Determining the Most Influencing Medical Conditions in MEDPRAT's SIN Directed Graph

Crew Health and Performance
Probabilistic Risk Assessment Project
(CHP-PRA)

Mona Matar, Ph.D.[1]
Hunter Rehm[2]
Lauren McIntyre[1]

1 NASA Glenn Research Center
2 Universities Space Research Association

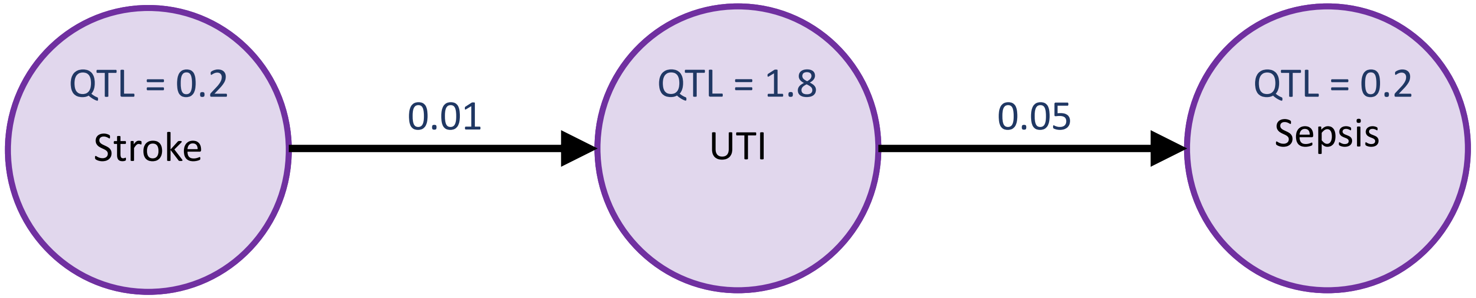
2023 HRP IWS - Feb 8, 2023
Galveston, TX

Background and Introduction

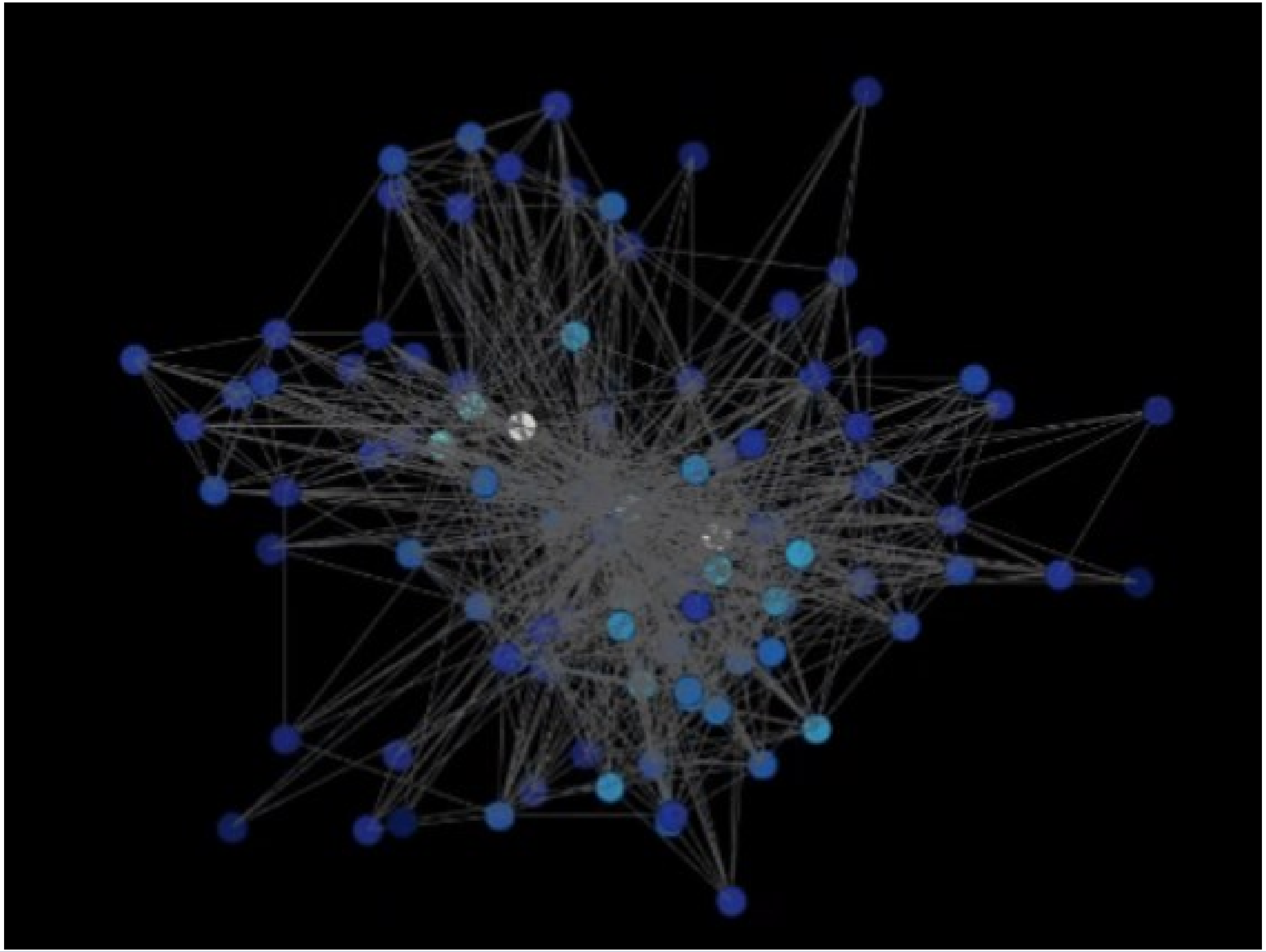
SIN Graph

The Susceptibility Inference Network (SIN)

- Part of MEDPRAT, an event driven, time dependent, stochastic tool
- Subject matter expert informed prototype
- Mathematical data structure to quantify relationships between medical conditions:
  - nodes: medical conditions
  - edges: directed, progression from one condition to the next
  - node weights: a statistic to measure the stand-alone risk, typically Quality Time Loss (QTL) or Loss Of Crew Life (LOCL)
  - edge weights: probability of progression to the next condition

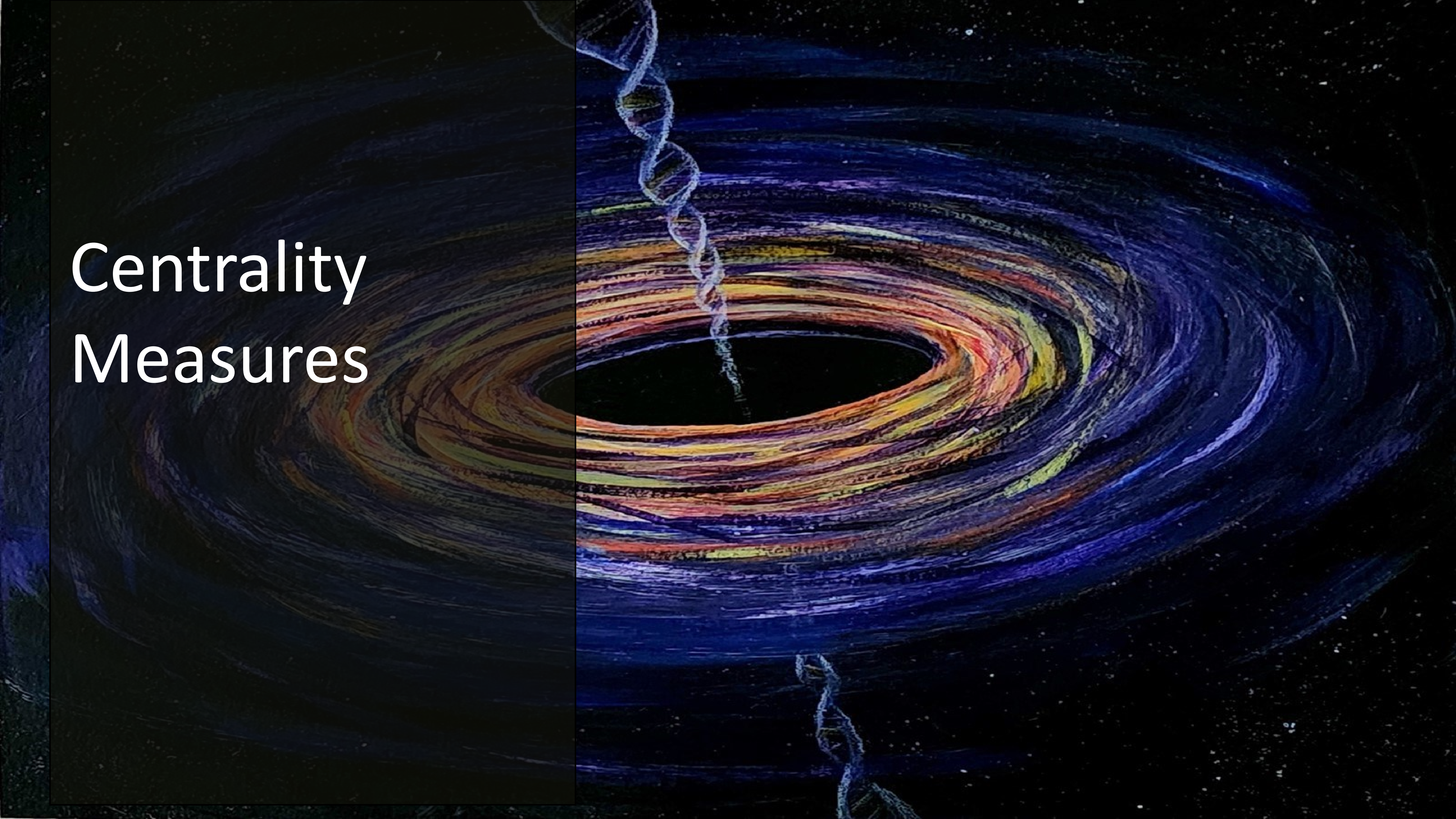


SIN Graph



- 99 medical conditions
- 1078 direct connections
- Anxiety, Atrial Fibrillation/ Flutter, Medication Overdose/Adverse Reaction responsible for 273 of the connections

Centrality Measures



Centrality Measures

- **<u>Our goal 1:</u>** compare how much each medical condition can progress through the network

- Measure of importance of a node
  - Local measures: degree centrality
  - Global measures: betweenness, eigenvector, Katz, PageRank Centrality

- We use: Katz Centrality

Graph Theory

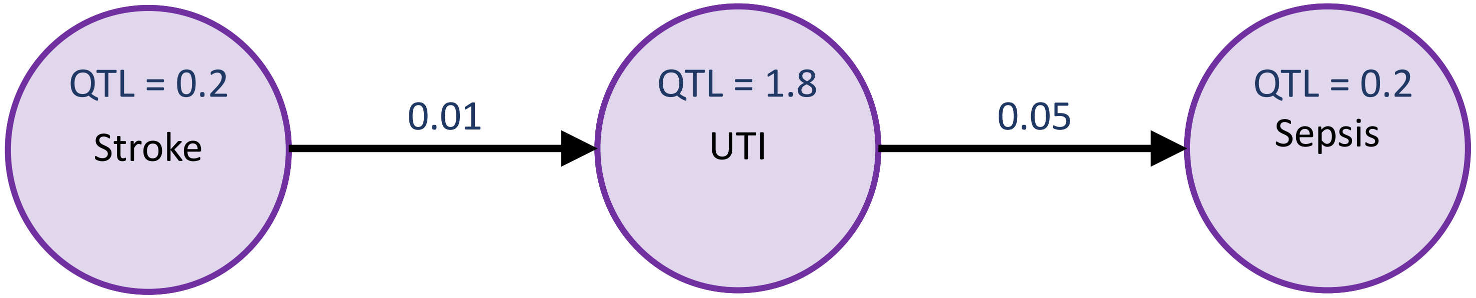
SIN Adjacency Matrix

- SIN adjacency matrix

$$A = \begin{array}{r} \\ \text{Stroke} \\ \text{UTI} \\ \text{Sepsis} \\ \ldots \end{array} \begin{array}{c} \begin{array}{cccc} \text{Stroke} & \text{UTI} & \text{Sepsis} & \ldots \end{array} \\ \begin{bmatrix} 0 & 0.01 & 0 & \ldots \\ 0 & 0 & 0.05 & \ldots \\ 0 & 0 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} \end{array}$$

- Stand-alone risks

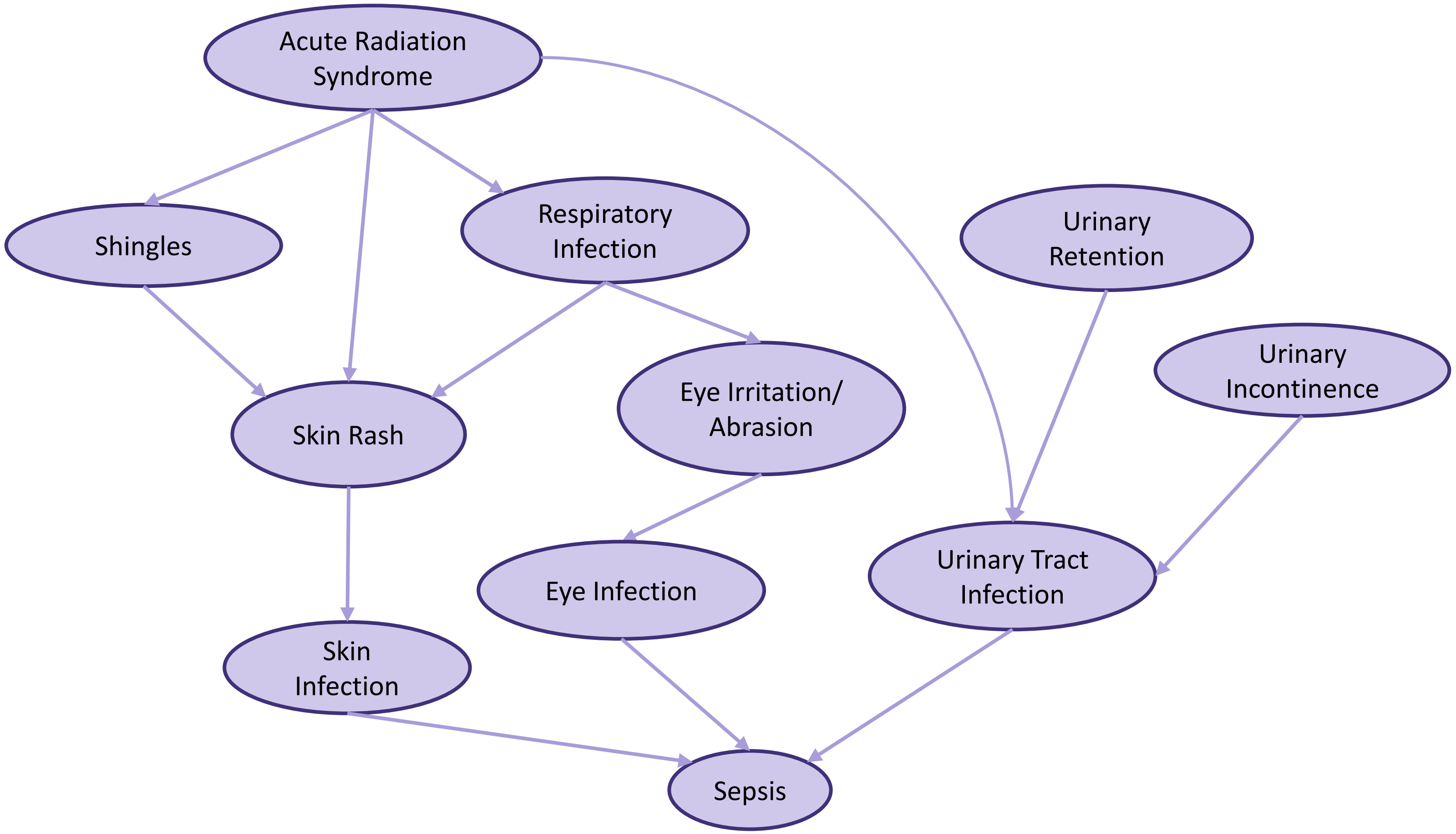
$$W = \begin{array}{r} \text{Stroke} \\ \text{UTI} \\ \text{Sepsis} \\ \ldots \end{array} \begin{bmatrix} 0.2 \\ 1.8 \\ 0.2 \\ \ldots \end{bmatrix}$$



Walk in SIN Network

- Walk of length k through k edges

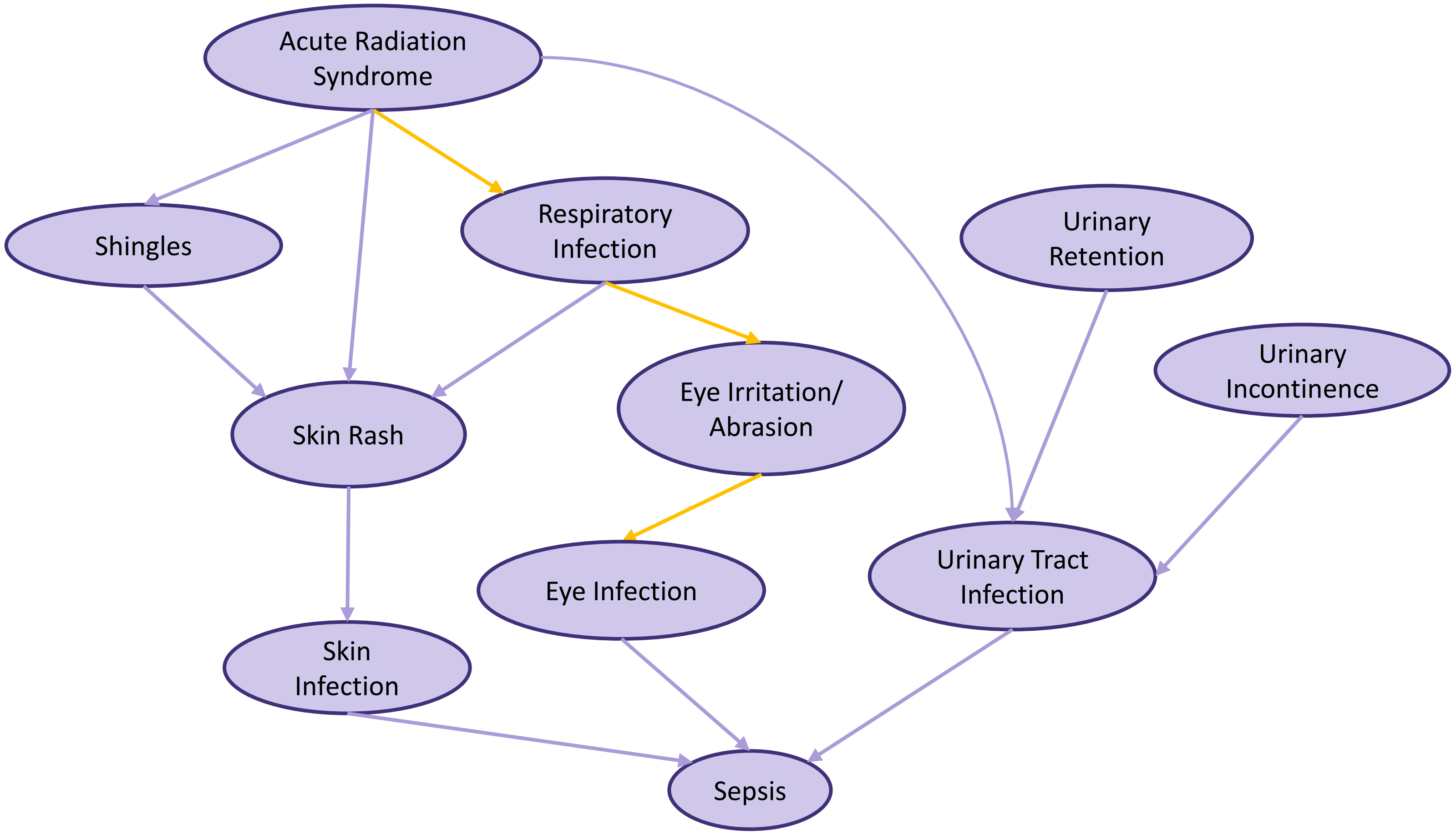
- Weight of walk
  = product of edge weights



Walk in SIN Network

Starting at Acute Radiation Syndrome

- Walk of length 3
- Weight of walk
  = product of edge weights

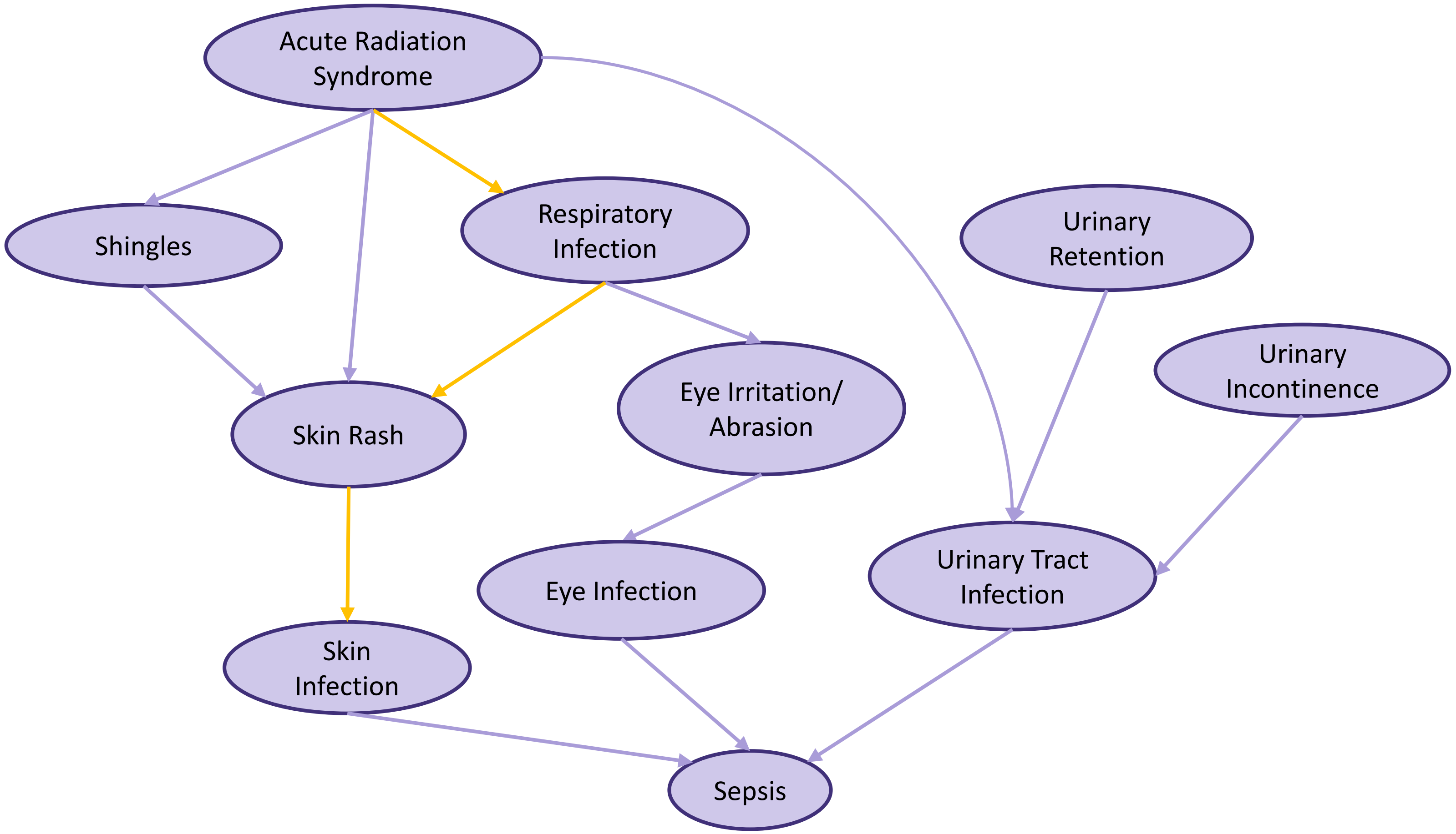


Walk in SIN Network

Starting at Acute Radiation Syndrome

- Walk of length 3

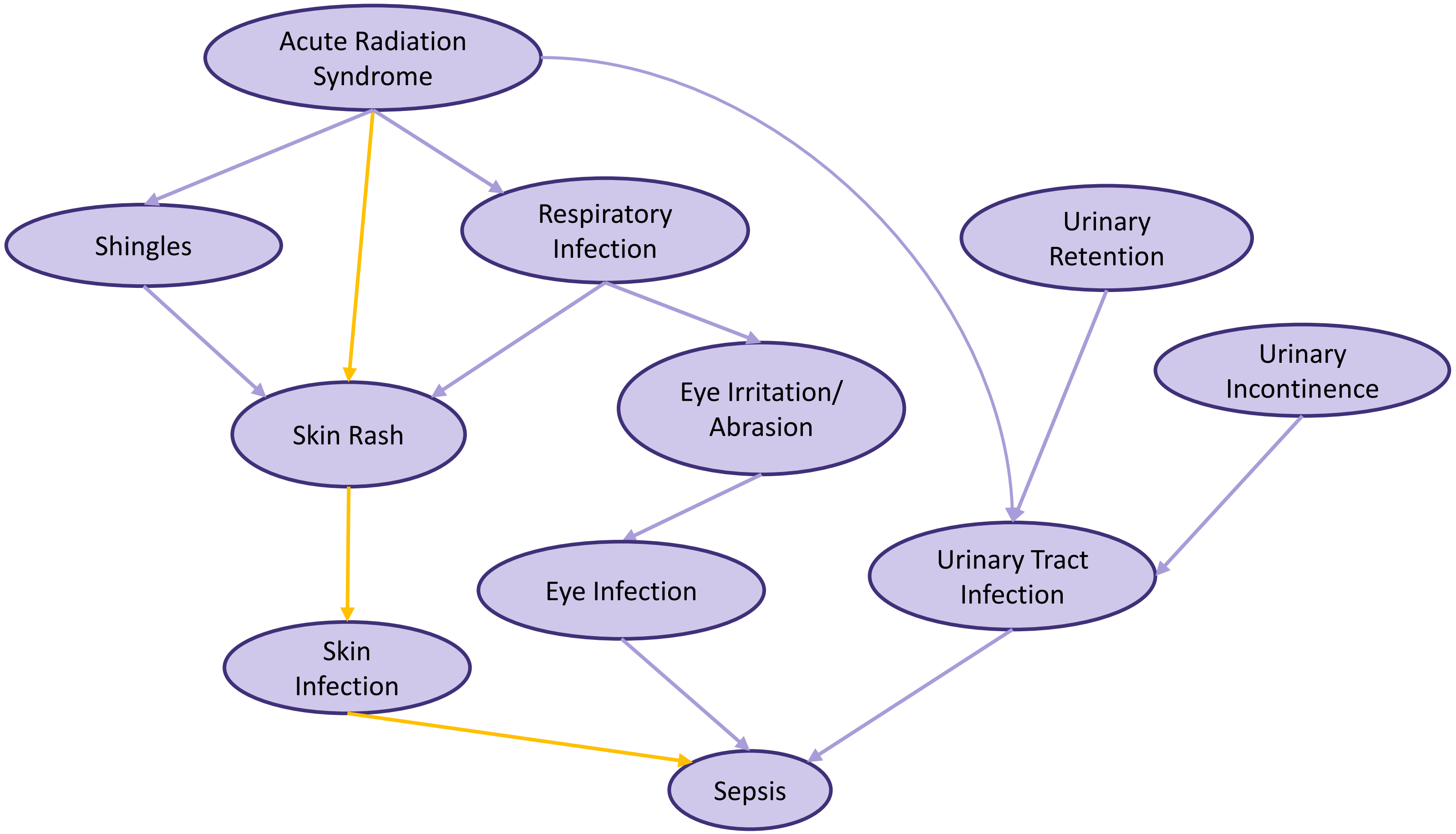
- Weight of walk
  = product of edge weights



Walk in SIN Network

Starting at Acute Radiation Syndrome

- Walk of length 3
- Weight of walk
  = product of edge weights

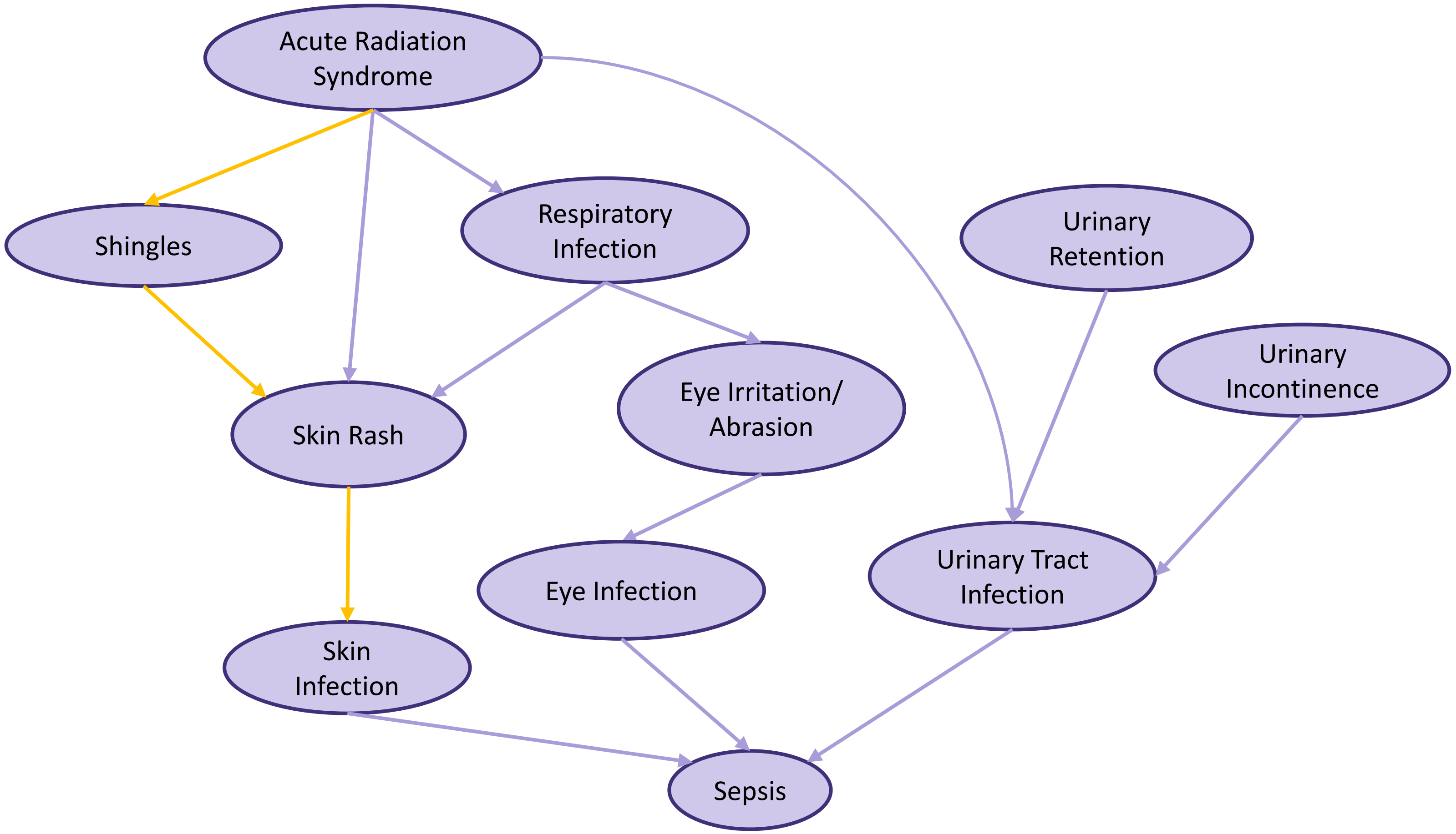


Walk in SIN Network

Starting at Acute Radiation Syndrome

- Walk of length 3

- Weight of walk
  = product of edge weights

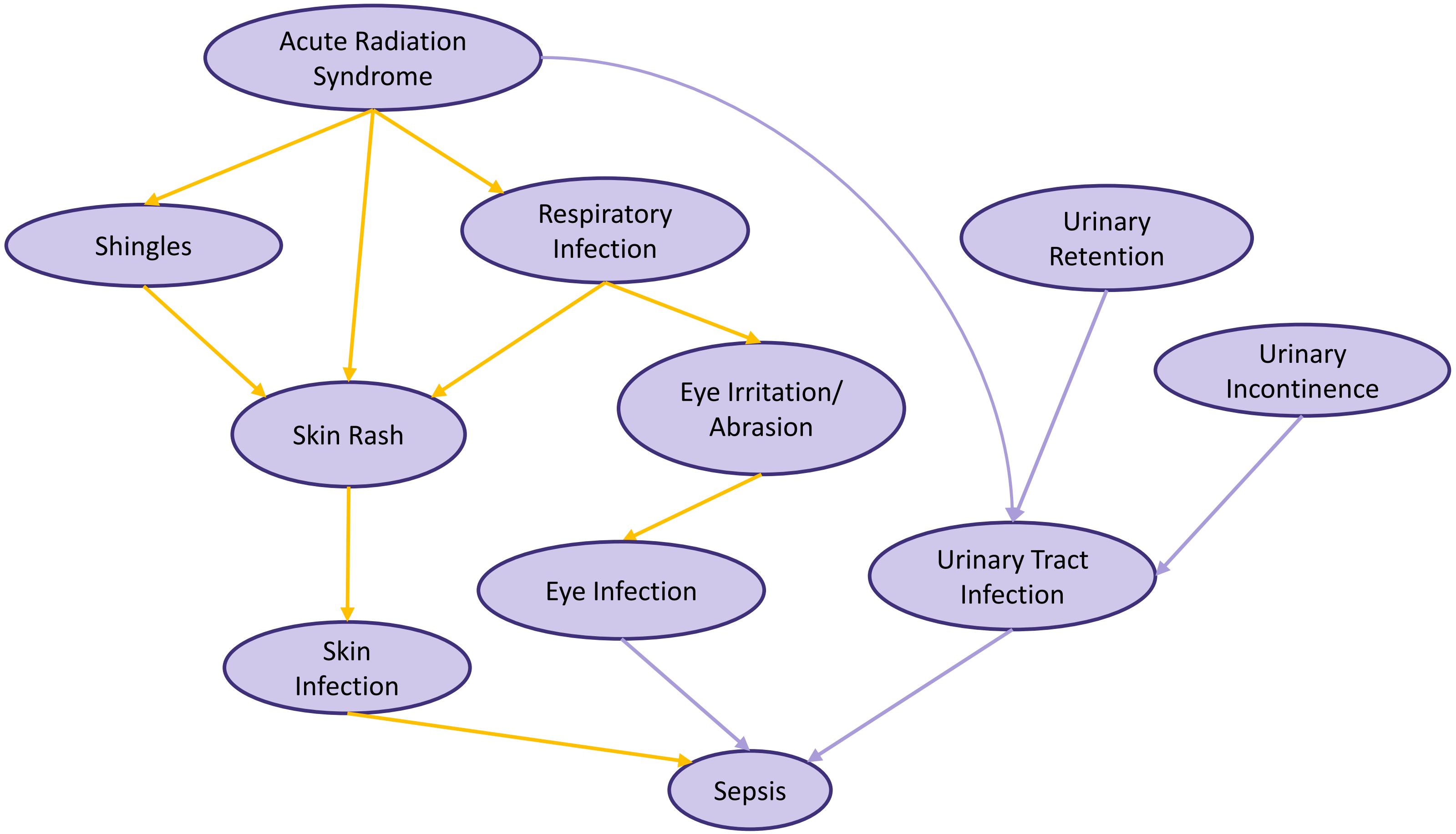


Walk in SIN Network

Starting at Acute Radiation Syndrome

- All walks of length 3
- Weighted number of walks of length 3

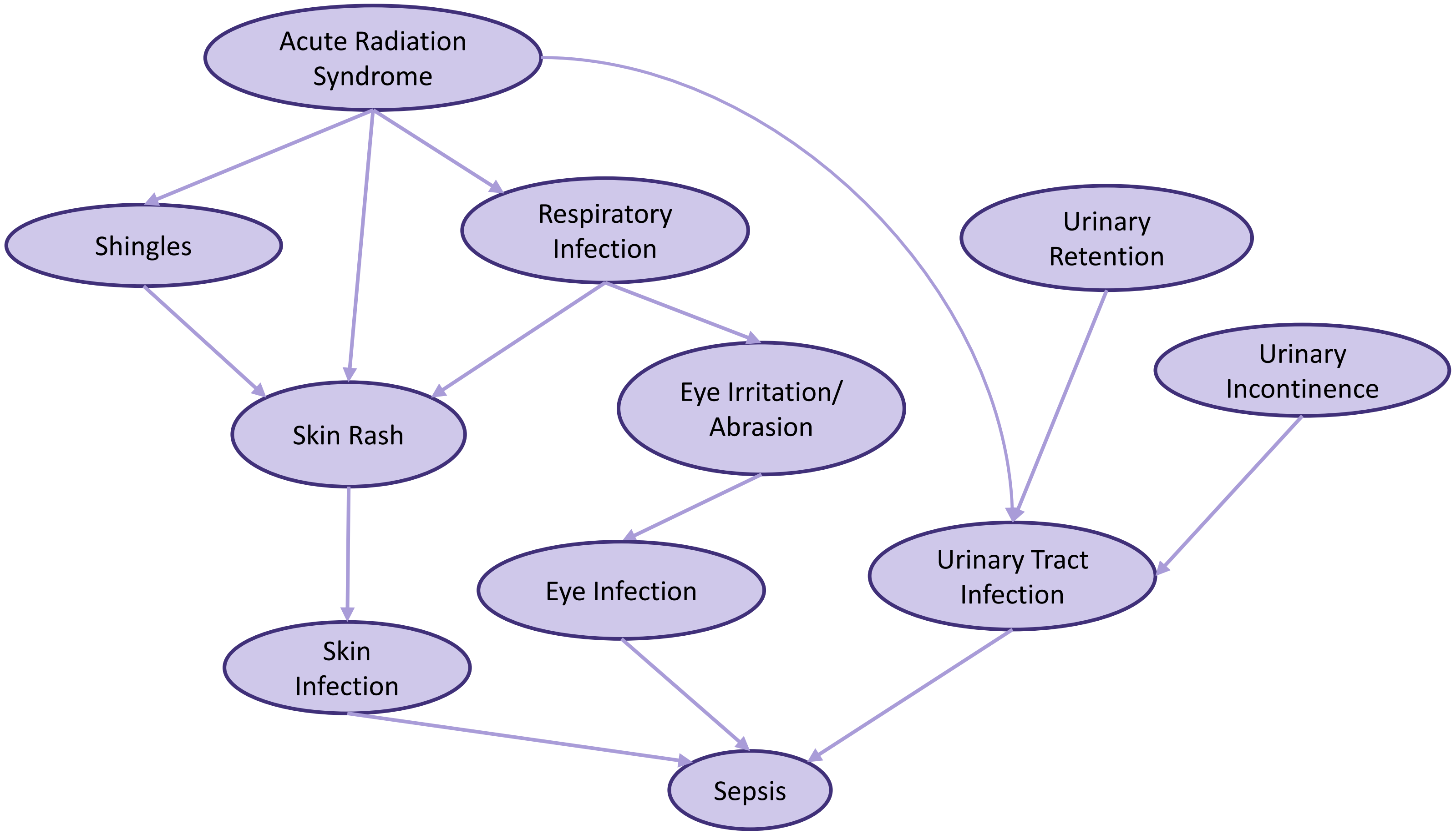
  = sum of weighted number of all walk of length 3



Walk in SIN Network

For adjacency matrix A

- Weighted number of walks of length k $=A^k \cdot 1$
- Sum of weighted number of walks of any length $=(A + A^2 + A^3 + \cdots) \cdot 1$



Katz Centrality

- The Katz Centrality of vertex $\boldsymbol{i}$ in a graph with adjacency matrix A:

$$\boldsymbol{C}(\alpha)_{\boldsymbol{i}} = (\alpha A + \alpha^2 A^2 + \alpha^3 A^3 + \cdots) \cdot W$$

$$= \left((I - \alpha A)^{-1} - I \right) \cdot W$$

- Katz parameter α to penalize longer walks

- Criteria to choose α not well defined in the literature

Katz Parameter

- **Our goal 2:** give guidance to the choice of α

- Based on the maximum path length of interest L when navigating the network

- Takeaway: for maximum path length L and error tolerance $\boldsymbol{\epsilon}$, choose $\boldsymbol{\alpha}$ that satisfies

$$\left\lceil \log_{\alpha\|A\|_2}\left(\frac{\epsilon}{\|C(\alpha)\|_2}\right)\right\rceil = L$$

- Our propositions and theorems:

Proposition 1 *Let C and C' be two centrality measures on a network N. If*

$$\|C - C'\|_\infty < \epsilon,$$

then C and C' ϵ-agree.

Lemma 2 *(Absolute Error Tolerance) Let $p \in \{1, 2, \infty\}$ and $\alpha \in (0, 1/\rho)$. Then*

$$\|C(\alpha) - C(\alpha, \ell)\|_\infty \leq (\alpha\|A\|_p)^\ell \|C(\alpha)\|_p := \epsilon_\ell.$$

Lemma 3 *(Error Tolerance Guarantee) Let $p \in \{1, 2, \infty\}$, $\alpha \in (0, 1/\|A\|_p)$ and $\epsilon > 0$. If*

$$\ell > \log_{\alpha\|A\|_p}\left(\frac{\epsilon}{\|C(\alpha)\|_p}\right) := L$$

then $\|C(\alpha) - C(\alpha, \ell)\|_\infty < \epsilon$.

Corollary 4 (Relative Error) *Let G be a graph, A the adjacency matrix of G, $\alpha = \alpha_0/\|A\|_2$, and $\epsilon = \epsilon_0\|c_\alpha\|_2$ where $\alpha_0, \epsilon_0 \in (0, 1)$. Then α-Katz centrality and (α, ℓ)-Katz centrality are ϵ-tolerant whenever $\ell > \log_{\alpha_0}(\epsilon_0) = L$.*

Corollary 5 (Local Katz Centrality) *Let G be a directed graph, $\alpha \in (0, 1/\|A\|_2)$, $\epsilon > 0$, and L be as defined in Lemma 2. Let H be the subgraphs formed by the L-hop neighborhood of v in G. Then the α-Katz centrality of $v \in G$ is within ϵ of the α-Katz centrality of $v \in H$.*

Application to SIN

Example Mission

Given short mission details:

1. 6 months long
2. L = 6
3. Error tolerance $\epsilon_0 = 0.001$

Use calculated $\alpha = 0.37$

Rank (QTL)	Condition
1	Acute Radiation Syndrome
2	Anaphylaxis
3	Allergic Reaction
4	Toxic Exposure
5	Abdominal Injury

Rank (LOCL)	Condition
1	Acute Radiation Syndrome
2	Burns Secondary to Fire
3	Appendicitis
4	Urinary Tract Infection
5	Sudden Cardiac Arrest

Example Mission

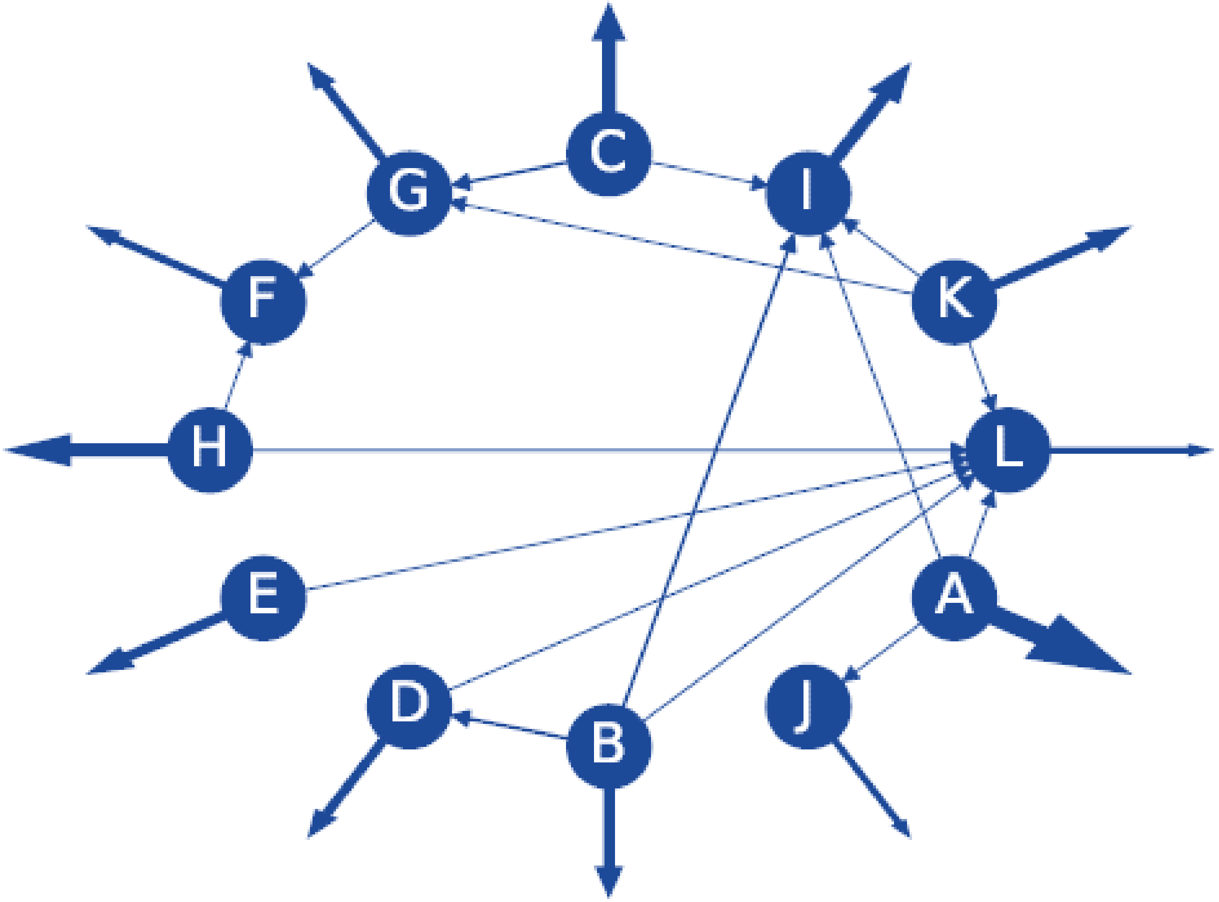
Given long mission details:

1. 3 years long
2. L = 15
3. Error tolerance $\epsilon_0 = 0.001$

Use calculated $\alpha = 0.80$

Rank (QTL)	Condition	Rank (LOCL)	Condition
1	Acute Radiation Syndrome	1	Acute Radiation Syndrome
2	Anaphylaxis	2	Burns Secondary to Fire
3	Toxic Exposure	3	Appendicitis
4	Allergic Reaction	4	Urinary Tract Infection
5	Abdominal Injury	5	Sudden Cardiac Arrest

Subnetwork of the SIN with Highest Centrality Nodes



For α=**0.37**

- 12 top ranked conditions in the SIN in contributing to the **QTL** risk
- Thickness of edges pointing outwards reflecting total edge weight towards nodes in the remainder of the network
- Node A in first rank: Acute Radiation Syndrome

Conclusions and Limitations

Conclusions and Limitations

Conclusions

- The nodes Katz centrality scores depend on Katz parameter α.
- We recommend an α based on the depth to navigate in the network.
- Medical conditions are ranked differently for different mission profiles, based on the number of progressions of concern during each mission.

Limitations

- α is bounded by the inverse of the spectral radius of the graph's adjacency matrix which limits the range of α to consider.
- We can get smaller margin of error computationally than analytically.

Thank you

mona.matar@nasa.gov