HALA ALBAHLOUL

Diseases Genes

The association between diseases and genes

How are diseases related?

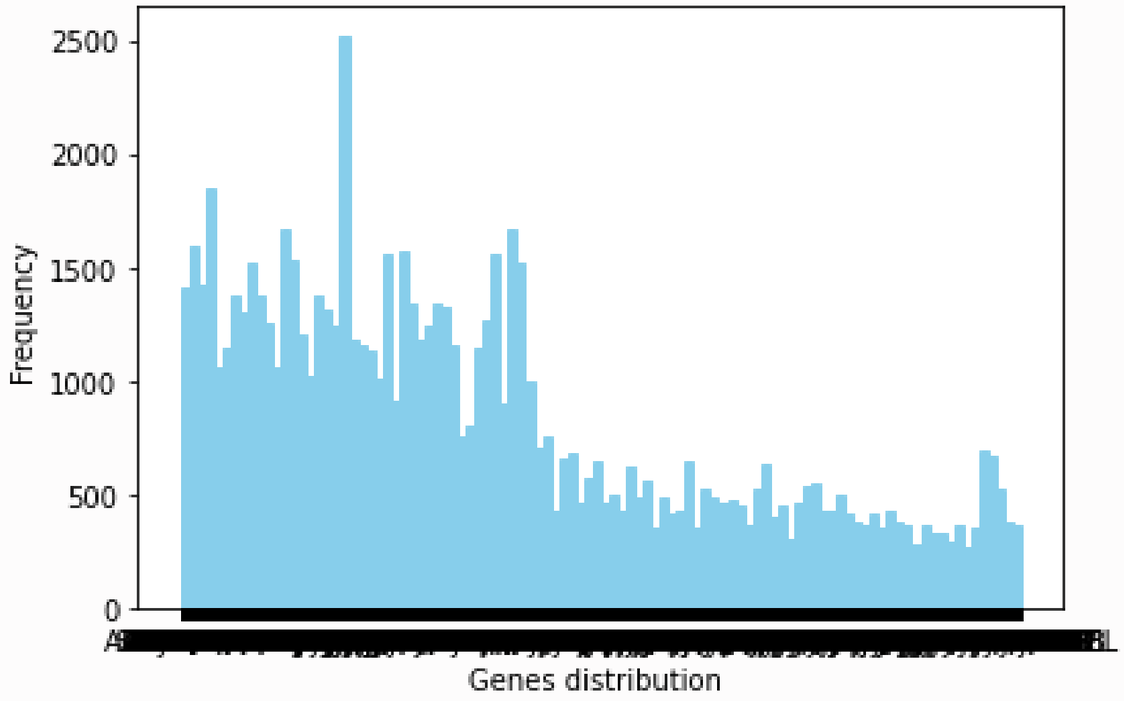
What diseases are you likely to get?

Based on Genes in common

DisGeNET

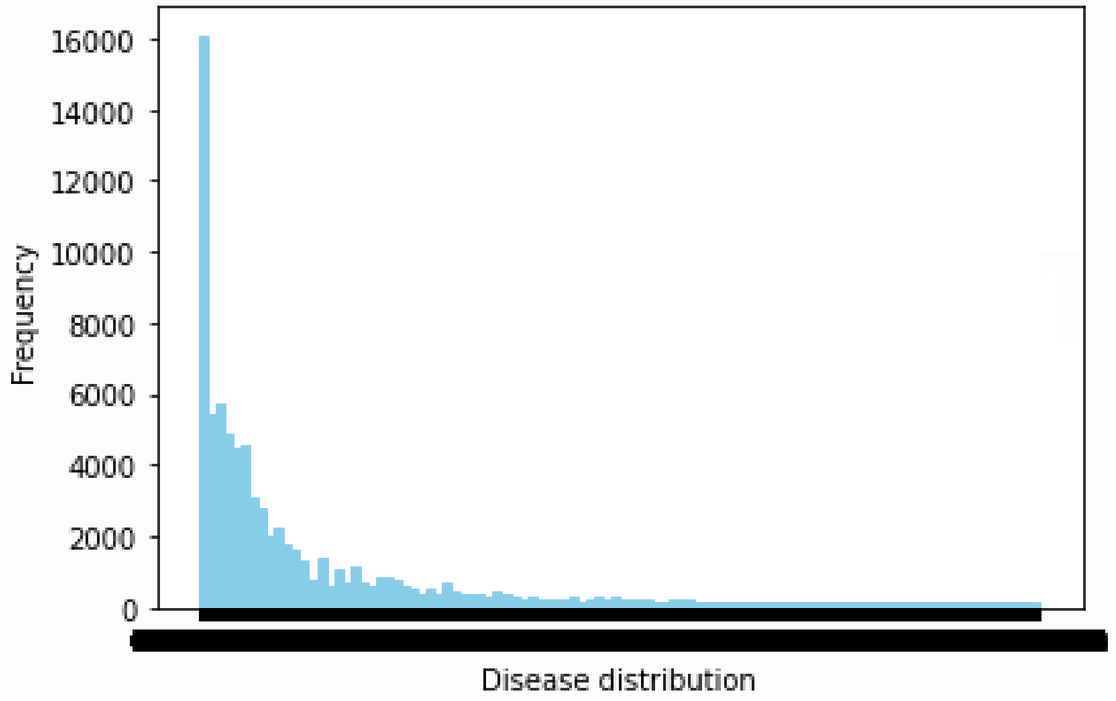
9411

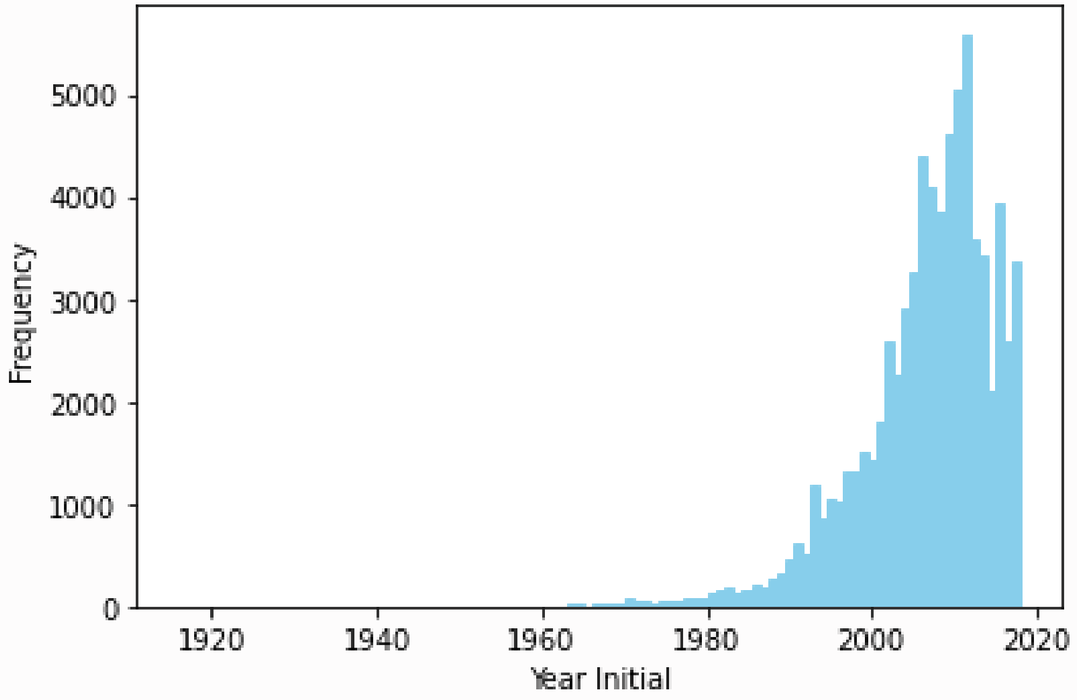
Genes

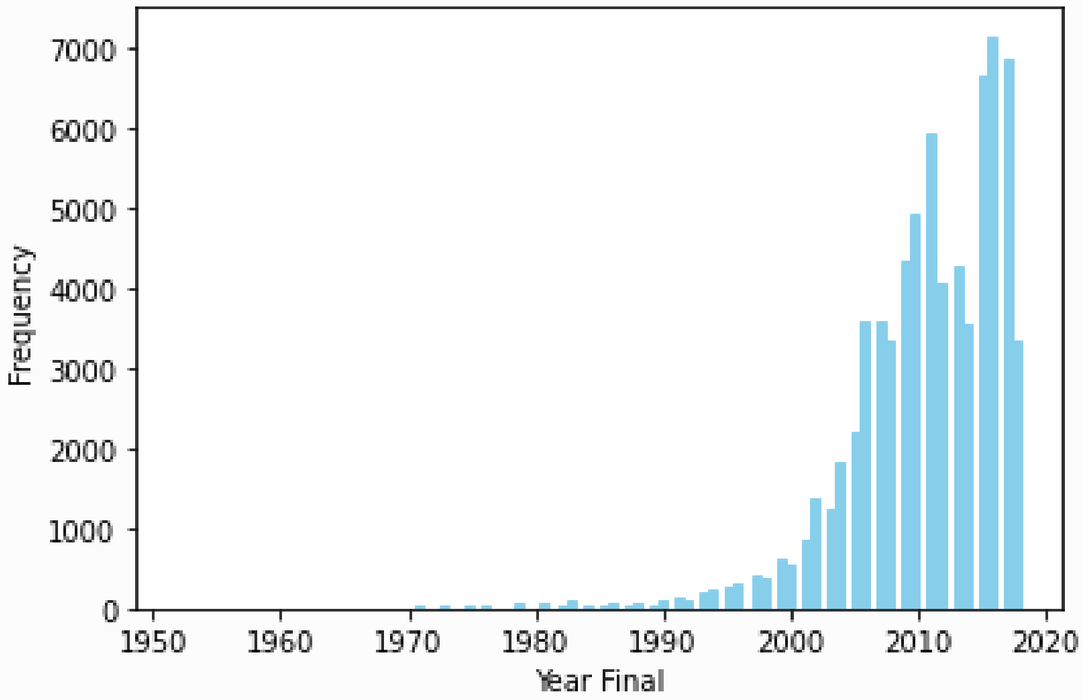


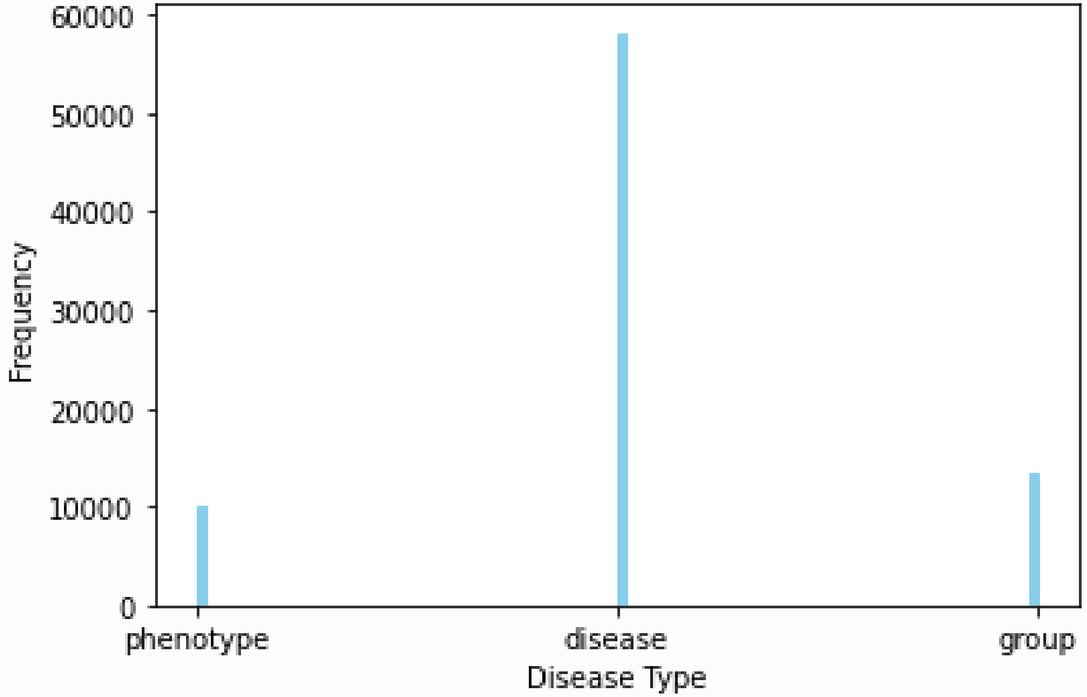
10370

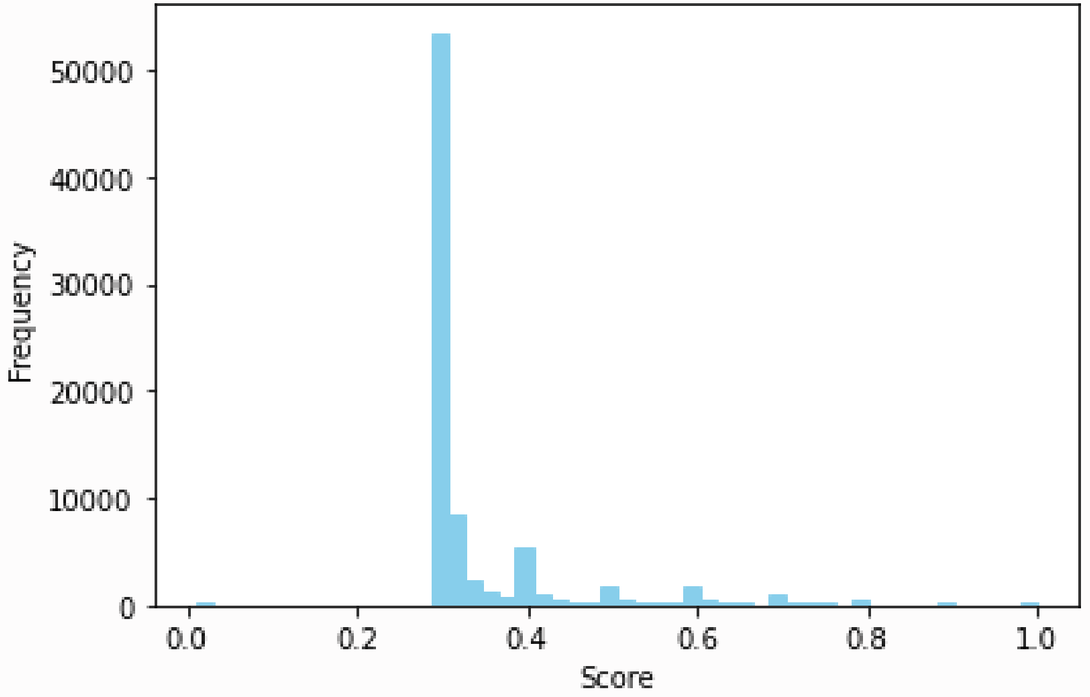
Diseases




Frequency
0
1000
2000
3000
4000
5000
1920
1940
1960
1980
2000
2020
Year Initial


Frequency
0
1000
2000
3000
4000
5000
6000
7000
1950
1960
1970
1980
1990
2000
2010
2020
Year Final


Frequency
0
10000
20000
30000
40000
50000
60000
phenotype
disease
group
Disease Type


Frequency
0
10000
20000
30000
40000
50000
0.0
0.2
0.4
0.6
0.8
1.0
Score

Process

How data is selected?

Score > 0.5
Recent years
Early years

Hypergeomitrec Filter

Filter the low probabilitis

Work with Gephi

Filter and draw

A Lucrative Industry

Diseases Network

Nodes: 177

Edges: 311

Node Size: betweenness centrality

Node Color: Disease Type

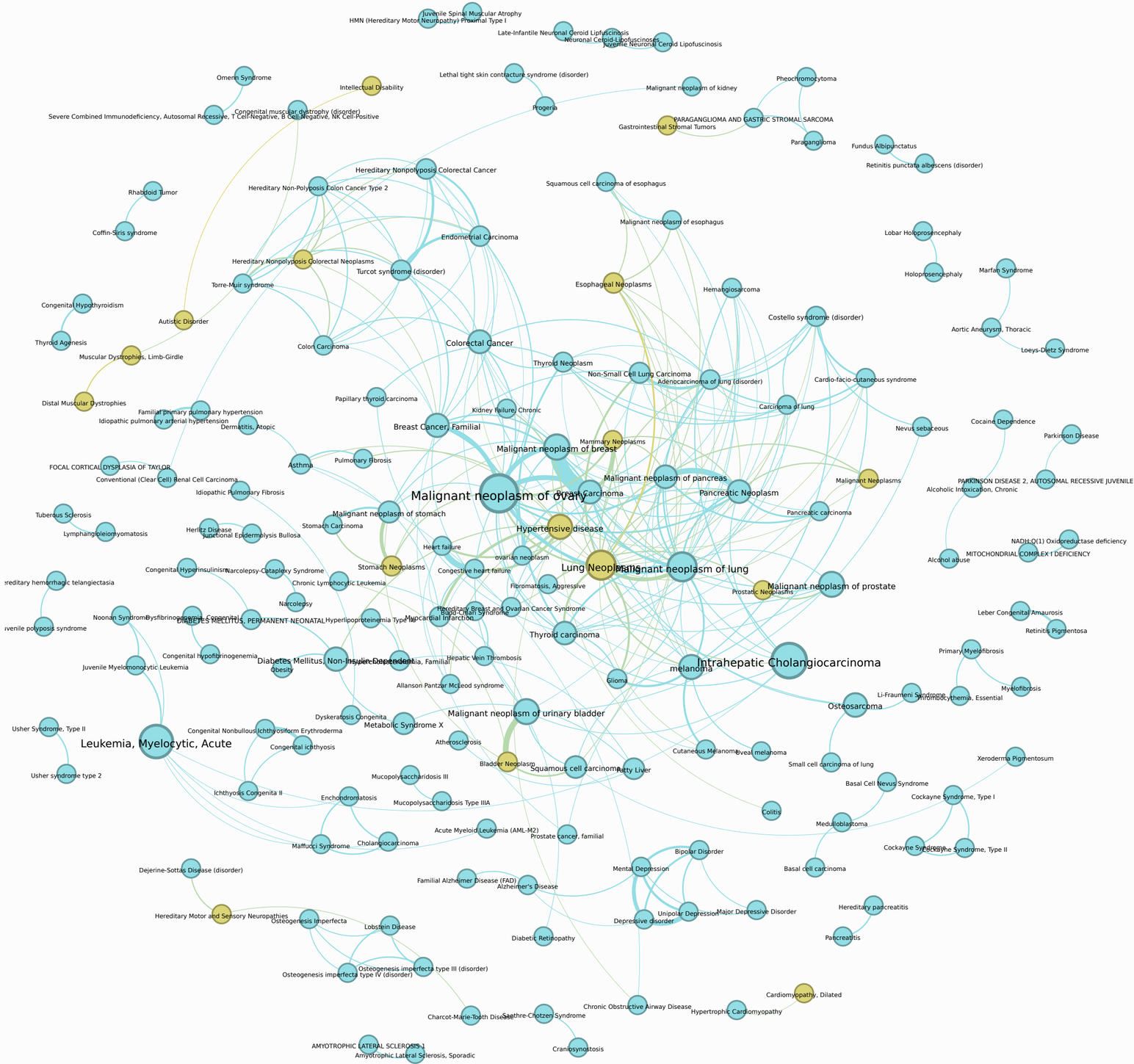
Filter:

Years>2014

Score> 0.5

Edge Wieght

Node Degree



Nodes: 177

Edges: 311

Node Size: betweenness centrality

Node Color: Disease Type

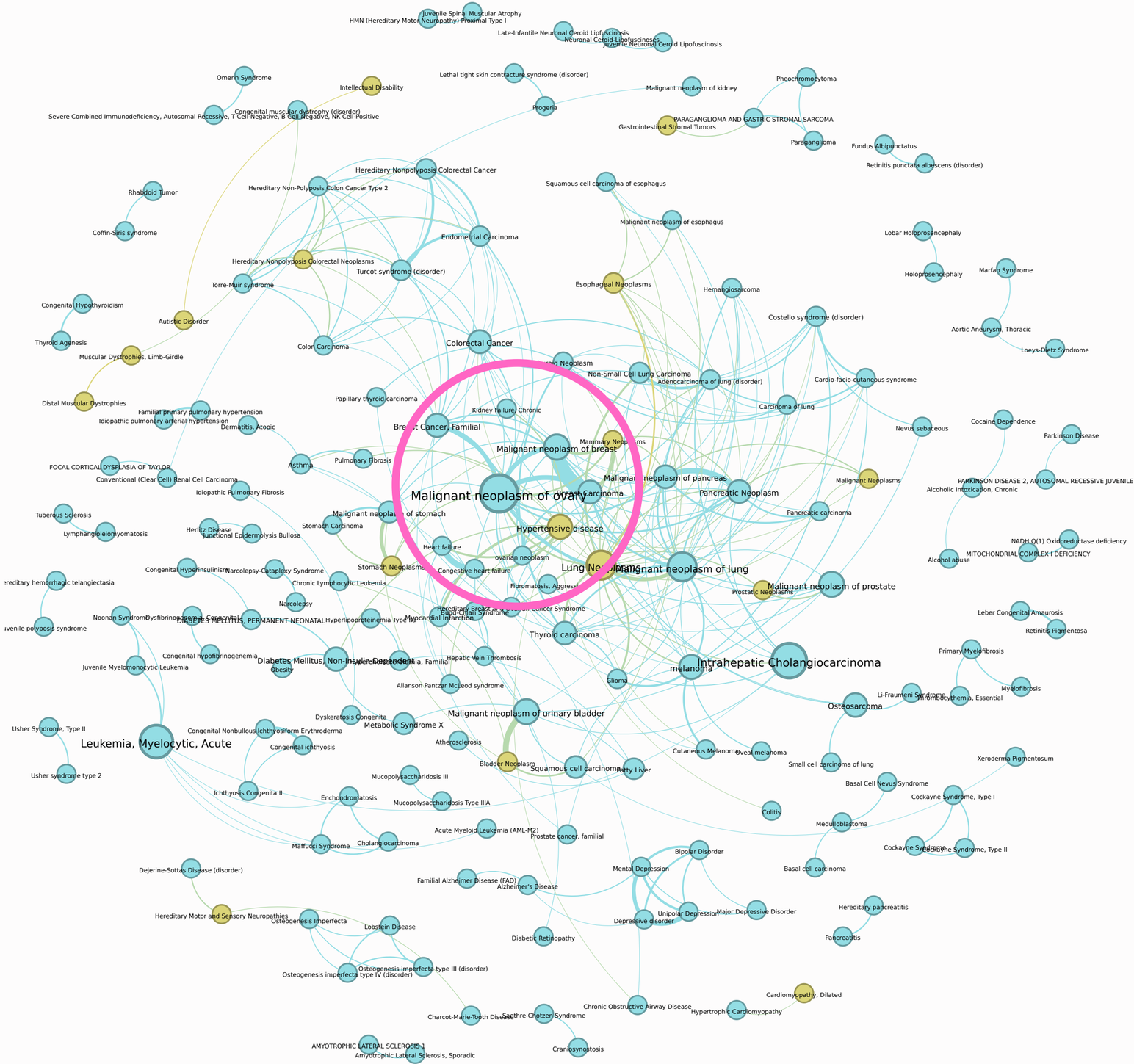
Filter:

Years>2014

Score> 0.5

Edge Wieght

Node Degree

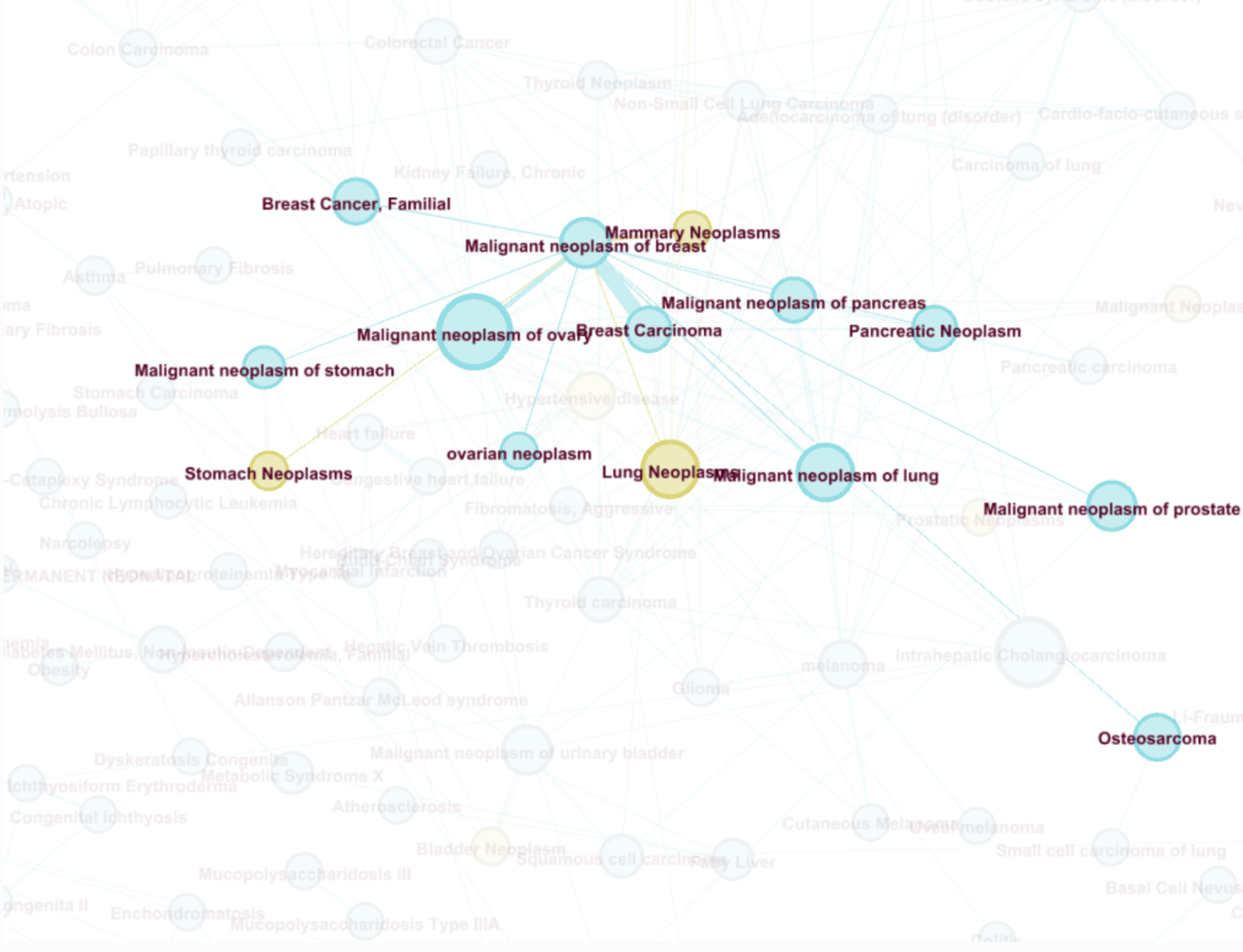


Malignant neoplasm of breast: 1983-2018

1025 Gene

2008-2018 26 Gene

1998-2008 16 Gene



Nodes: 79

Edges: 311

Node Size: betweenness centrality

Node Color: Disease Type

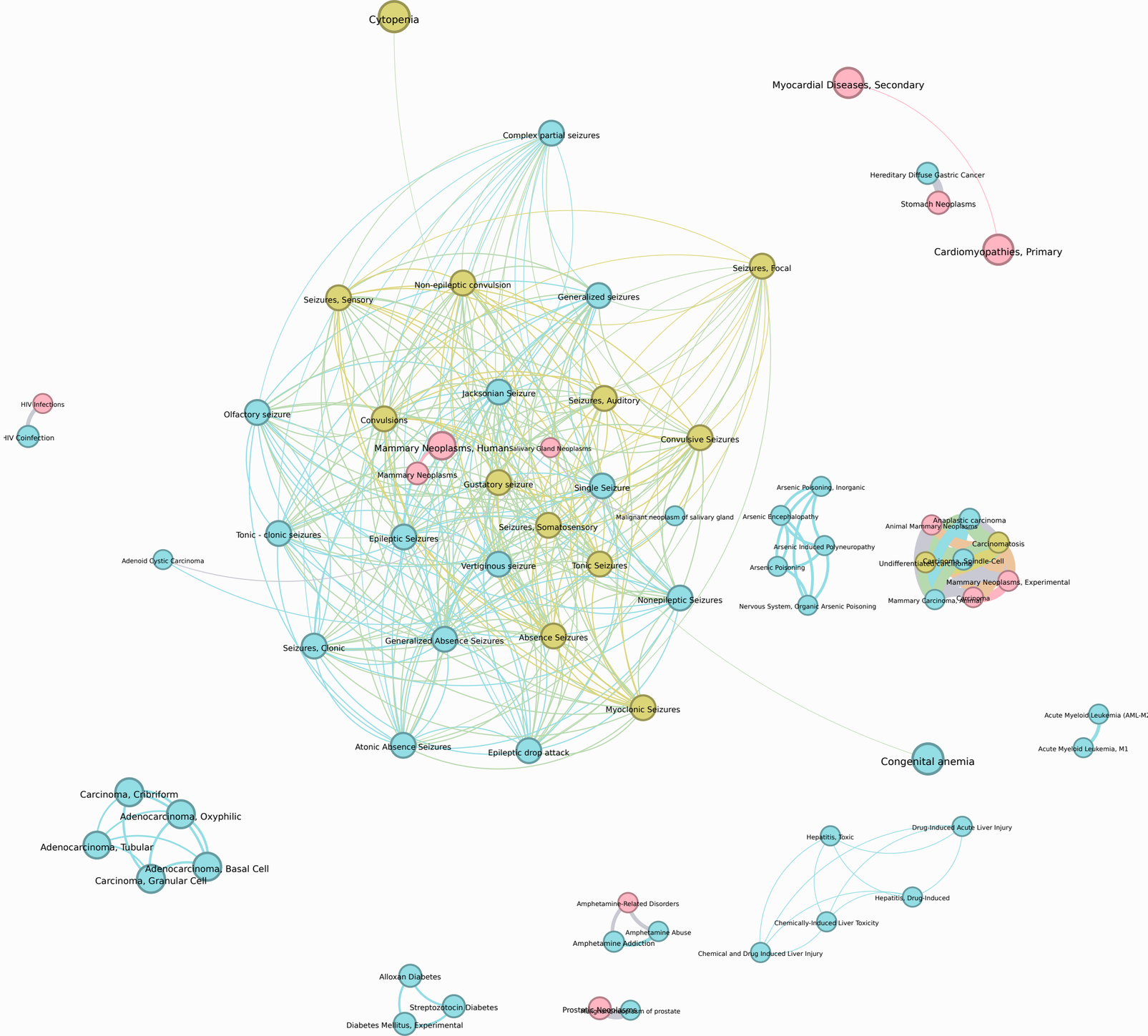
Filter:

Years: 1998-2008

Score > 0.5

Edge Wieght

Node Degree



Nodes: 79

Edges: 311

Node Size: betweenness centrality

Node Color: Disease Type

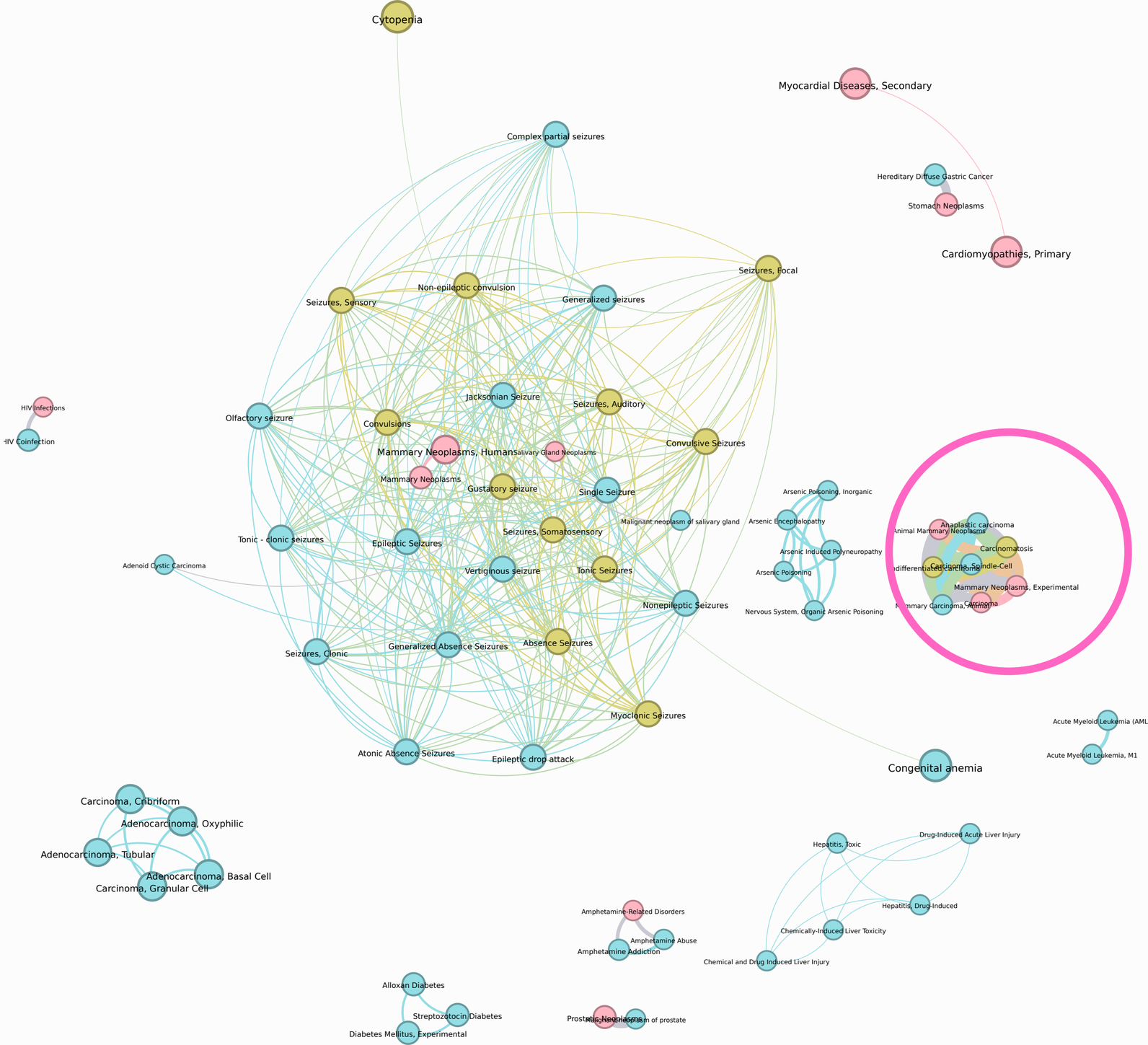
Filter:

Years: 1998-2008

Score > 0.5

Edge Wieght

Node Degree



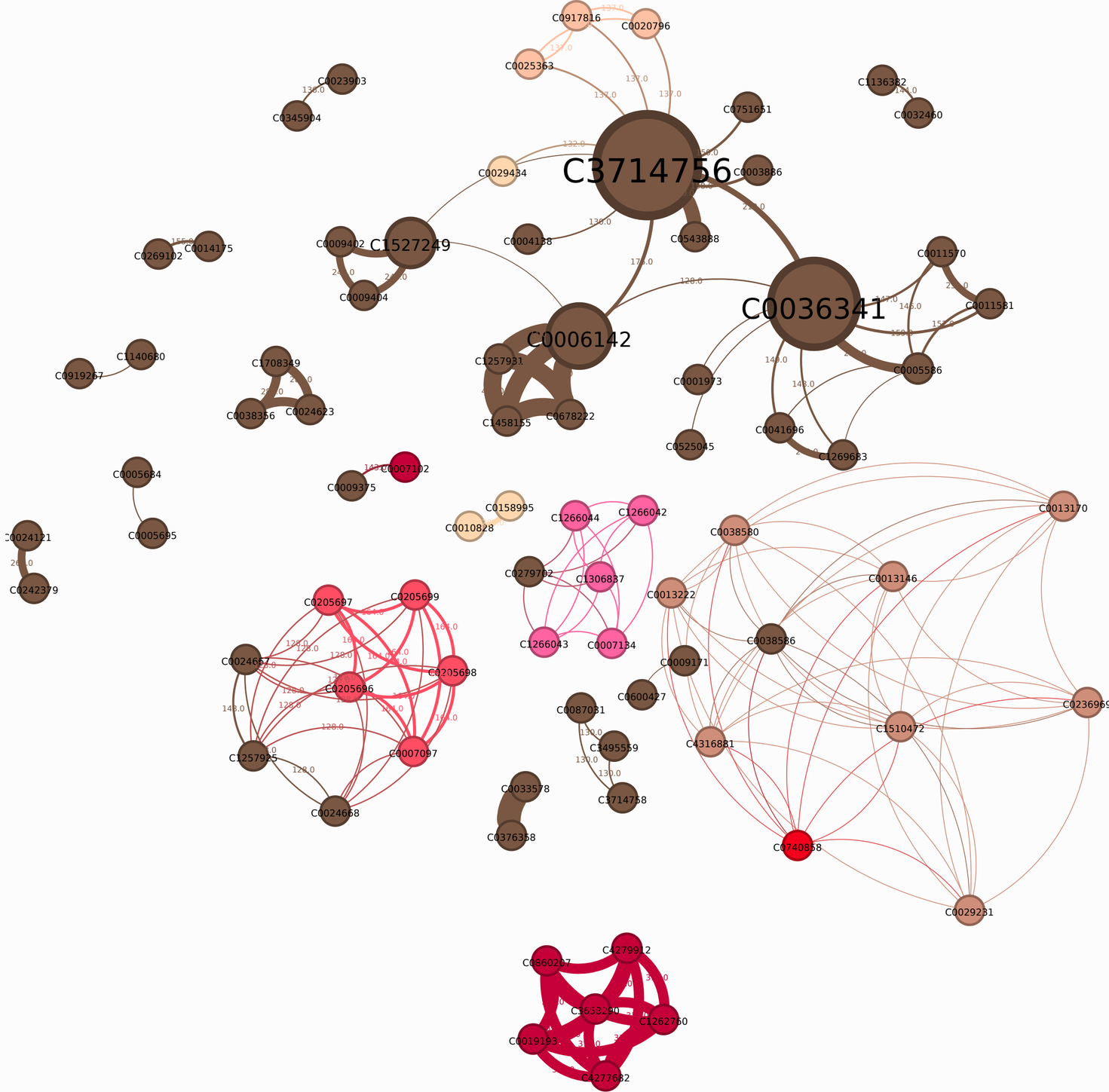
Filter:

Edge Wieght

Node Degree

Node Size: betweenness centrality

Node Color: Degree



Github Repository

https://github.com/halaalbahloul/DisgenetDataAnalysis

Thank you for your attension!

How is the score for a gene-disease pair, or variant-disease pair computed?

The DisGeNET score for GDAs takes into account the number and type of sources (level of curation, organisms), and the number of publications supporting the association, while the score for the VDAs takes into account sources, and number of papers. The scores range from 0 to 1.