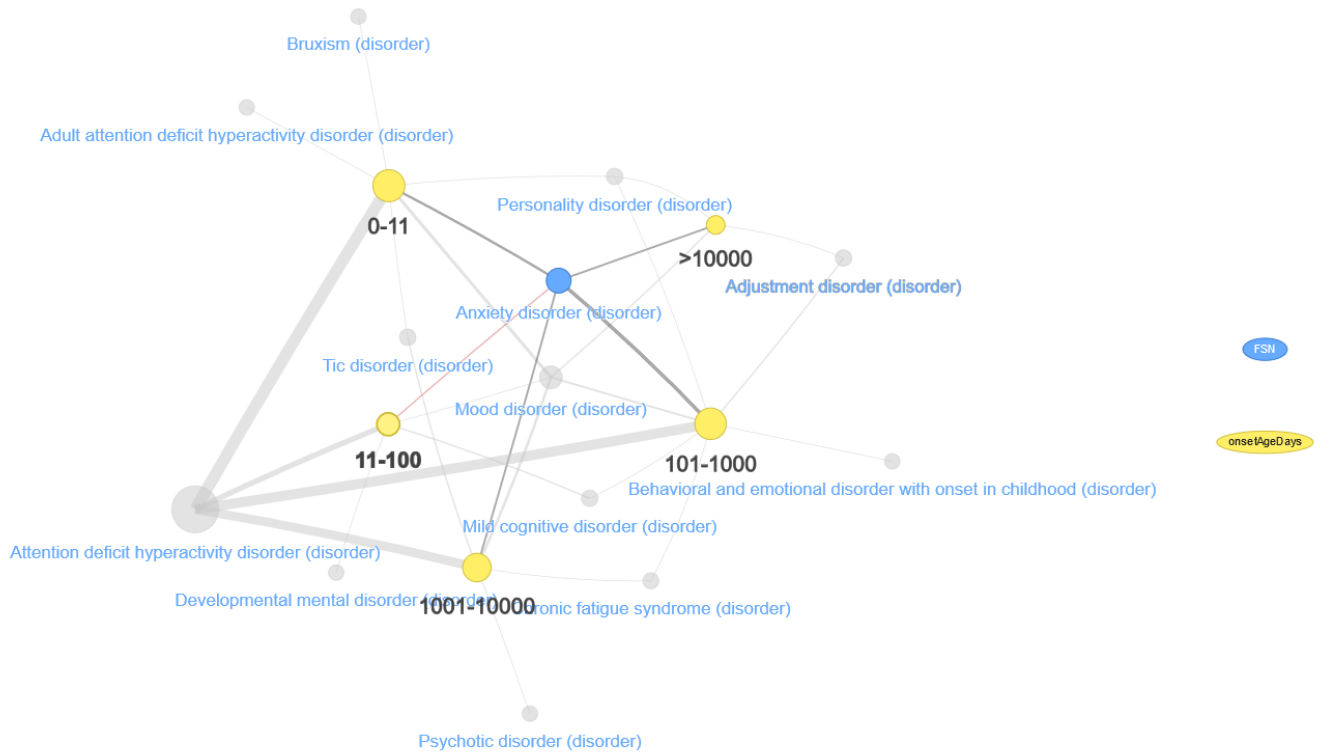Figure 3: Graph with Mental Disorder concepts (blue) and Participant UCDDx (yellow) as nodes.

Figure 4: Graph with Mental Disorder concepts (blue) and Participant Age at Onset (in days, yellow) as nodes.

Figure 5: Graph with Mental Disorder concepts (blue) and Participant Age at Onset (yellow) as nodes. Anxiety Disorder subset with neighborhood highlighted.
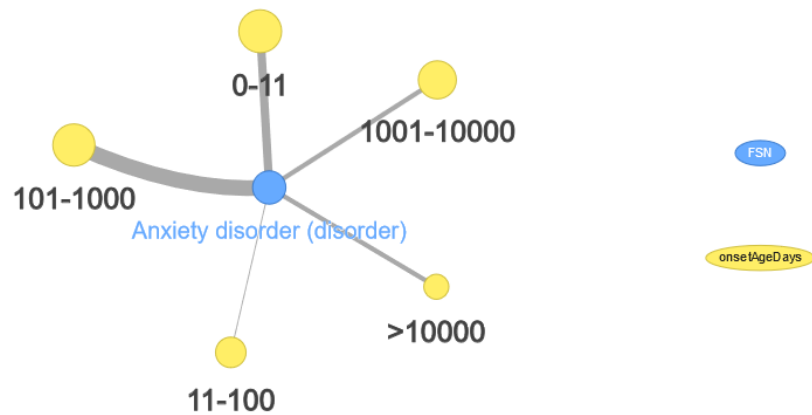

Figure 6: Graph with Mental Disorder concepts (blue) and Participant Age at Onset (yellow) as nodes. Graph truncated to Anxiety Disorder neighborhood.

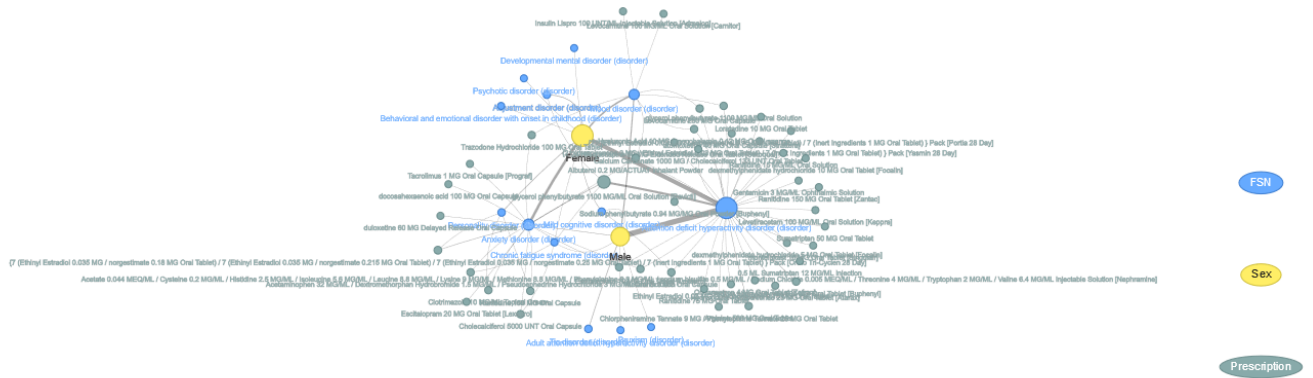# 2 Graphs relating mental disorder concepts with participant features and prescriptions

Figure 7: Graph with Mental Disorder concepts (blue), Participant Sex (yellow), and Prescription (green) as nodes. Complete prescription text. Prescriptions joined to Concept.
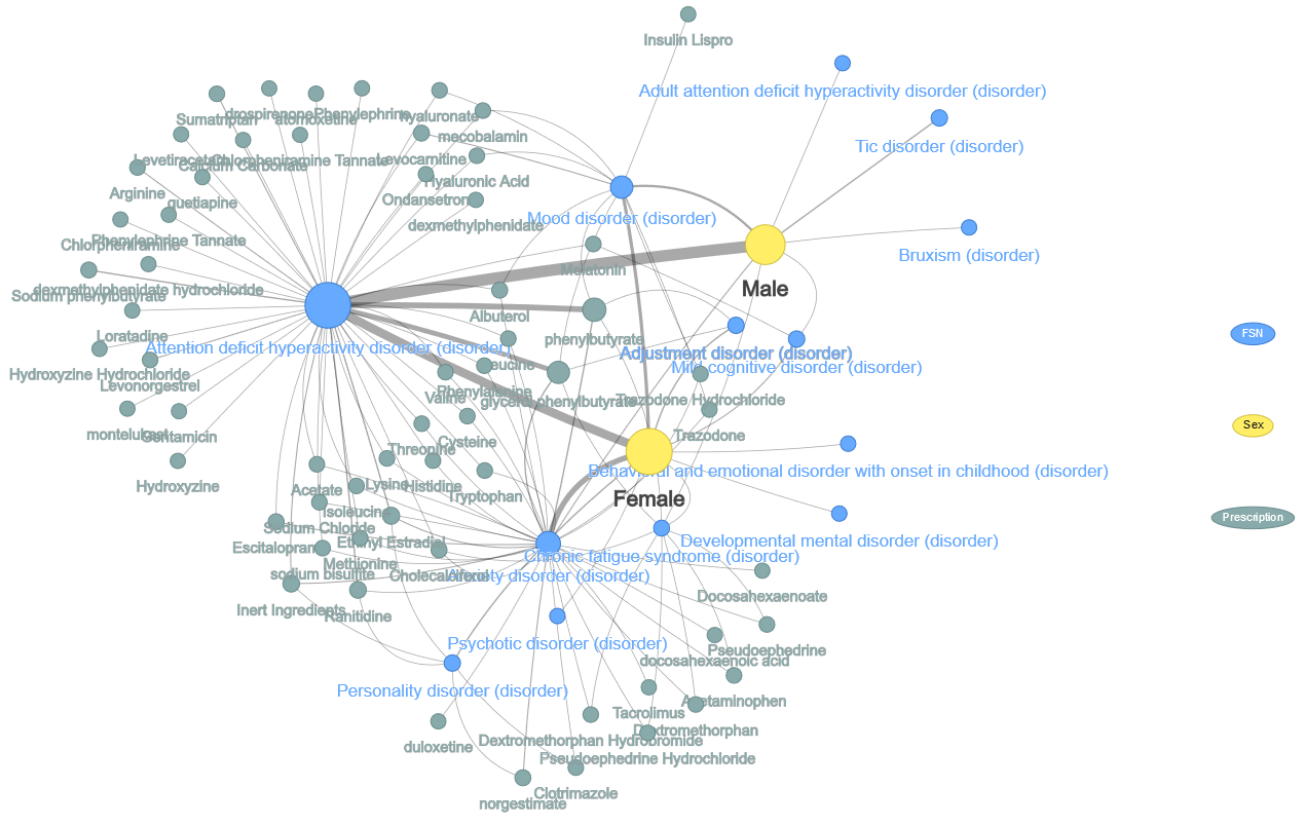

Figure 8: Graph with Mental Disorder concepts (blue), Participant Sex (yellow), and Prescription (green) as nodes. Subsuming Prescription text replaces complete text. Prescriptions connected to Concept.
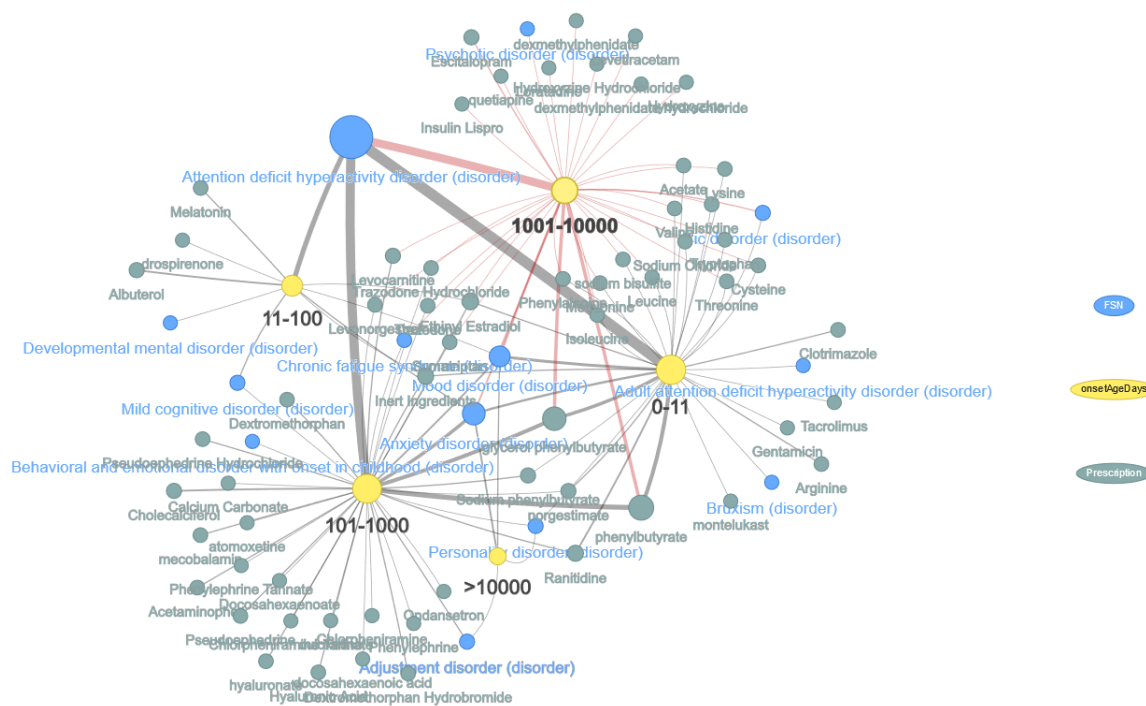
Figure 9: Graph with Mental Disorder concepts (blue), Participant Age at Onset (yellow), and Prescription (green) as nodes. Subsumed prescription text. Prescriptions joined to Age.
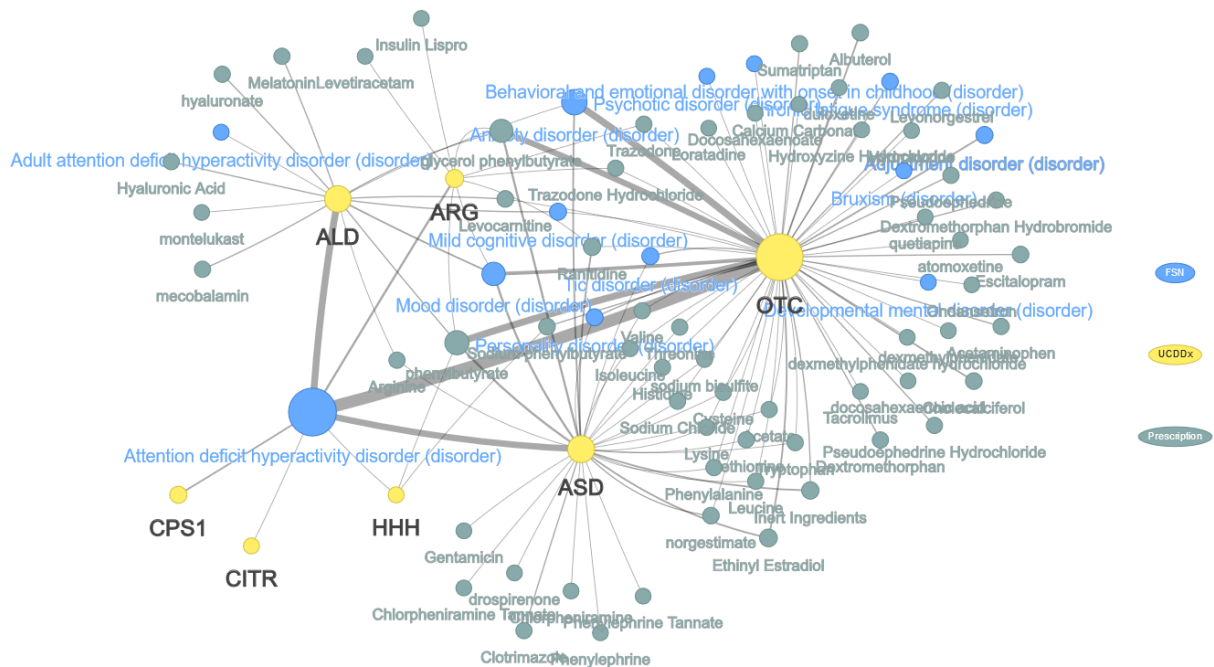

Figure 10: Graph with Mental Disorder concepts (blue), Participant Age at Onset (yellow), and Prescription (green) as nodes. Subsumed prescription text. Prescriptions joined to UCDDx.
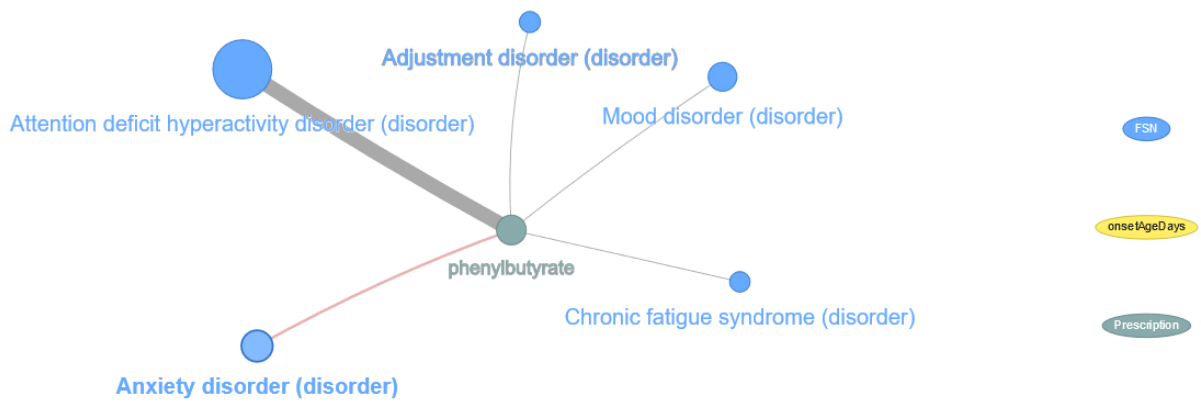
Figure 11: Truncated graph subset with Mental Disorder concepts (blue) and selected subsumed Prescription (Phenylbutyrate, green) as node.
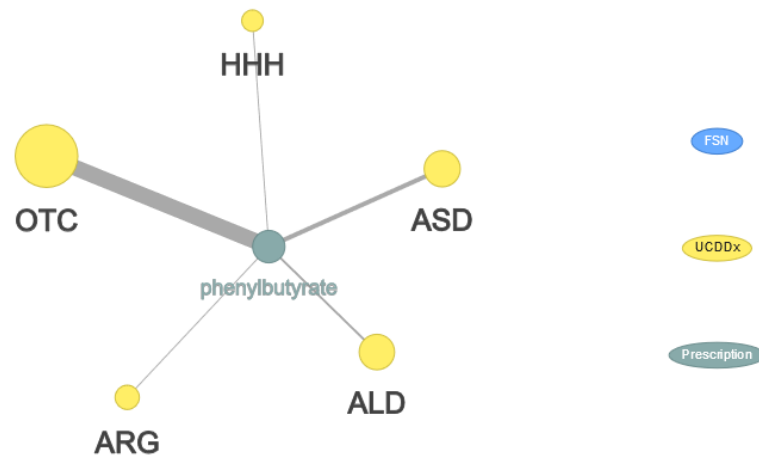

Figure 12: Truncated graph subset with UCDDx (yellow) and selected subsumed Prescription (Phenylbutyrate, green) as node.