Evaluation of Learning-Based Screening Policies for Early Detection of Hepatocellular Carcinoma (HCC)

DSL Seminar
December 7th, 2018

Ryerson University

Data Science Laboratory

Outline

- Background and Context Overview
- Policies
- Simulation Overview
- Toy Example
- Paper Results: Performance Measures Report
- My Replication Results: Challenges and Performance Measures Report

Background and Context Overview

Ryerson University

Early Detection of HCC

- Early Detection: important factor in survival
  - Less than 10% of late-stage cancer diagnosed patients survive beyond 5 years
  - 50% of early-stage cancer diagnosed patients survive beyond 5 years
- Risk for HCC is on the rise
  - US: rate has tripled from 1975 to 2005
  - Canada: tripled in men and doubled in women since 1970
  - Worldwide: 3rd leading cause of cancer death

Current Screening Practice

- In US, screen all at-risk patients every six months
- At-risk patients: chronic hepatitis C and advanced fibrosis
- Screening Process:
  - Ultrasound image of patient liver
  - Blood-test measure for alphafeto-protein level (AFP)

Current Screening Practice Challenges

- Interval Screening
  - Current screening does not consider past AFP patterns
  - It has been shown that AFP past patterns significantly improve estimation of patient's risk of developing HCC: Standard Deviation and Rate of Rise

Current Screening Practice Challenges

- Screening Capacity
  - Number of patients at risk greater than the number of screening resources available
  - Applicable in Third World Countries and in face of soaring healthcare costs (US)

Overcoming the Challenges

- Interval Screening
  - Solution approach: Take into account ‘sequential learning’
  - At each screening, compute: standard deviation and rate of rise of AFP based on past screening values of AFP
- Screening Capacity
  - Solution approach: choose the best subset of patients to screen based on whoever is most likely to develop HCC
  - **How to choose the best subset? Determined by policy**

Learning-Based Policy

- A *policy*: determines which action should be performed in each state

- In our context: Policy refers to the ‘action’ of sub-setting the best patients to screen

- The aim: Test different policies and assess their performance

Data Source

HALT-C: Hepatitis C Antiviral Long-Term Treatment against Cirrhosis

- The HALT-C study includes patients that have chronic Hepatitis C and fibrosis
- They follow the patients for around 5.3 years
- Obtain the following measurement at specified times:
    - AFP Screening
    - Ultrasound imaging
    - Liver Biopsy

Data publically available (requires access authorization): https://repository.niddk.nih.gov/studies/halt-c/

Research Data

- 967 of HALT-C subjects are selected
- HCC cases were match to control without HCC in 1:3 ration

- Clinical and Demographic Variables
  - Static Data: measured upon enrollment into surveillance (age, blood platelets, smoker etc.)
  - Dynamic Data: AFP values measured at each screening time t

Data Science Laboratory

Ryerson University

HCC Probability Equation

$$P(\mathrm{HCC})_i = [1 + exp(-c_1 B_i - c_2 SD_i - c_3 RR_i)]^{-1}$$

- B_i: vector of all static variables
- SD_i: standard deviation amongst a patient's recorded AFP readings
- RR_i: Rate of AFP rise over time amongst a patient's recorded AFP readings

- c_1: vector of corresponding regression coefficients for all static risk factors
- c_2, c_3: regression coefficients for the AFP standard deviation and rate of AFP rise over time

Data Science Laboratory

Ryerson University

Policies

Ryerson University

Policy 1: Myopic Policy

Data Science Laboratory

1. Until t = T
2. Rank the n patients in descending order with respect to

$$P(\mathrm{HCC})_i = [1 + exp(-c_1 B_i - c_2 SD_i - c_3 RR_i)]^{-1}$$

3. Screen the k patients at the top of this list
4. Advance to time t+1

Ryerson University

Policy 2: ϵ-greedy strategies

- Choose ϵ belonging to [0,1]

1. Until t = T

2. Rank the N patients in non-descending order with respect to

$$P(\mathrm{HCC})_i = [1 + exp(-c_1 B_i - c_2 SD_i - c_3 RR_i)]^{-1}$$

3. Screen the (1- ϵ).k patients at the top of this list
4. Screen ϵ.k patients from the remaining patients randomly

5. Advance to time t+1

Policy 3: Interval Estimation Strategies

- Choose *z* belonging to [0, ∞)

1. Until t = T

2. Rank the N patients in non-descending order with respect to

$$c_1 \hat{B}_i + c_2(\hat{SD}_{i,t} + z \cdot \sqrt{v_{i,t}}) + c_3(\hat{RR}_{i,t} + z \cdot \sqrt{w_{i,t}})$$

3. Screen the k patients at the top of this list

4. Advance to time t+1

Policy 4: Boltzmann Exploration Strategies

- Choose τ belonging to (0, ∞)

1. Until t = T

2. Screen the k patients where patient i is screened with probability $\boxed{\frac{e^{x_{i,t}/\tau}}{\sum_{i'=1}^{n} e^{x_{i',t}/\tau}}}$ where the original risk score $x_{i,t}$ is:

$$x_{i,t} = (c_1 B_i + c_2 \tilde{SD}_{i,t} + c_3 \tilde{RR}_{i,t})$$

3. Advance to time t+1

Ryerson University

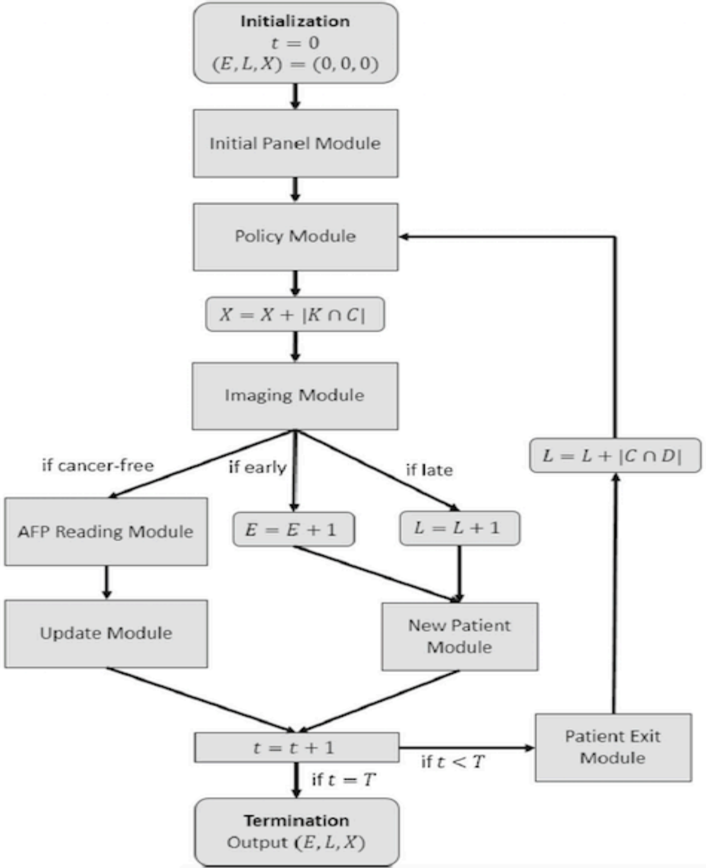
Method of Policy Evaluation: A Simulation

- Reinforcement Learning Component:
  - P(HCC) learning probability dynamically as AFP measures are updated
- Simulation input:
  - List of n patients, k: number of k patients to subset, T: iteration time
- Simulation Output:
  - **E**: number of early stage cancers detected, **L**: number of late stage cancer detected, **X**: number of screenings spent on patients who would eventually develop HCC

Simulation Overview

Ryerson University

Simulation Overview



Performance Measures:

1. Proportion of cancers detected in early stage: **E/(E+L)**

2. Proportion of resources spent on patients who eventually develop cancer: **X/(k*T)**

Ryerson University

Initial Panel Module

- Select n number of patients
- Create:
    - C: set of patients who will develop cancer
    - N: set of patients who will not develop cancer

Policy Module

- Receives knowledge of n patients as input
- Chooses subset k based on Policy

- Additionally:
    - X increments by number of screenings spent on $|k \cap C|$
    - $|k \cap C|$: subset of k chosen patients that will actually develop cancer

Imaging Module

- Query Current Cancer State
    - If patient never developed cancer ➔ assign cancer-free state
    - If patient has tumor of size $\bar{s}$ at time $\bar{t}$ tbar
        - Estimate tumor: $s = 2^{\frac{t-\bar{t}}{\delta}} \cdot \bar{s}$
        - Assign state as per

$$\text{State} = \begin{cases} \text{Early} & \text{if } t \geq \bar{t} \text{ and } 1 \leq s \leq 5 \\ \text{Late} & \text{if } t \geq \bar{t} \text{ and } 5 < s \\ \text{Cancer-Free} & \text{if Otherwise} \end{cases}$$

Ryerson University

Imaging Module Output

- If Cancer-Free: Run AFP Module and Update Module
  - Query new AFP readings
  - Calculate RR and SS
  - Update patient information
- If Early or Late: Run New Patient Module
  - Increment E or L
  - Patient exits simulation and is replaced by new patient

Patient Exit Module

- Find subset of k: D ➔ Departing Patients
    - Time of screening t >= time of simulation T
    - Death
    - Voluntary Withdrawal

- Patients in D are eliminated and replaced

- Increment L by $|D \cap C|$
    - $|D \cap C|$: The patients who are eliminated without being detectd with early cancer but who would eventually have cancer are considered late detections

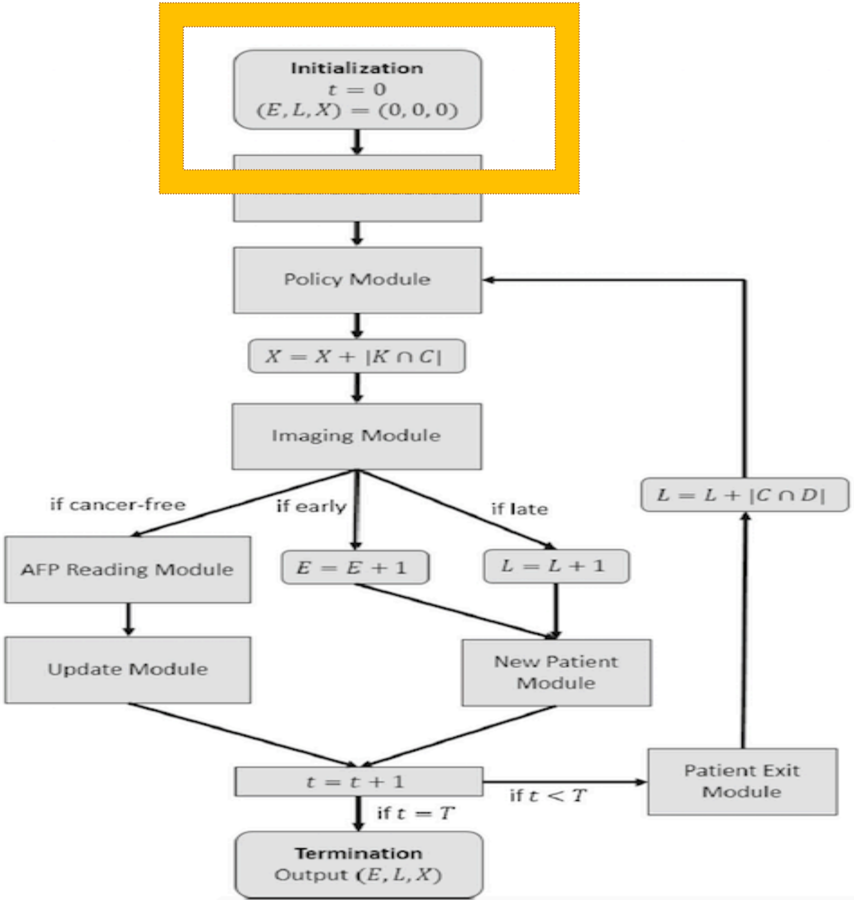
Toy Example

Run the Simulation until time T = 4

Initialization

E = 0
L = 0
X = 0
t = 0

T = 4
k = 3
n = 10



Initial Panel Module

```
> N_Patient_list
$`1`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
1   1  95  0  0    2    2      1 1   C

$`2`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
2   2  80  0  0   -1    2      2 1   C

$`3`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
3   3  81  0  0    1    1      3 1   C

$`4`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
4   4  94  0  0    1    0      0 1   N
```

$n = \{P_{1,}\ P_{2,}\ P_{3,}\ P_4,\ P_{5,}\ P_{6,}\ P_7,\ P_8,\ P_9,\ P_{10)}$

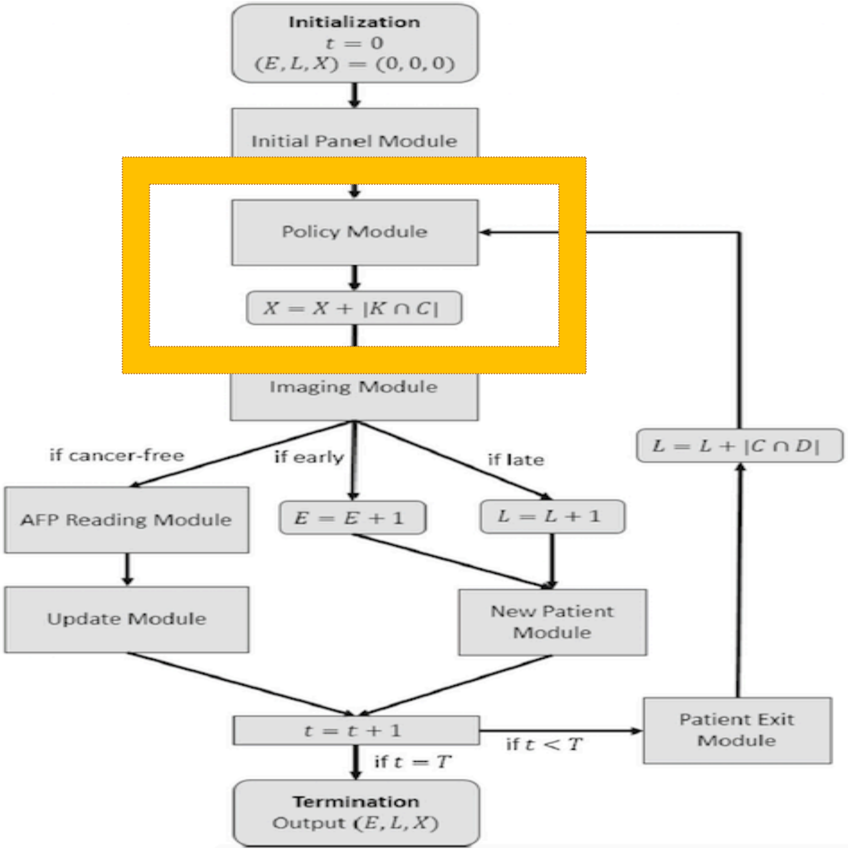
$C = \{P_{1,}\ P_{2,}\ P_3\}$

$N = \{P_4,\ P_{5,}\ P_{6,}\ P_7,\ P_8,\ P_9,\ P_{10}\}$

Ryerson University

Data Science Laboratory

Policy Module at t=0



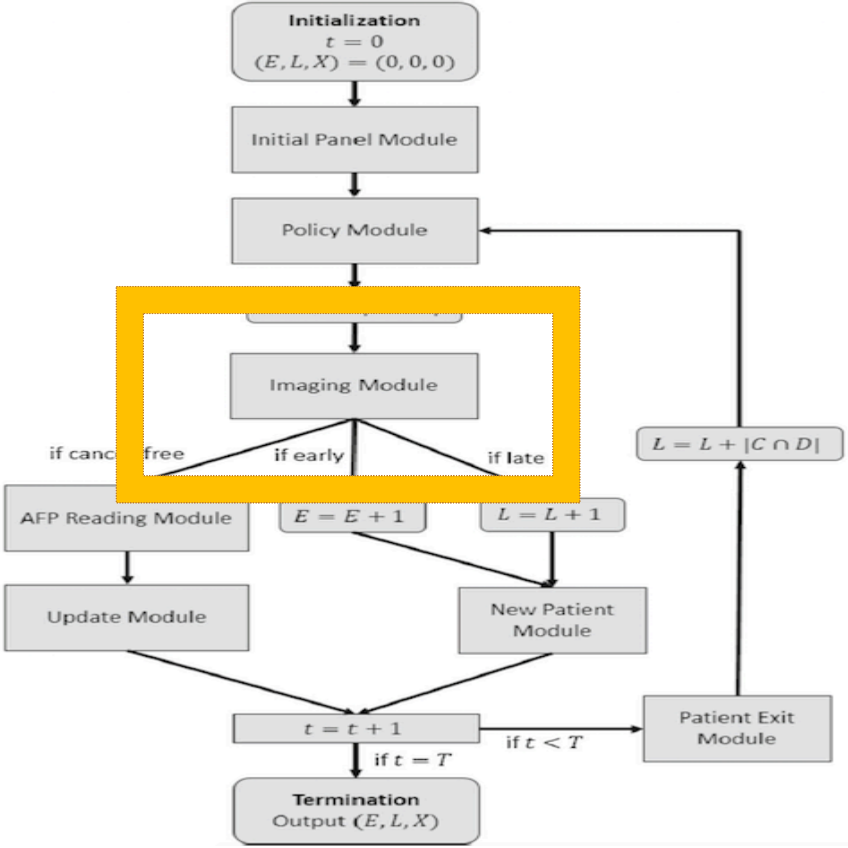
Policy Module Output

- Result: k patient selection ➔ k subset

```
> policy_modHCC(N_Patient_list, 10)
        1         3         4        10         7         8         9         2         5         6 
0.8807971 0.7310586 0.7310586 0.7310586 0.3543437 0.3543437 0.3100255 0.2689414 0.1824255 0.1824255
```

- Increment X
    - $|K \cap C| = \{P_1, P_3\}$
    - $X = X + |K \cap C| = 2$

Imaging Module at t=0



Ryerson University

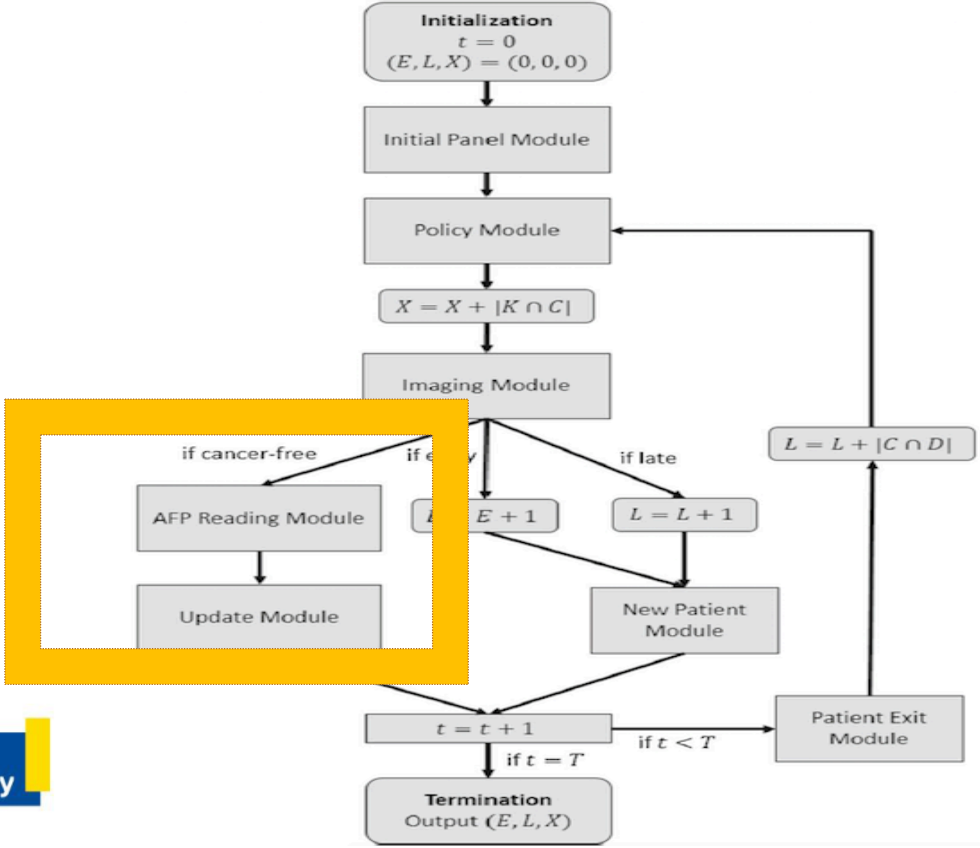
Data Science Laboratory

Imaging Module Output

- At time t=0, every patient is cancer-free as the dataset used was a clinical trial whose enrollment criteria included being cancer-free at the beginning of surveillance

- Output
    - P_1 ➔ Cancer-Free
    - P_3 ➔ Cancer-Free
    - P_4 ➔ Cancer-Free

Imaging Module Cancer-Free Output at t=0

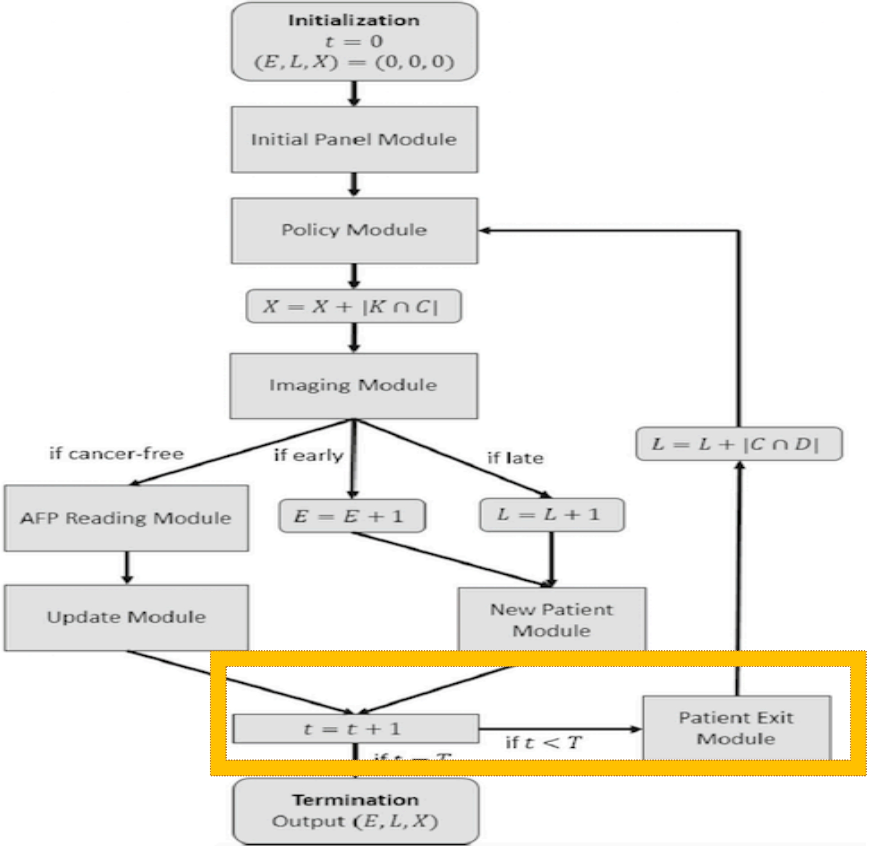


Imaging Module Output: Cancer-Free

```
> Cancer_Free
[[1]]
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
1   1  95  0  0    2    2      1 1   C

[[2]]
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
3   3  81  0  0    1    1      3 1   C

[[3]]
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
4   4  94  0  0    1    0      0 1   N
```

```
> N_Patient_list
$`1`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C

$`2`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
2   2  80  0  0   -1    2      2 1   C

$`3`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
3   3  81 0.000000  0    1    1      3 1   C
1   3  88 4.949747  7    1    1      3 2   C

$`4`
  Pid AFP        SS RR c1Bi sbar t_sbar t C_N
4   4  94 0.0000000  0    1    0      0 1   N
1   4  93 0.7071068 -1    1    0      0 2   N
```

Advance time and Patient Exit Module at t=1



Ryerson University

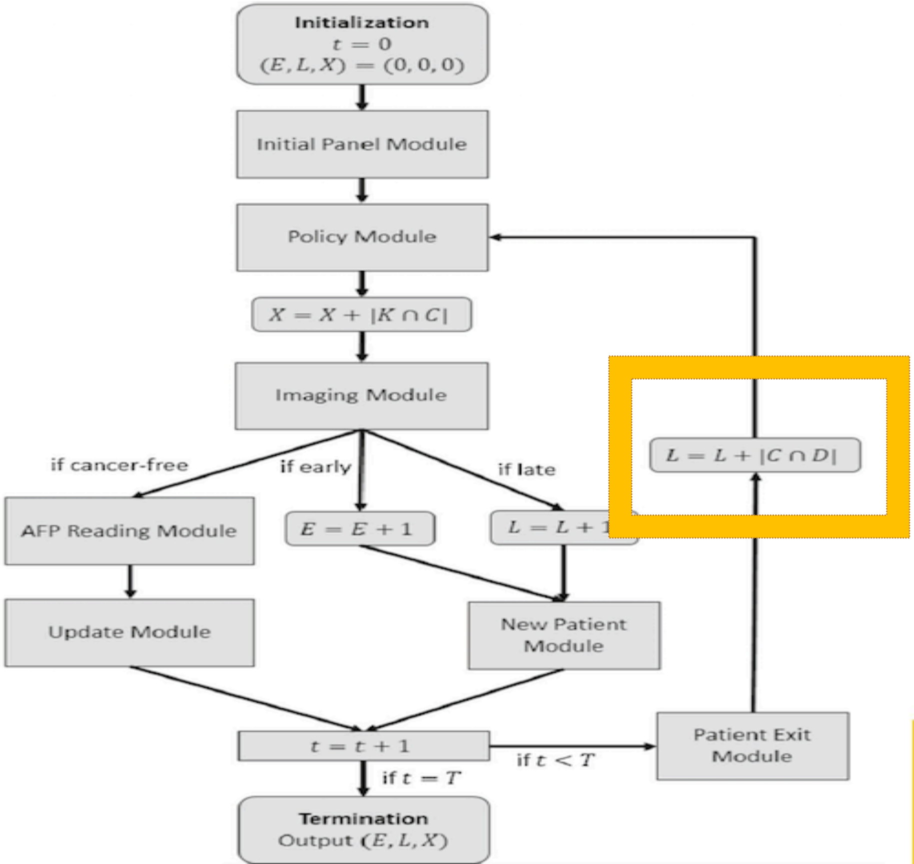
Patient Exit Module

```
$`1`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C

$`3`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
3   3  81 0.000000  0    1    1      3 1   C
1   3  88 4.949747  7    1    1      3 2   C

$`4`
  Pid AFP        SS RR c1Bi sbar t_sbar t C_N
4   4  94 0.0000000  0    1    0      0 1   N
1   4  93 0.7071068 -1    1    0      0 2   N
```

D = {}

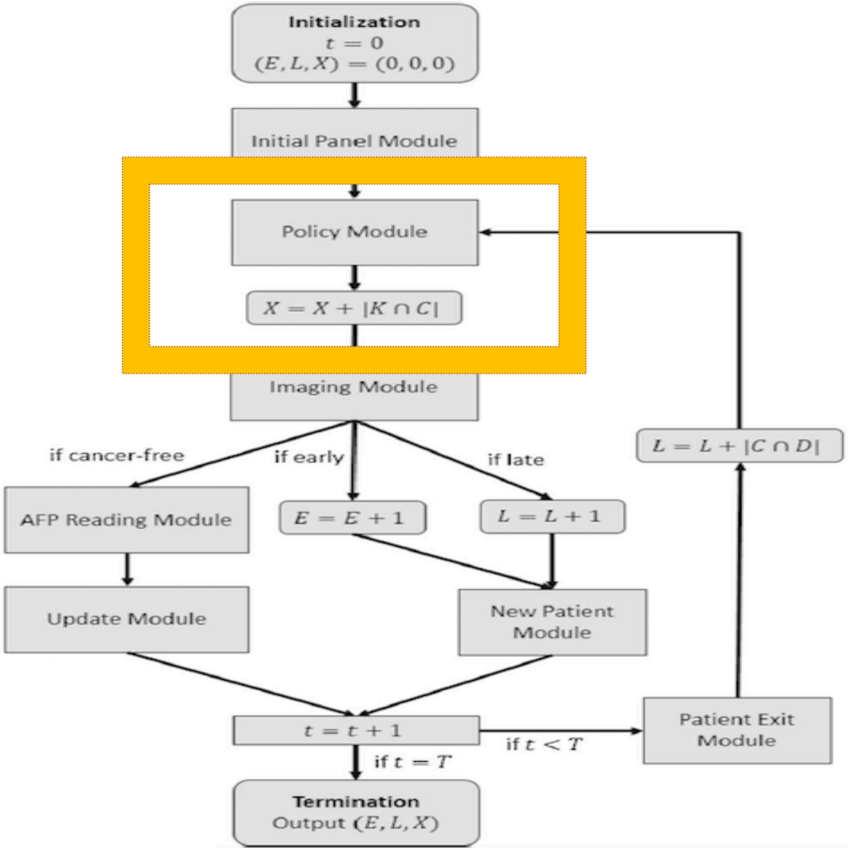
No departing patients

Incrementing L at t=1



No Departing Patients, L does not increment

Policy Module at t=1



Policy Module Output

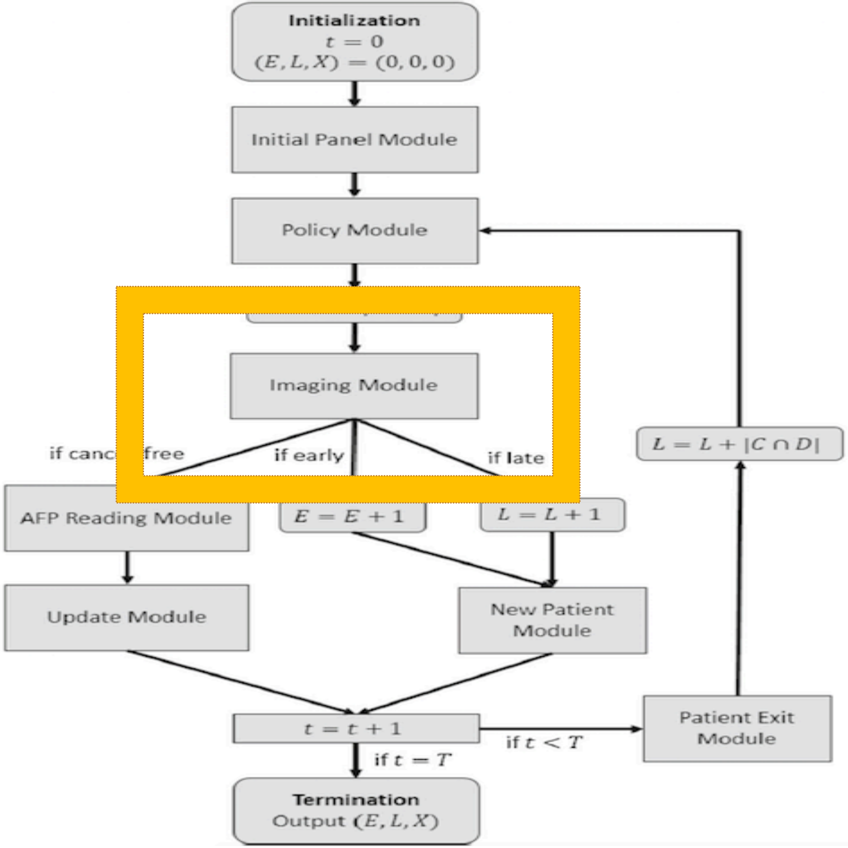
- Result: k patient selection ➔ k subset

```
> policy_modHCC(N_Patient_list, 10)
        3         1         4         7         8         9        10         2         5         6 
0.9999992 0.9999877 0.6413547 0.3543437 0.3543437 0.3100255 0.3100255 0.2689414 0.1824255 0.1824255 
```

- Increment X
    - $|K \cap C| = \{P_1, P_3\}$
    - $X = 2 + |K \cap C| = 4$

Ryerson University

Imaging Module at t=1



Imaging Module Output

- If patient belongs to N → Assign Cancer-Free
- If patient belongs to C → Estimate tumor and assign

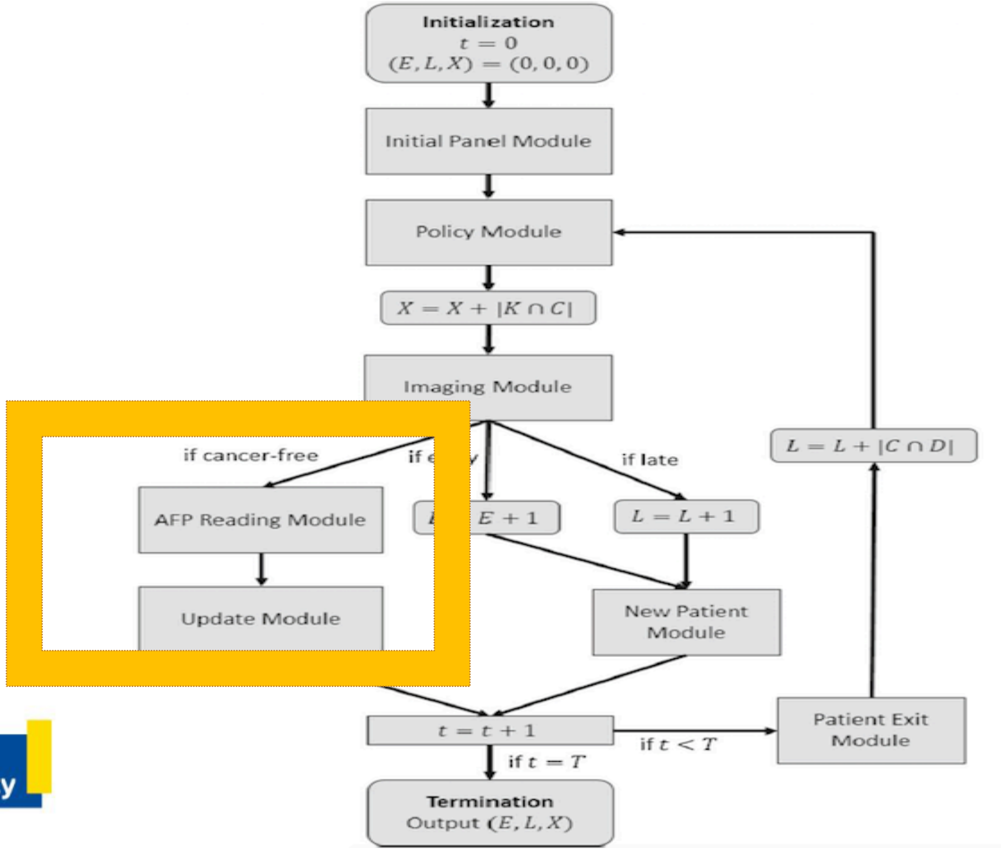
$$\text{State} = \begin{cases} \text{Early} & \text{if } t \geq \bar{t} \text{ and } 1 \leq s \leq 5 \\ \text{Late} & \text{if } t \geq \bar{t} \text{ and } 5 < s \\ \text{Cancer-Free} & \text{if Otherwise} \end{cases}$$

- Output
  - P_4 ➔ Cancer-Free
  - P_3 ➔ Estimated tumor = 0.06 ➔ Cancer-Free
  - P_1 ➔ Estimated tumor = 2 ➔ Early Cancer

Data Science Laboratory

Ryerson University

Imaging Module Cancer-Free Output at t=1



Imaging Module Cancer-Free Output

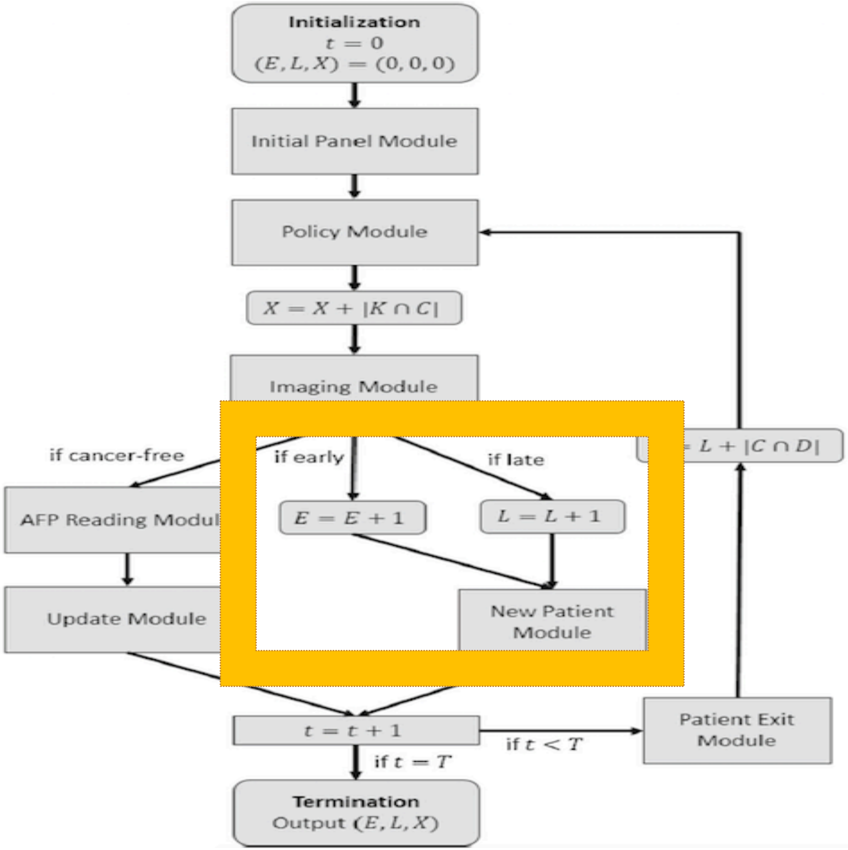
```
> Cancer_Free
[[1]]
  Pid AFP        SS RR c1Bi sbar t_sbar t C_N
3   3  81 0.000000  0    1    1      3 1   C
1   3  88 4.949747  7    1    1      3 2   C

[[2]]
  Pid AFP         SS RR c1Bi sbar t_sbar t C_N
4   4  94 0.0000000  0    1    0      0 1   N
1   4  93 0.7071068 -1    1    0      0 2   N
```

```
> N_Patient_list
$`1`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C

$`2`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
2   2  80  0  0   -1    2      2 1   C

$`3`
   Pid AFP       SS RR c1Bi sbar t_sbar t C_N
3    3  81 0.000000  0    1    1      3 1   C
1    3  88 4.949747  7    1    1      3 2   C
11   3  88 4.041452  0    1    1      3 3   C

$`4`
   Pid AFP        SS RR c1Bi sbar t_sbar t C_N
4    4  94 0.0000000  0    1    0      0 1   N
1    4  93 0.7071068 -1    1    0      0 2   N
11   4  94 0.5773503  1    1    0      0 3   N
```

Ryerson University

Data Science Laboratory

Imaging Module Cancer Output at t=1



Ryerson University

Data Science Laboratory

Imaging Module Cancer Output

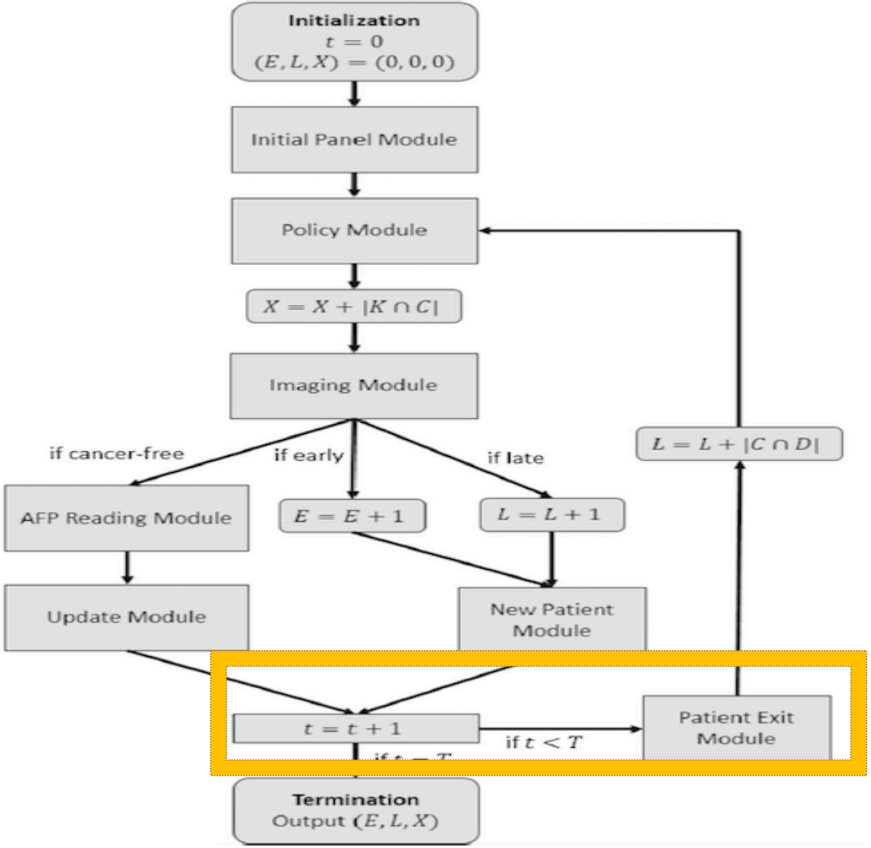
- Increment E and L ➔ E = E+ 1 = 0 + 1
- Replace Patients

```
> Replace_Patient
[[1]]
  Pid AFP        SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C
```

```
> N_Patient_list
$`1`
   Pid AFP SS RR c1Bi sbar t_sbar t C_N
15  99  94  0  0    1    2      2 1   C
```

Ryerson University

Advance time and Patient Exit Module at t=2



Data Science Laboratory

Ryerson University

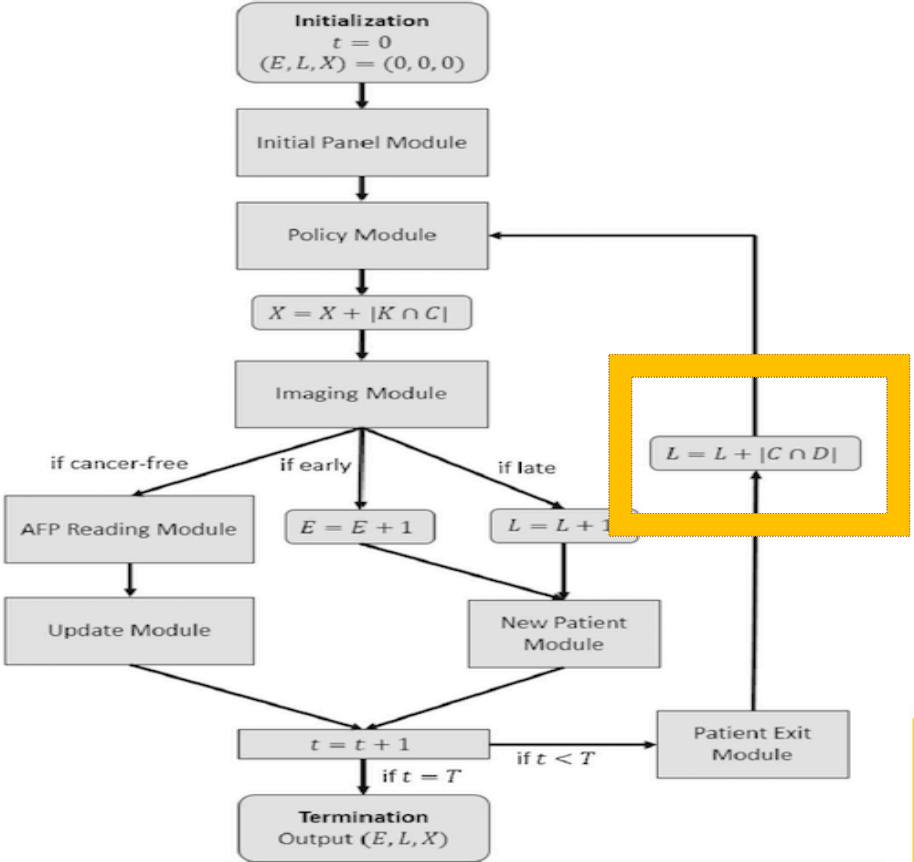
Patient Exit Module

```
$`1`
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C

$`3`
   Pid AFP       SS RR c1Bi sbar t_sbar t C_N
3    3  81 0.000000  0    1    1      3 1   C
1    3  88 4.949747  7    1    1      3 2   C
11   3  88 4.041452  0    1    1      3 3   C

$`4`
   Pid AFP        SS RR c1Bi sbar t_sbar t C_N
4    4  94 0.0000000  0    1    0      0 1   N
1    4  93 0.7071068 -1    1    0      0 2   N
11   4  94 0.5773503  1    1    0      0 3   N
```

$D = \{\}$

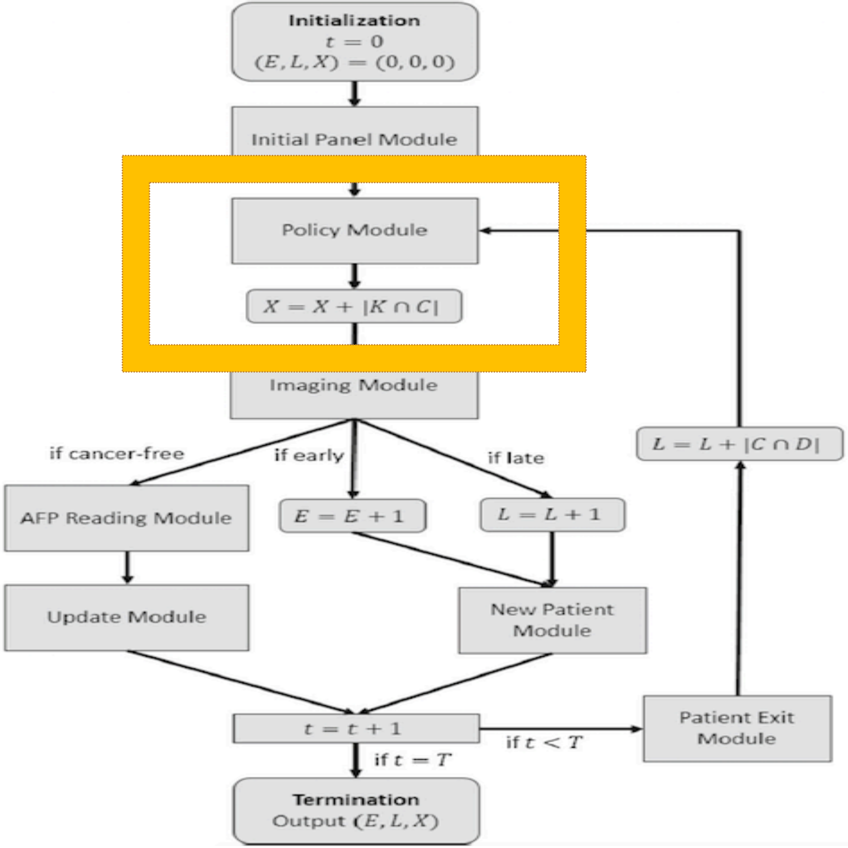
No departing patients

Incrementing L at t=2



No Departing Patients, L does not increment

Policy Module at t=2



Ryerson University

Data Science Laboratory

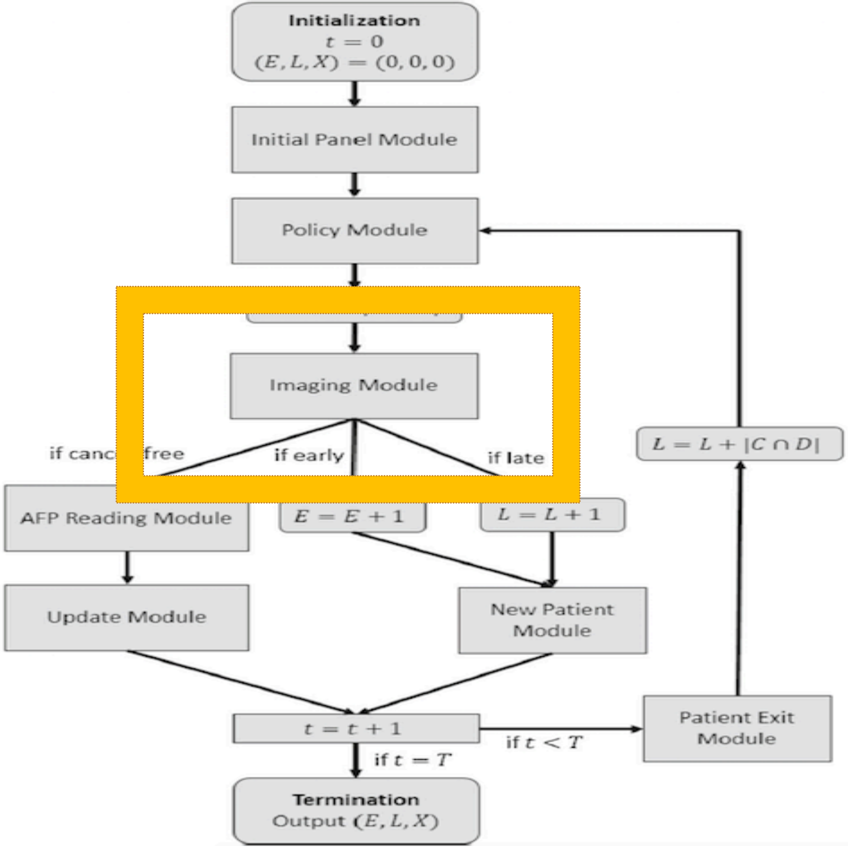
Policy Module Output

- Result: k patient selection ➔ k subset

```
> policy_modHCC(N_Patient_list, 10)
        3         4         1         7         8         9        10         2         5         6
0.9940729 0.9387104 0.7310586 0.3543437 0.3543437 0.3100255 0.3100255 0.2689414 0.1824255 0.1824255
```

- Increment X
  - $|K \cap C| = \{P_{99}, P_3\}$
  - $X = 2 + |K \cap C| = 6$

Imaging Module at t=2



Ryerson University

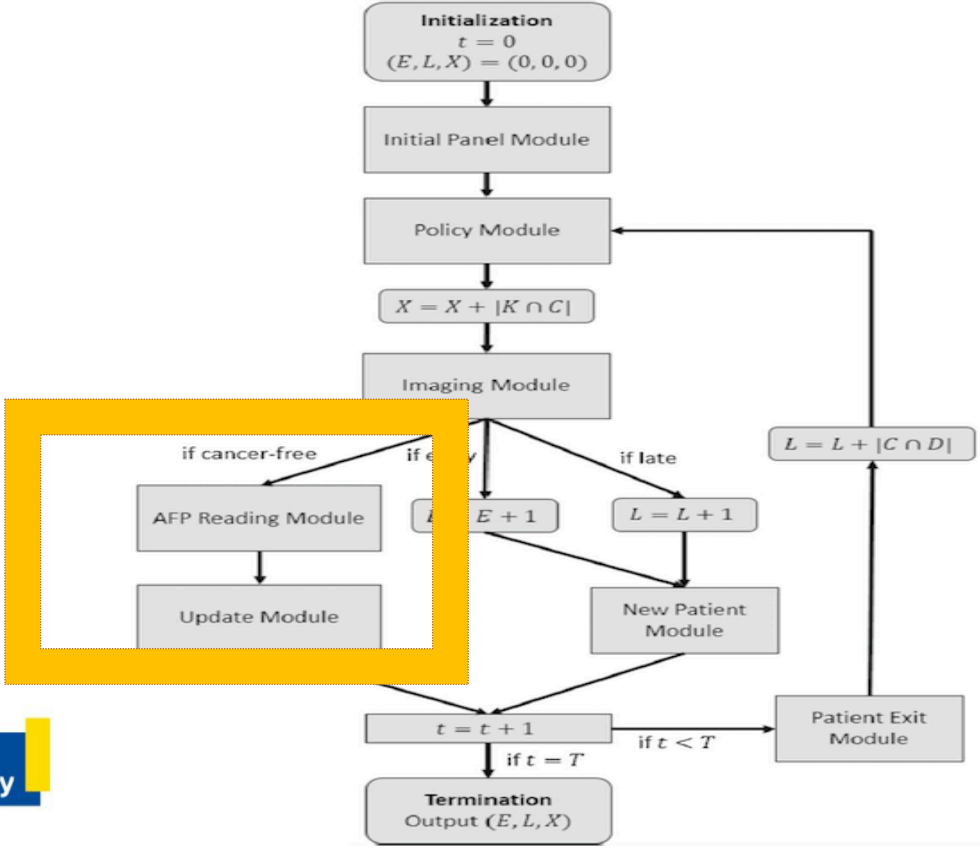
Imaging Module Output

- If patient belongs to N → Assign Cancer-Free
- If patient belongs to C → Estimate tumor and assign

$$\text{State} = \begin{cases} \text{Early} & \text{if } t \geq \bar{t} \text{ and } 1 \leq s \leq 5 \\ \text{Late} & \text{if } t \geq \bar{t} \text{ and } 5 < s \\ \text{Cancer-Free} & \text{if Otherwise} \end{cases}$$

- Output
  - P_4 ➔ Cancer-Free
  - P_3 ➔ Estimated tumor = 0.25 ➔ Early Cancer
  - P_{99} ➔ Estimated tumor = 2 ➔ Early Cancer

Imaging Module Cancer-Free Output at t=2



Ryerson University

Data Science Laboratory

Imaging Module Cancer-Free Output

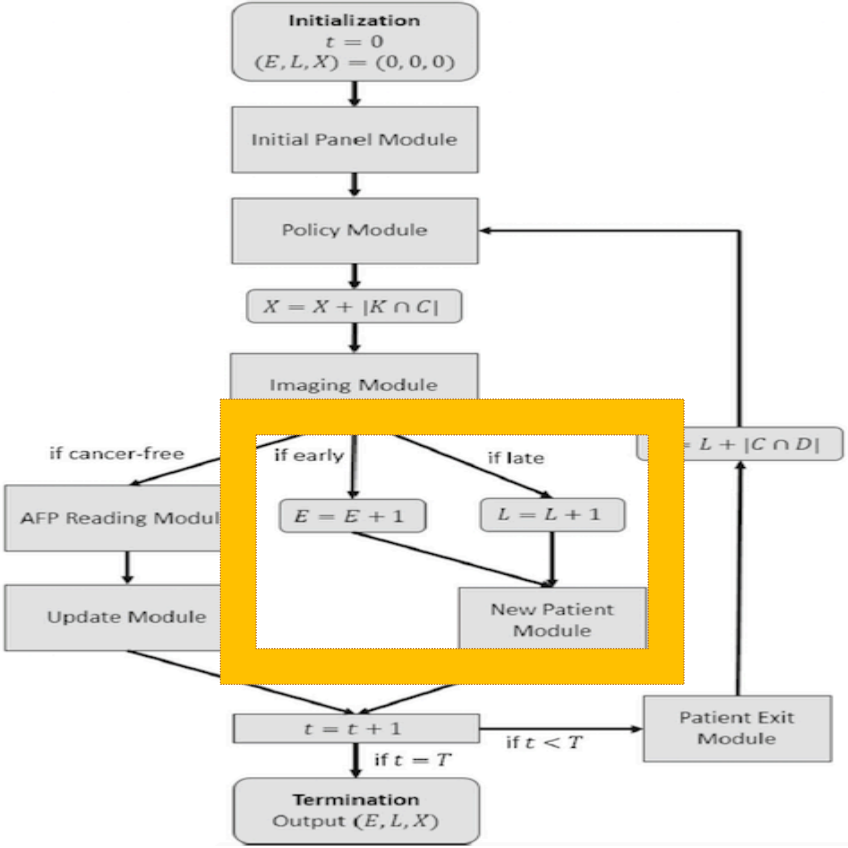
Data Science Laboratory

```
> Cancer_Free
[[1]]
   Pid AFP        SS RR c1Bi sbar t_sbar t C_N
3    3  81 0.000000  0    1    1      3 1   C
1    3  88 4.949747  7    1    1      3 2   C
11   3  88 4.041452  0    1    1      3 3   C

[[2]]
   Pid AFP         SS RR c1Bi sbar t_sbar t C_N
4    4  94 0.0000000  0    1    0      0 1   N
1    4  93 0.7071068 -1    1    0      0 2   N
11   4  94 0.5773503  1    1    0      0 3   N
```

```
> N_Patient_list
$`1`
   Pid AFP SS RR c1Bi sbar t_sbar t C_N
15  99  94  0  0    1    2      2 1   C

$`2`
  Pid AFP SS RR c1Bi sbar t_sbar t C_N
2   2  80  0  0   -1    2      2 1   C

$`3`
   Pid AFP        SS RR c1Bi sbar t_sbar t C_N
3    3  81 0.000000  0    1    1      3 1   C
1    3  88 4.949747  7    1    1      3 2   C
11   3  88 4.041452  0    1    1      3 3   C
12   3  88 3.500000  0    1    1      3 4   C

$`4`
   Pid AFP         SS RR c1Bi sbar t_sbar t C_N
4    4  94 0.0000000  0    1    0      0 1   N
1    4  93 0.7071068 -1    1    0      0 2   N
11   4  94 0.5773503  1    1    0      0 3   N
12   4  94 0.5000000  0    1    0      0 4   N
```

Ryerson University

Imaging Module Cancer Output at t=2



Ryerson University

Imaging Module Cancer Output

- Increment E and L ➔ E = E+1 = 2 ; L = 0
- Replace Patients

```
> Replace_Patient
[[1]]
  Pid AFP       SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C
```

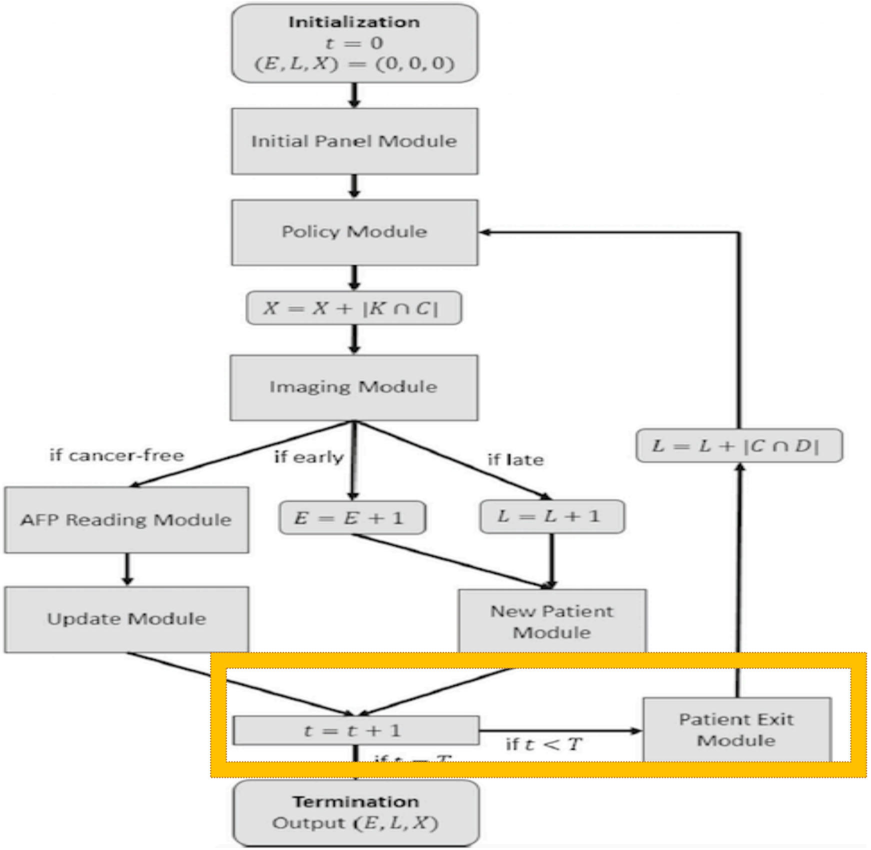
➔

```
> N_Patient_list
$`1`
   Pid AFP SS RR c1Bi sbar t_sbar t C_N
15  99  94  0  0    1    2      2 1   C
```

Ryerson University

Data Science Laboratory

Advance time and Patient Exit Module at t=3



Patient Exit Module

```
$`1`
  Pid AFP        SS RR c1Bi sbar t_sbar t C_N
1   1  95 0.000000  0    2    2      1 1   C
2   1 100 3.535534  5    2    2      1 2   C

$`3`
   Pid AFP        SS RR c1Bi sbar t_sbar t C_N
3    3  81 0.000000  0    1    1      3 1   C
1    3  88 4.949747  7    1    1      3 2   C
11   3  88 4.041452  0    1    1      3 3   C
12   3  88 3.500000  0    1    1      3 4   C

$`4`
   Pid AFP         SS RR c1Bi sbar t_sbar t C_
4    4  94 0.0000000  0    1    0      0 1
1    4  93 0.7071068 -1    1    0      0 2
11   4  94 0.5773503  1    1    0      0 3
12   4  94 0.5000000  0    1    0      0 4
```

$D = \{P_3, P_4\}$

$(t = 4) >= (T=4)$

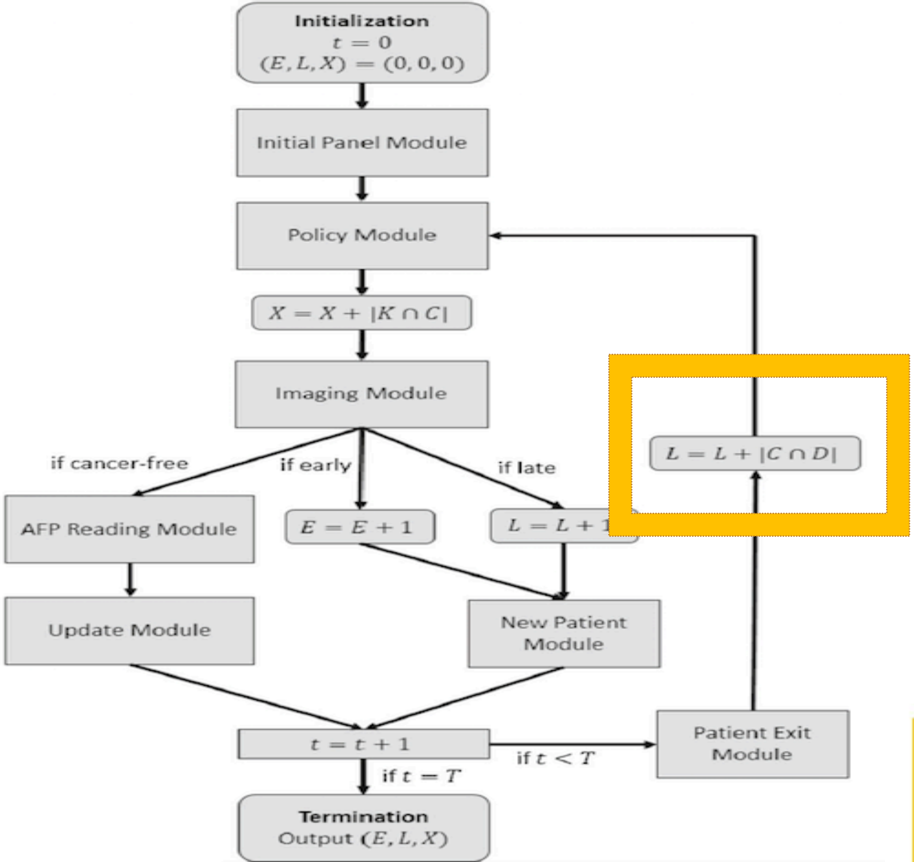
Two departing patients:

$P_3 \rightarrow C$

$P_4 \rightarrow N$

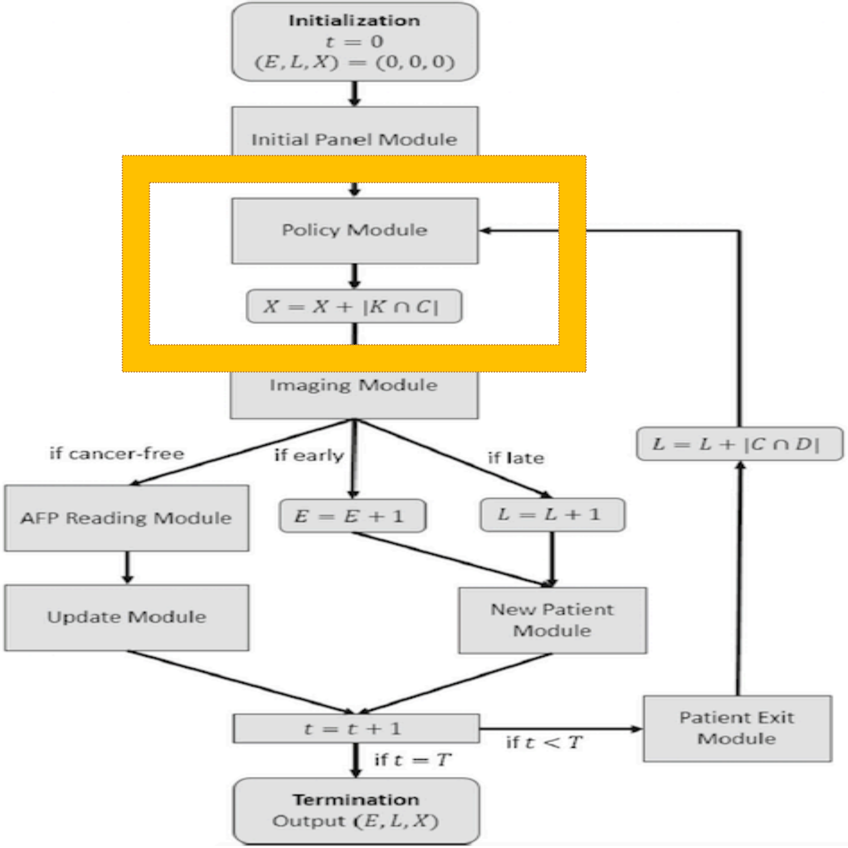
$|D \cap C| = \{P_3\}$

Incrementing L at t=3

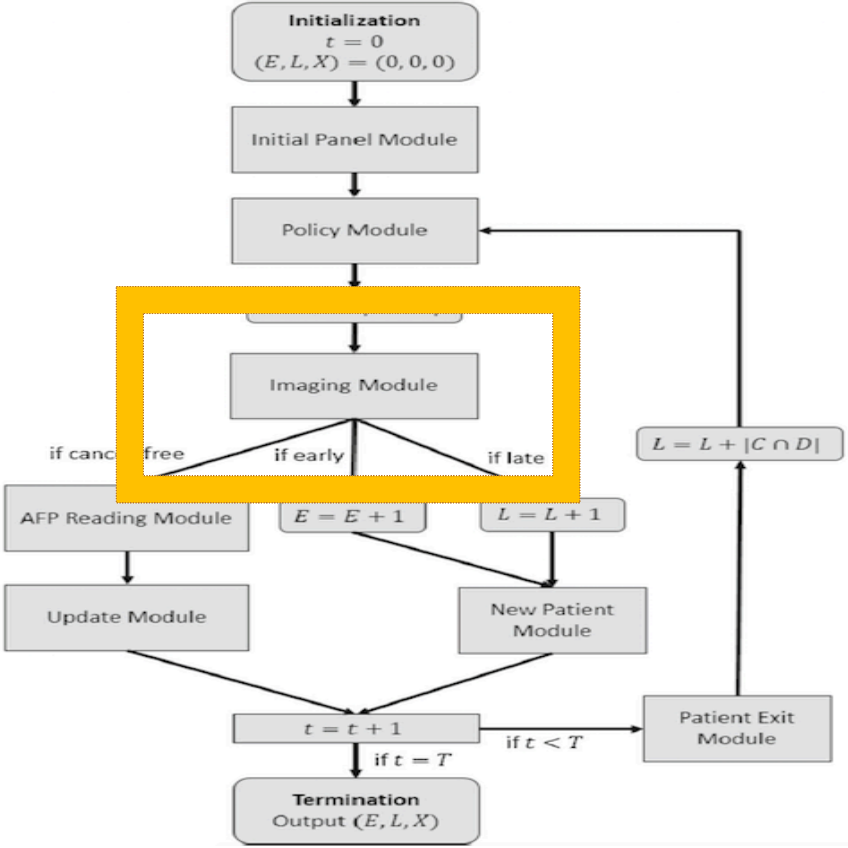


Increment L:
L = 0 + 1 = 1

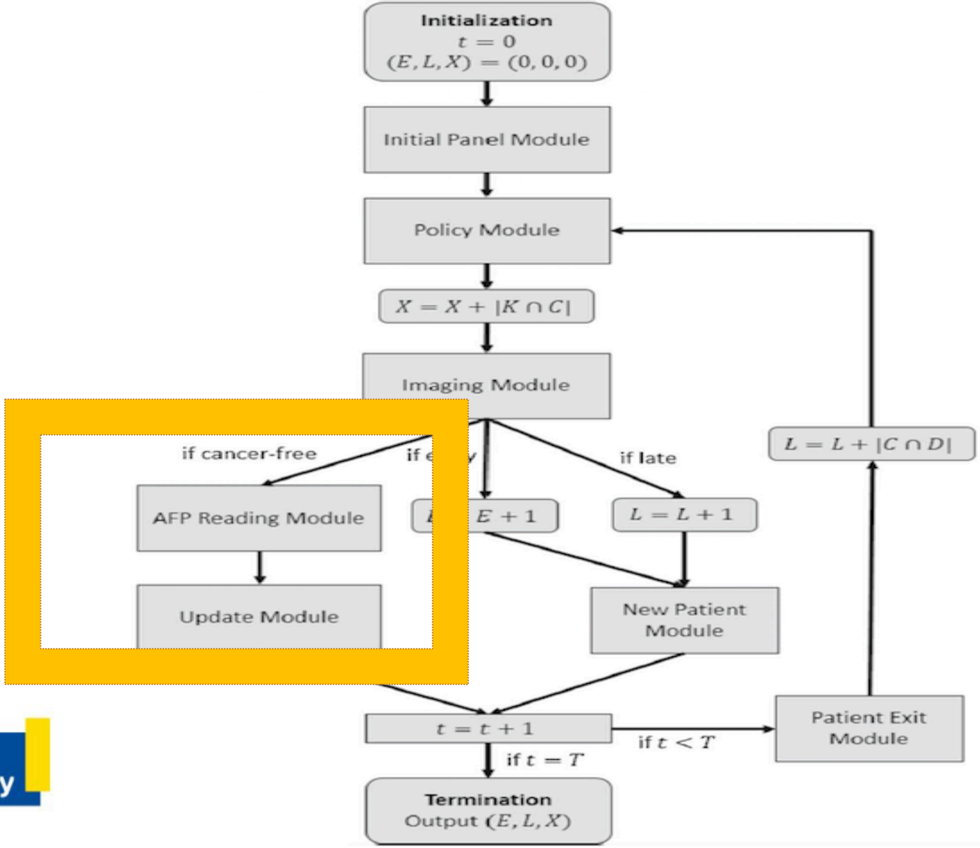
Policy Module at t=3



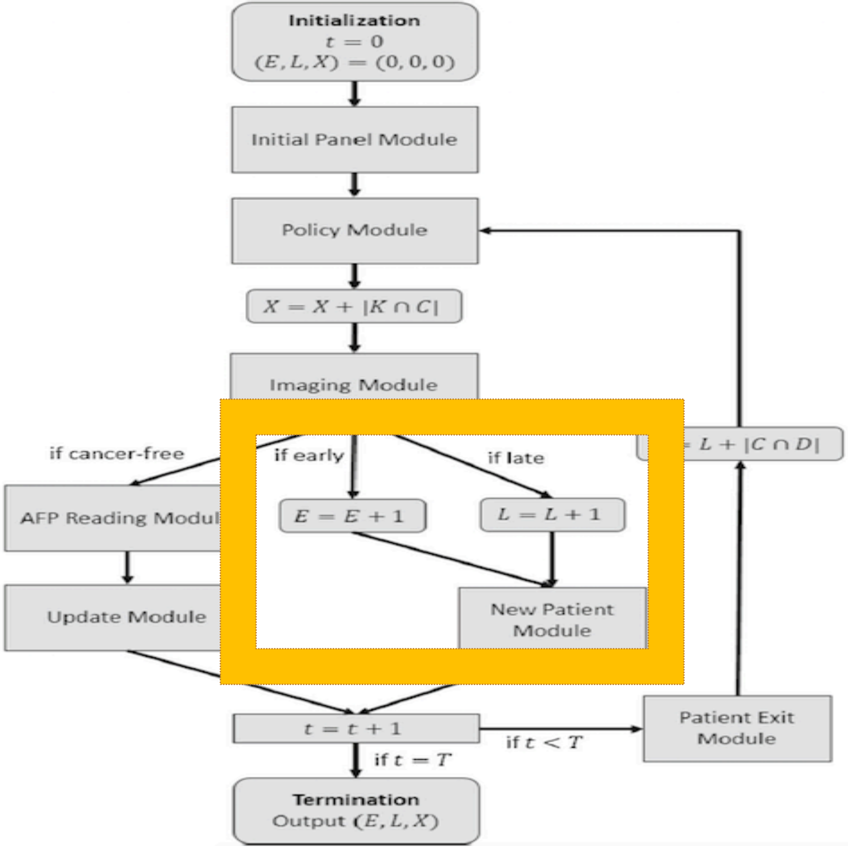
Imaging Module at t=3



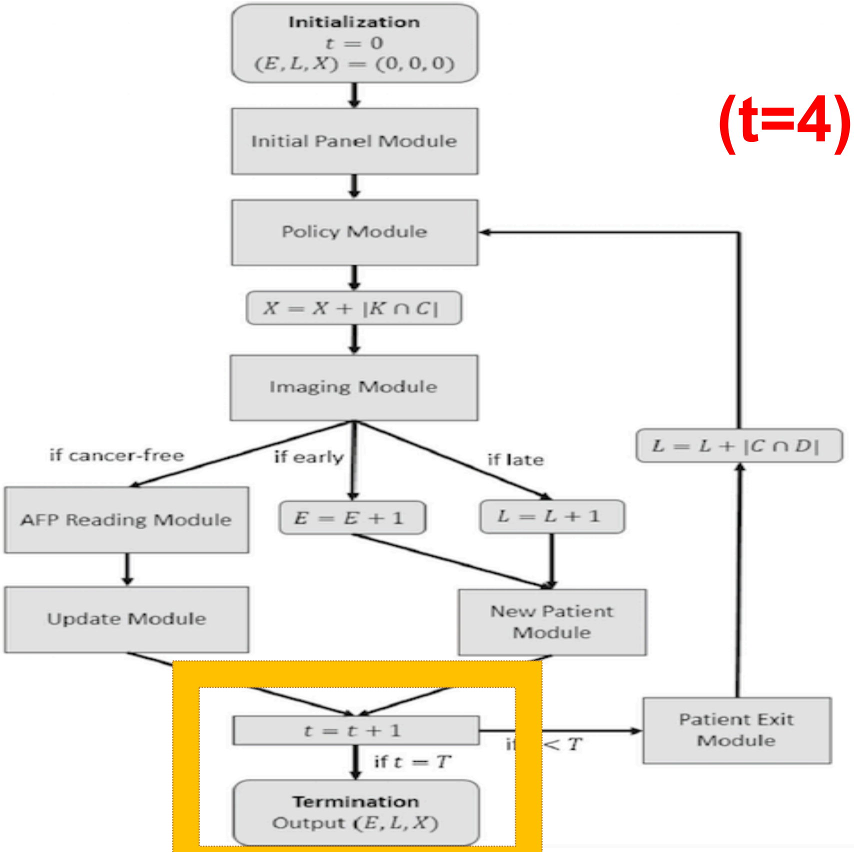
Imaging Module Cancer-Free Output at t=3



Imaging Module Cancer Output at t=3



Advance time and Patient Exit Module at t=4



(t=4) = (T=4) ➔ Termination

Performance Measures

- Termination Output:
  - E = 2
  - X = 6
  - L = 1
- Performance:
  - Early Cancer Detection: E/(E+L) = 2/(2+1) = 0.6
  - Screening Resource Spent on Cancer Patients: X/(k*T) = 6/(3*4) = 0.5

Paper Results

Performance Measure Report

Ryerson University

Optimal Tuning Parameter

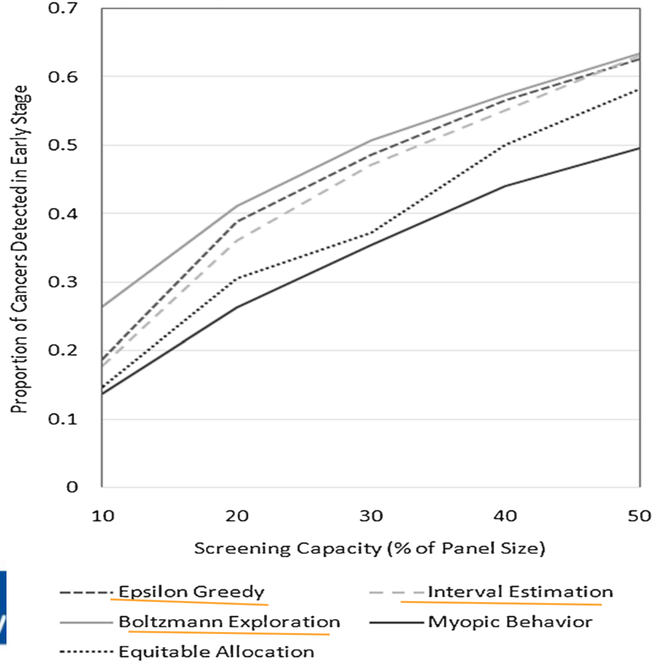
- Tuning Parameter: For each policy, at each constraint level, the optimal parameter used and presented in the paper

Table 2 Optimal tuning parameters determined by the indifference zone method

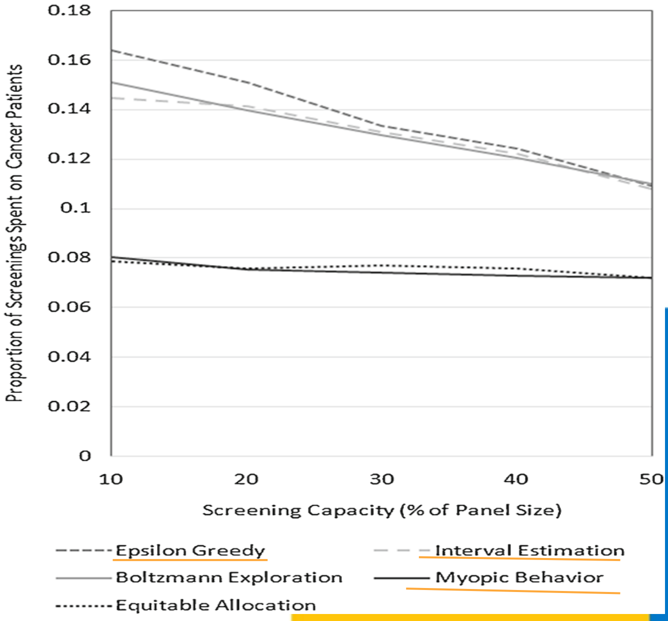
Resource constraint level k/n	ϵ-greedy ϵ	Interval estimation z	Boltzmann exploration τ
10 %	0.025	1	0.250
20 %	0.05	1	0.250
30 %	0.10	1	0.300
40 %	0.10	1	0.325
50 %	0.25	5	0.400

Paper Results

Proportion of Cancers Detected in Early Stage



Proportion of Screenings Spent on Cancer Patients



Ryerson University

Data Science Laboratory

My Replication Results

Challenges and Performance Measure Report

Ryerson University

Challenges

- First attempt: my results Dynamic AFP values should follow a specific RR and SS that is different for each subgroup (HCC and non-HCC)

- However, my AFP values are randomly generated ➔ my results did not match paper

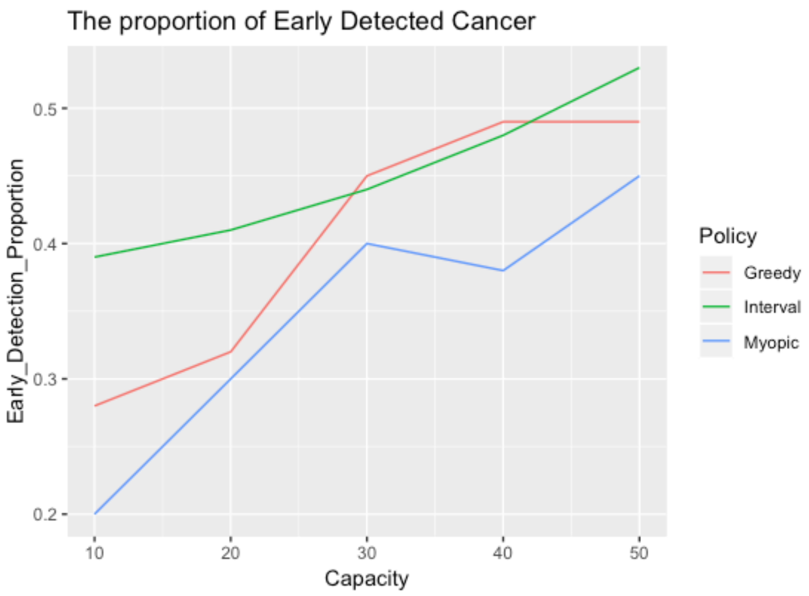
- Approach: instead of generating AFP values, I generate SS and RR values

Table 1 Characteristics of patients with and without HCC

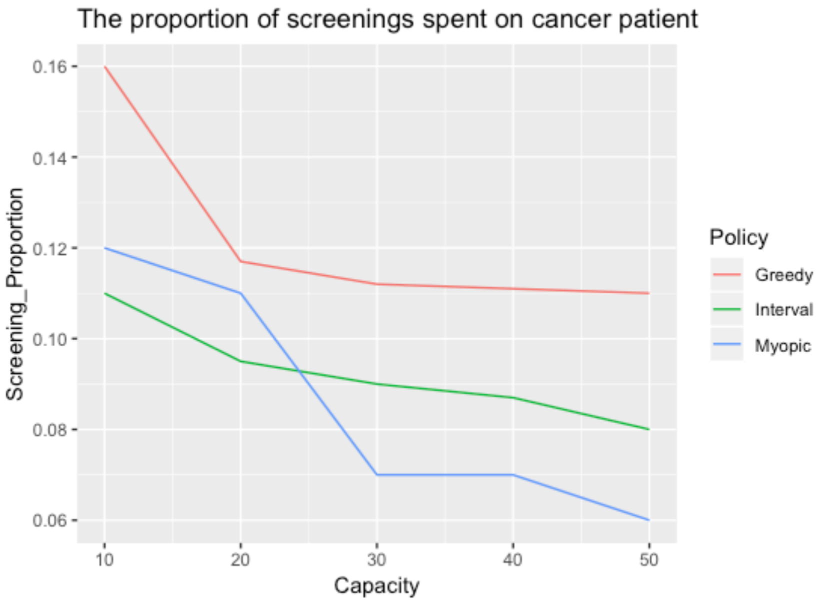
Characteristic	HCC (N = 82)	NO HCC (N = 885)	*P* Value
Age at Baseline (years)	53±7	50±7	< 0.01
Black (binary)	24 %	18 %	0.10
Platelets at Baseline $\times 1000/mm^3$	126±51	169±65	< 0.01
Ever Smoked (binary)	41 %	24 %	< 0.01
Alkaline Phosphatase at Baseline (U/L)	117±59	97±43	< 0.01
Esophageal Varices (binary)	4 %	34 %	0.01
Standard Deviation of AFP (ng/mL)	51 ± 86	9 ± 19	< 0.01
Rate of AFP Rise (90*ng/mL)	5±11	0.11±2.1	< 0.01

My Results

Proportion of Cancers Detected in Early Stage

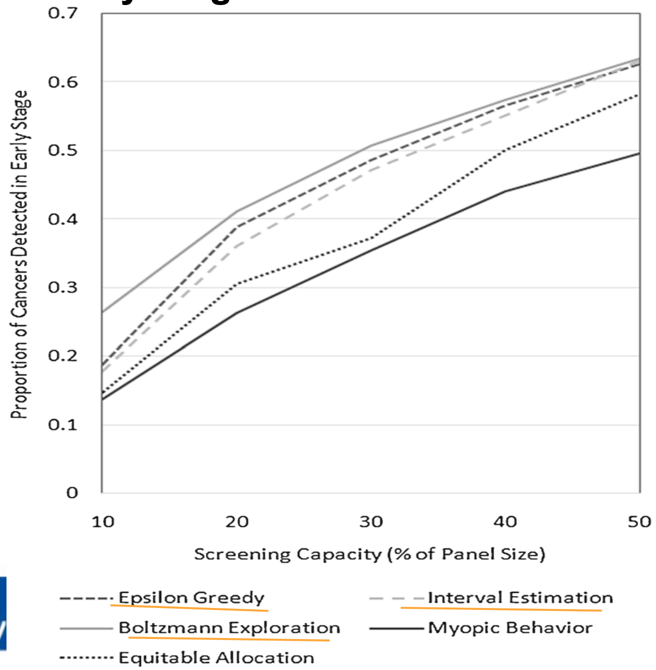


Proportion of Screenings Spent on Cancer Patients

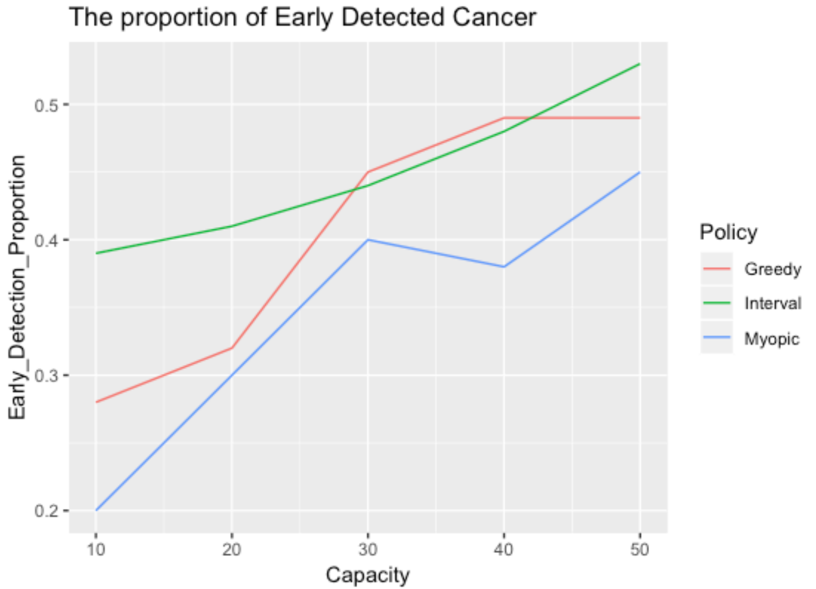


Paper Results Vs. My Results

Proportion of Cancers Detected in Early Stage

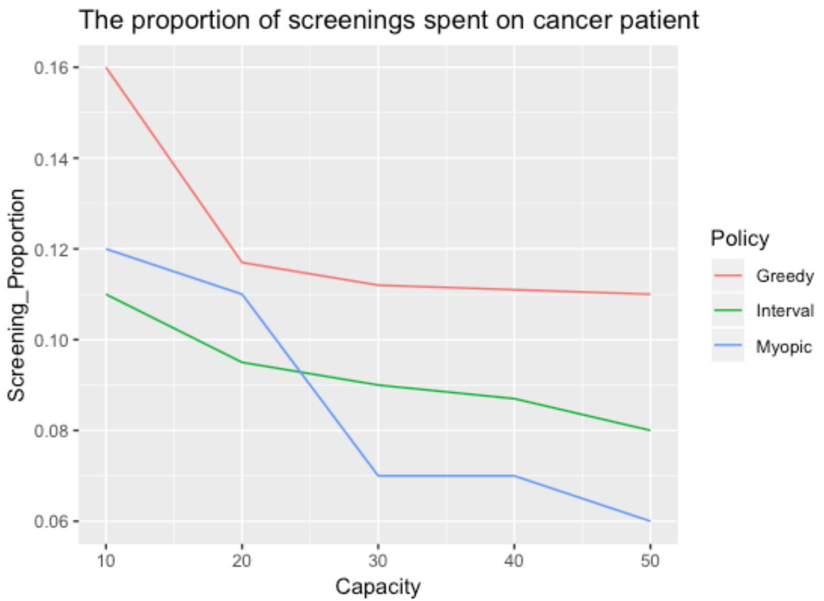


Proportion of Cancers Detected in Early Stage

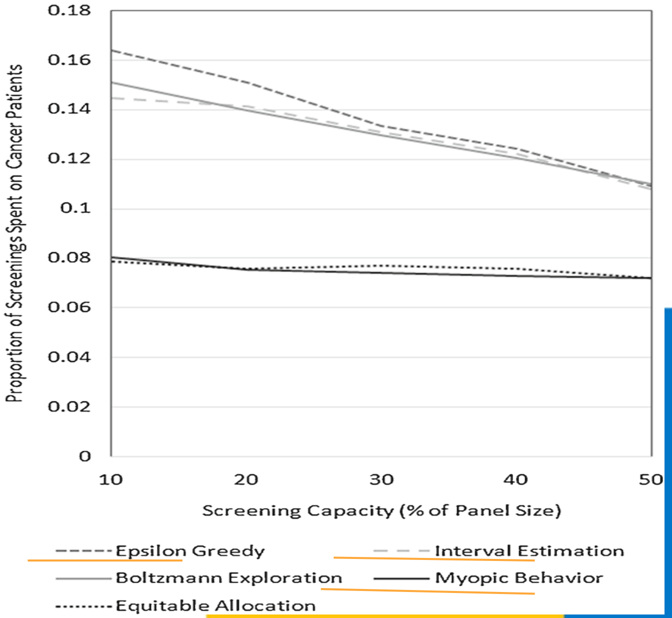


Paper Results Vs. My Results

Proportion of Screenings Spent on Cancer Patients



Proportion of Screenings Spent on Cancer Patients



Possible Reasons Why Results May Differ

- Parameters I used are based on the best parameters reported in the paper
- Other parameters may be more well suited to our synthetic dataset and may result in a more accurate performance

Table 2 Optimal tuning parameters determined by the indifference zone method

Resource constraint level k/n	ϵ-greedy ϵ	Interval estimation z	Boltzmann exploration τ
10 %	0.025	1	0.250
20 %	0.05	1	0.250
30 %	0.10	1	0.300
40 %	0.10	1	0.325
50 %	0.25	5	0.400

Thank you!

Questions?

Ryerson University

Data Science Laboratory

References

[1] Lee, Elliot, et al. "Applying reinforcement learning techniques to detect hepatocellular carcinoma under limited screening capacity." *Health care management science* 18.3 (2015): 363-375.

[2] Lee, Elliot, et al. "Improving screening for hepatocellular carcinoma by incorporating data on levels of α-fetoprotein, over time." *Clinical Gastroenterology and Hepatology* 11.4 (2013): 437-440.

[3] O'Connor, S., et al. "Hepatocellular carcinoma-United States, 2001-2006." *Morbidity and Mortality Weekly Report* 59.17 (2010): 517-520.

[4] Okada, S., et al. "Follow-up examination schedule of postoperative HCC patients based on tumor volume doubling time." *Hepato-gastroenterology* 40.4 (1993): 311-315.

Ryerson University

Data Science Laboratory