Exploring Regional Variation and Survival in Colon Cancer Pathway Concordance

Research Roundtable: Data Analytics in Healthcare

Rotman School of Management
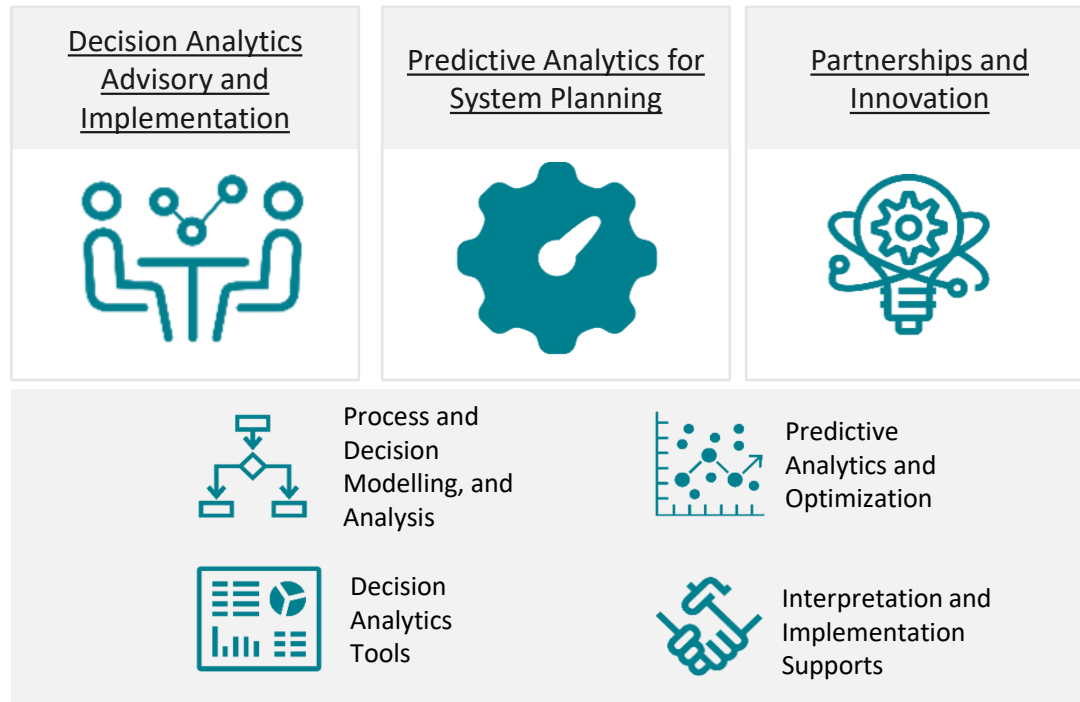
March 9, 2020

Luciano Ieraci, Maria Eberg

Data and Decision Sciences, Ontario Health (CCO)

Ontario Health
Cancer Care Ontario

D+DS Provides Advanced Analytics Competency for Health System Management

Decision Analytics Advisory and Implementation	Predictive Analytics for System Planning	Partnerships and Innovation

- Process and Decision Modelling, and Analysis
- Decision Analytics Tools
- Predictive Analytics and Optimization
- Interpretation and Implementation Supports

Ontario Health
Cancer Care Ontario

Background to Disease Pathway Measurement

What Are Pathway Maps?

Flowcharts that describe the care cancer patients should receive based on **best available evidence**

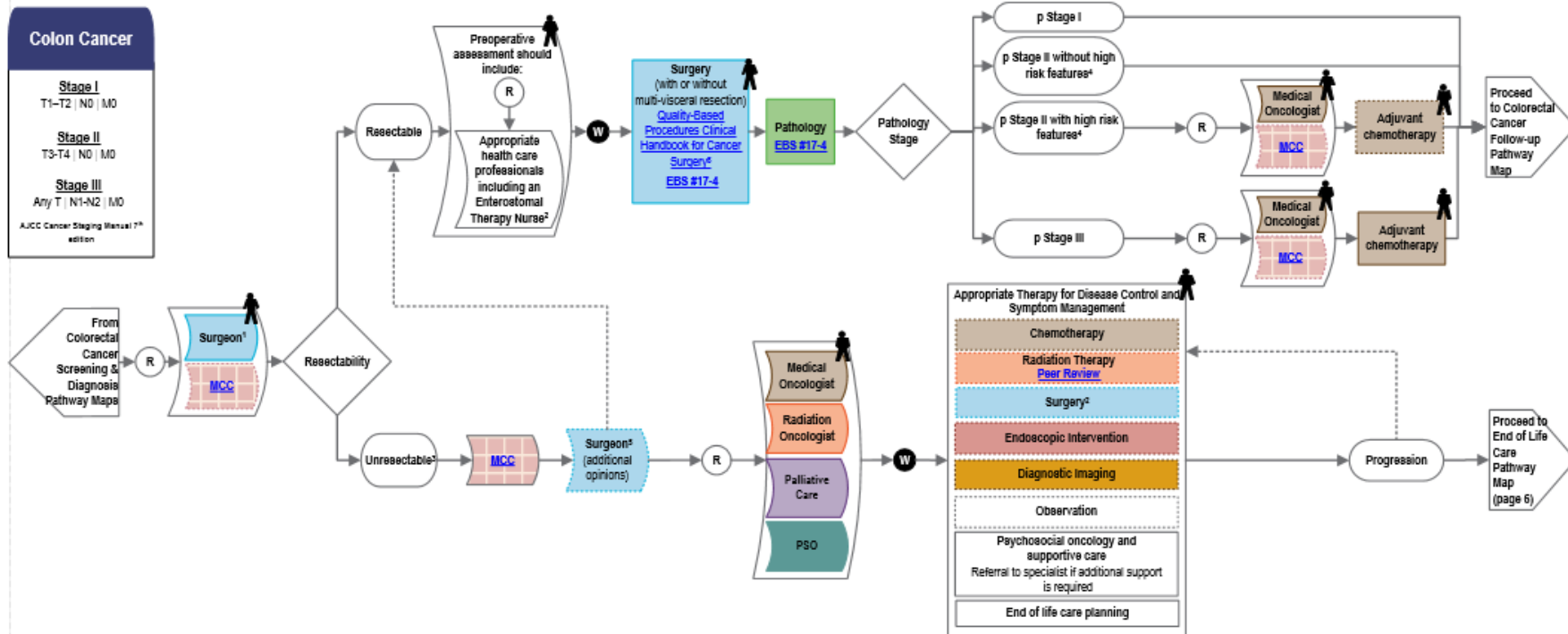Quality improvement tool for the **Ontario cancer system**

Population based, organized by cancer type (e.g. lung cancer) and span the entire care continuum

Sequences of interventions with care teams, specialties, programs – *does not include specific protocols, regimens, costs, etc.*

Ontario Health
Cancer Care Ontario

Example – Colorectal Cancer Pathway Map

Why Do We Develop Pathway Maps?

GOAL

- Facilitate improvements in quality, access, appropriateness, and coordination of care across the cancer continuum.

- Set care expectations for cancer patients in Ontario, based on best scientific evidence

- Reduce undesired variation in care

- Promote discussion and collaboration between care providers, health administrators, system planners, and educators

Ontario Health
Cancer Care Ontario

Introduction to Pathway Concordance Measurement

Pathway Concordance Objectives

- Quantifies agreement between actual care received and care recommended by the pathway map

Are we following the ideal path?

What, where & why are there deviations?

Ontario Health
Cancer Care Ontario

Pathway Concordance Objectives

1. Develop methodology to measuring pathway concordance using CCO's current data holding at a population and system level. Focus at the system level or specific section of the pathway. These methods do not exist currently.

2. Create capability **to link to outcomes of interest** to monitor system performance
- Clinical outcomes – survival, recurrence, patient-reported outcomes
- Health system costs

3. Establish metric for ongoing **evaluation** and **performance measurement, system monitoring**
- Identify deviations from pathway map and effects on outcomes
- Support quality initiatives and performance management

Ontario Health
Cancer Care Ontario

Approach and Methods

# Population and Data Sources

Cohort

- Ontario Cancer Registry
    - Accrual from Jan 1, 2012 to Dec 31, 2016
    - Follow-up until death or Mar 31, 2019

- Incident stage II and III colon cancer
    - Pathologically-confirmed

- Exclusions:
    - Non-Ontario resident / invalid health card
    - Had more than one primary tumour
    - Evidence of non-curative treatment before surgery

Health services

- Radiation, chemotherapy (Activity Level Reporting, Ontario Drug Benefit population)

- Physician services (Ontario Claims History Database)

- Hospital services (National Ambulatory Care Reporting System , Discharge Abstract Database)

Pathway – Colon Cancer

Resectable, curative pathway only
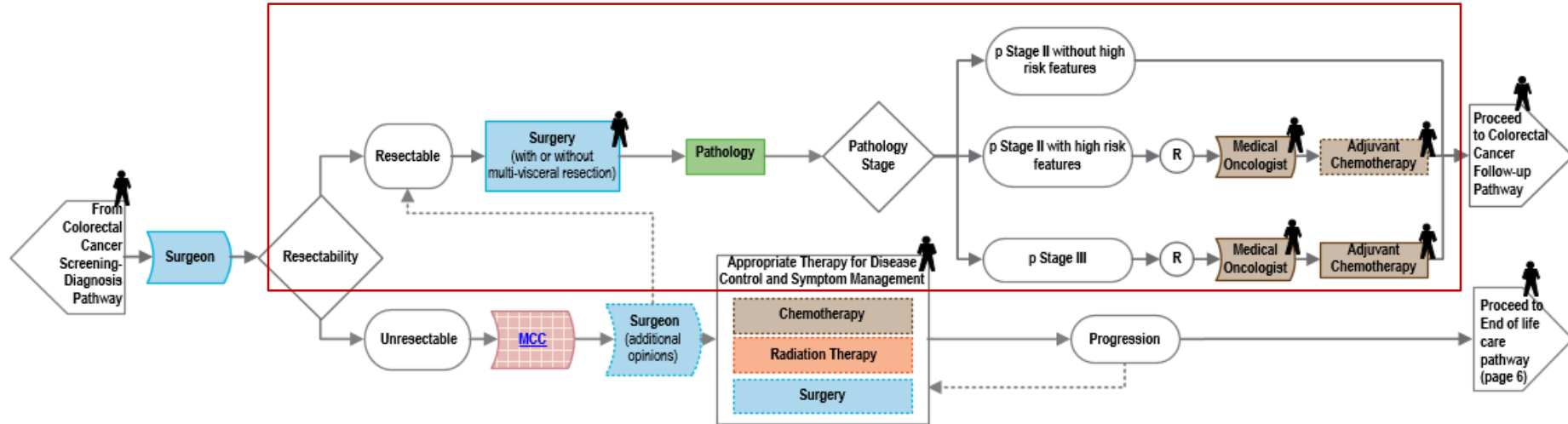
Metric Development

Simplified Reference Pathways

- Stages IIA, IIB/C correspond to average and higher risk of cancer recurrence, respectively

Colon Cancer Stage II	Colon Cancer Stage III
Stage IIA (A or B) – C – D – (E optional)	(A or B) – C – D – E – F
Stage IIB / IIC (A or B) – C – D – E – (F optional)	

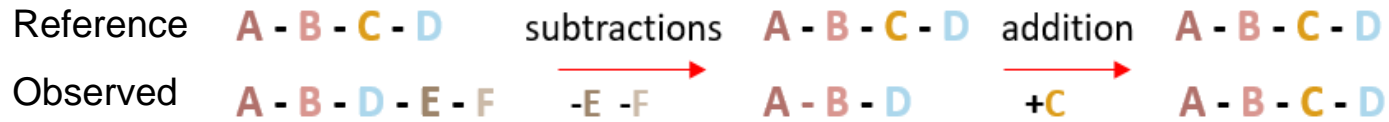A or B – Endoscopy full or partial
C – Imaging: abdominal CT, pelvis CT, chest imaging
D – Surgical resection
E – Oncology consultation
F – Chemotherapy

Levenshtein Distance Metric

- Calculate Levenshtein distance metric for each subject

Therefore Levenshtein distance = # Operations = 3

- **Levenshtein** counts the number of "edits" required to transform observed pathway: counts the number of additional (edit = "insert") and missing (edit = "delete") events along pathway

- Normalized Levenshtein used to represent "percent similarity" (higher scores are concordant)

$$d' = 1 - \frac{d(x,y)}{max(x,y)}$$

  - where $d(x,y)$ is the Levenshtein distance and $max(x,y)$ is the maximum path length

Pathway Concordance and Survival

Objective and Methods

Objective

- Validate the developed concordance score by investigating its association with patient survival at the population level

Crude analyses

o Kaplan-Meier survival curves by *tercile groups*

o Common edit operations / sources of discordance – *lowest tercile group*

Adjusted analyses

o Extension of Cox proportional hazards model for time-dependent variables

o Time-dependent covariates:

Concordance score, unplanned ED visits, chemotherapy treatments

o Baseline covariates:

Socio-demographic variables, history of health services utilization, comorbidity history, variables related to cancer and its treatment

**Ontario Health**
Cancer Care Ontario

Additional / Missing Events

| Event type	Stage II – low concordance tercile (n = 1,646)	Stage III – low concordance tercile (n = 1,782)
Additional activity		
Additional imaging test, n (%)	1,642 (99.8%)	1,779 (99.8%)
Abdomen CT, median (IQR)	2 (1 – 3)	3 (2 – 4)
Chest CT, median (IQR)	1 (1 – 2)	2 (1 – 3)
Chest X-ray, median (IQR)	4 (2 – 7)	4 (2 – 7)
Additional consultation, n(%)	1,531 (93.0%)	1,624 (91.1%)
Medical oncologist, median (IQR)	2 (1 – 3)	2 (1 – 3)
Surgical oncologist, median (IQR)	2 (1 – 2)	2 (1 – 2)
Additional endoscopy, n (%)	704 (42.8%)	782 (43.9%)
Additional radiation treatment, n (%)	0 (0%)	385 (21.6%)
Missing activity		
Missing endoscopy, n (%)	416 (25.3%)	507 (28.5%)
Missing imaging test, n (%)	186 (11.3%)	232 (13.0%)
Missing chemotherapy, n (%)	N/A	684 (38.4%)

Survival by Concordance Tercile

- Stage II – Kaplan-Meier survival curves
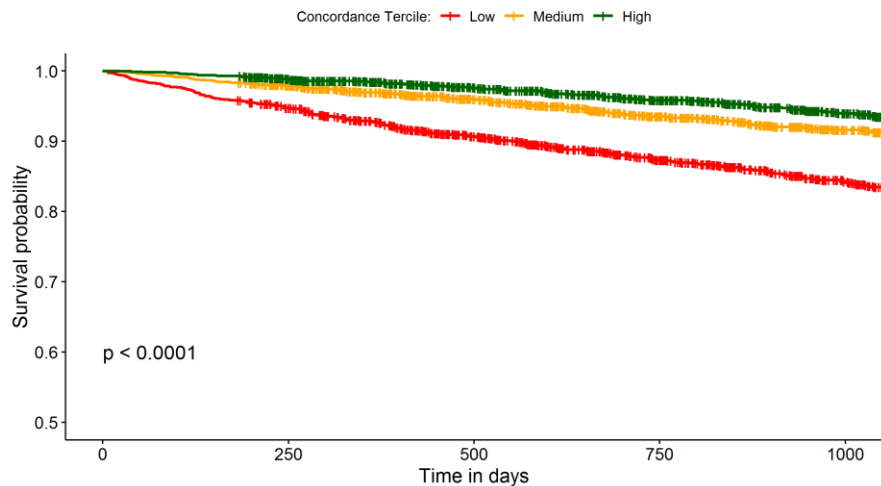
- Stage III – Kaplan-Meier survival curves
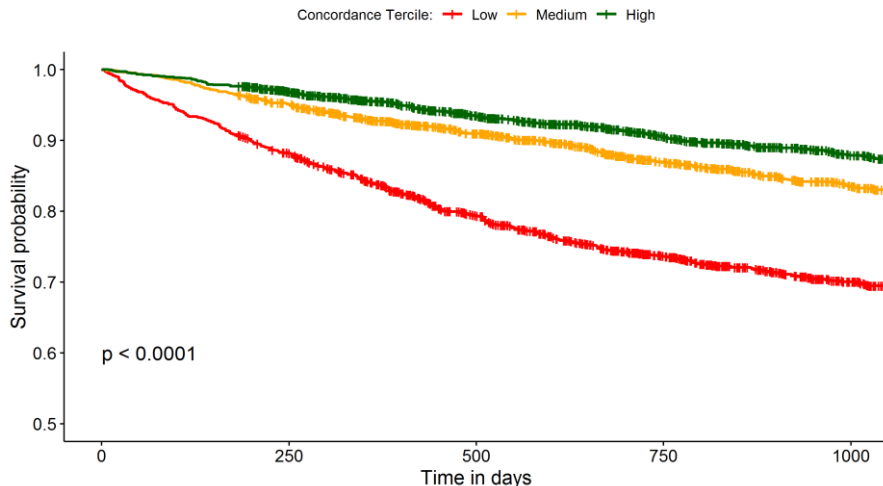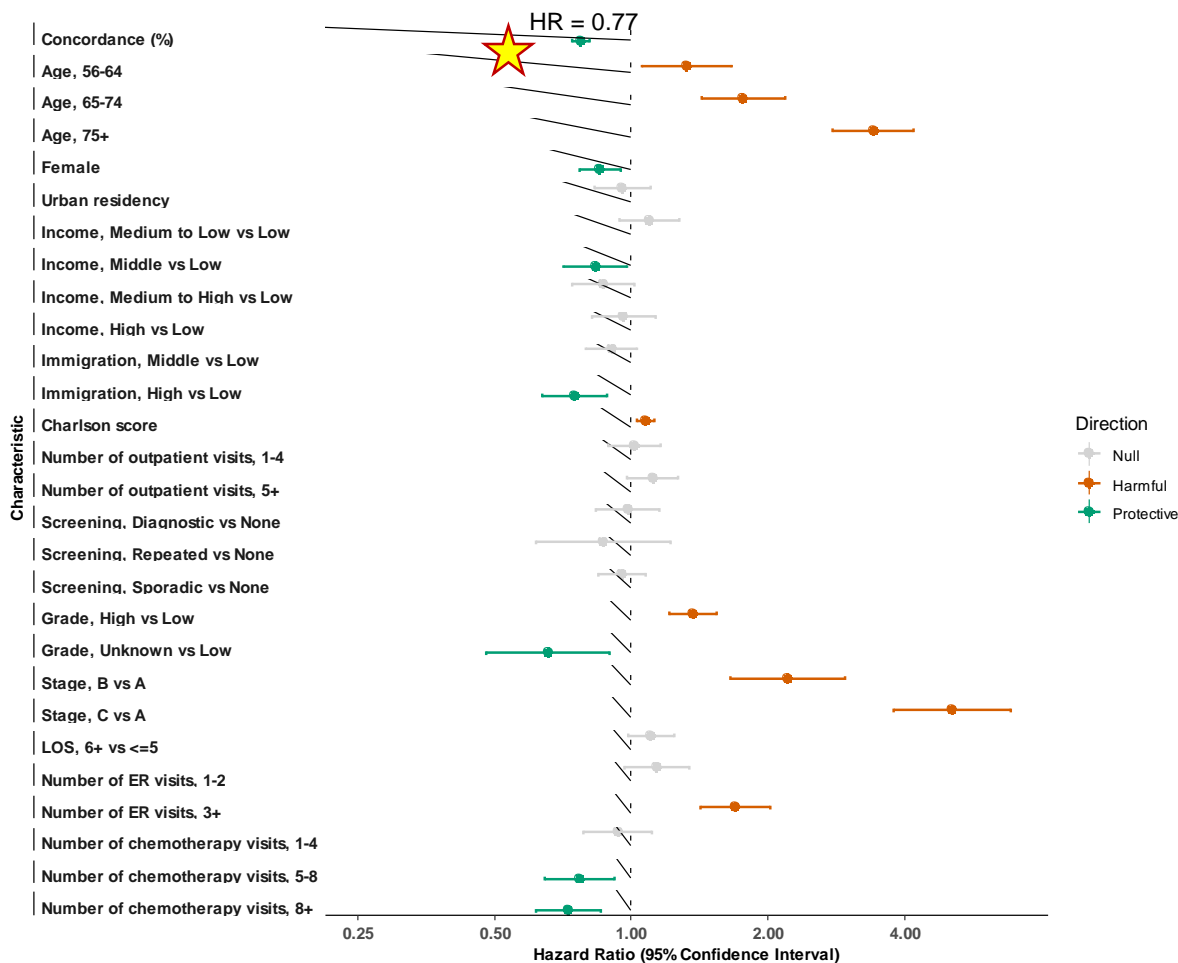
Stage II – Survival Analysis Results
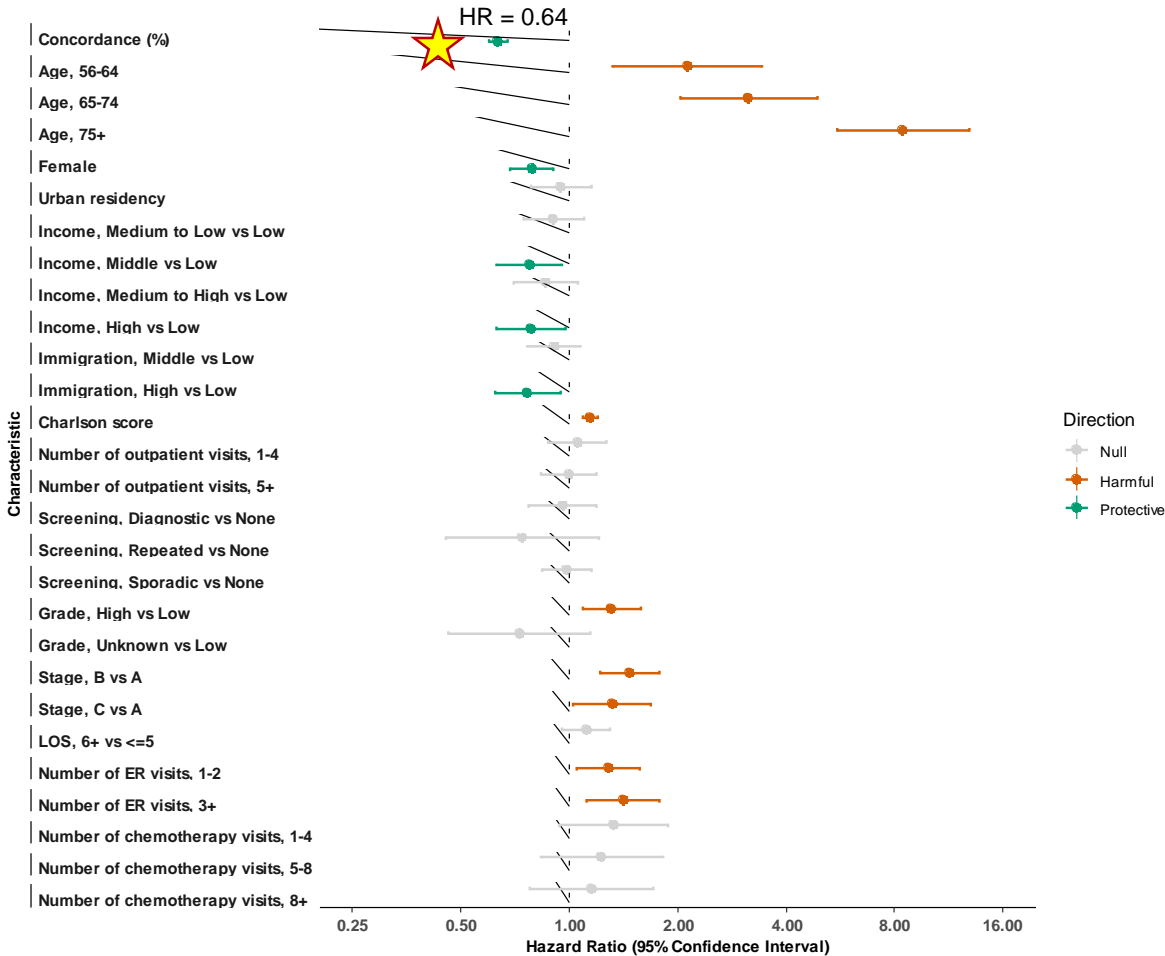
Higher rate of death:

- Older age
- Charlson score
- Higher stage
- High tumor grade
- ED visits

Lower rate of death:

- Female
- Highest immigrant population
- **Pathway concordance**

18

Stage III – Survival Analysis Results

Clustering and Segmentation Analysis

Deriving Actionable Insights

How can we use concordance measurement to inform quality improvement?
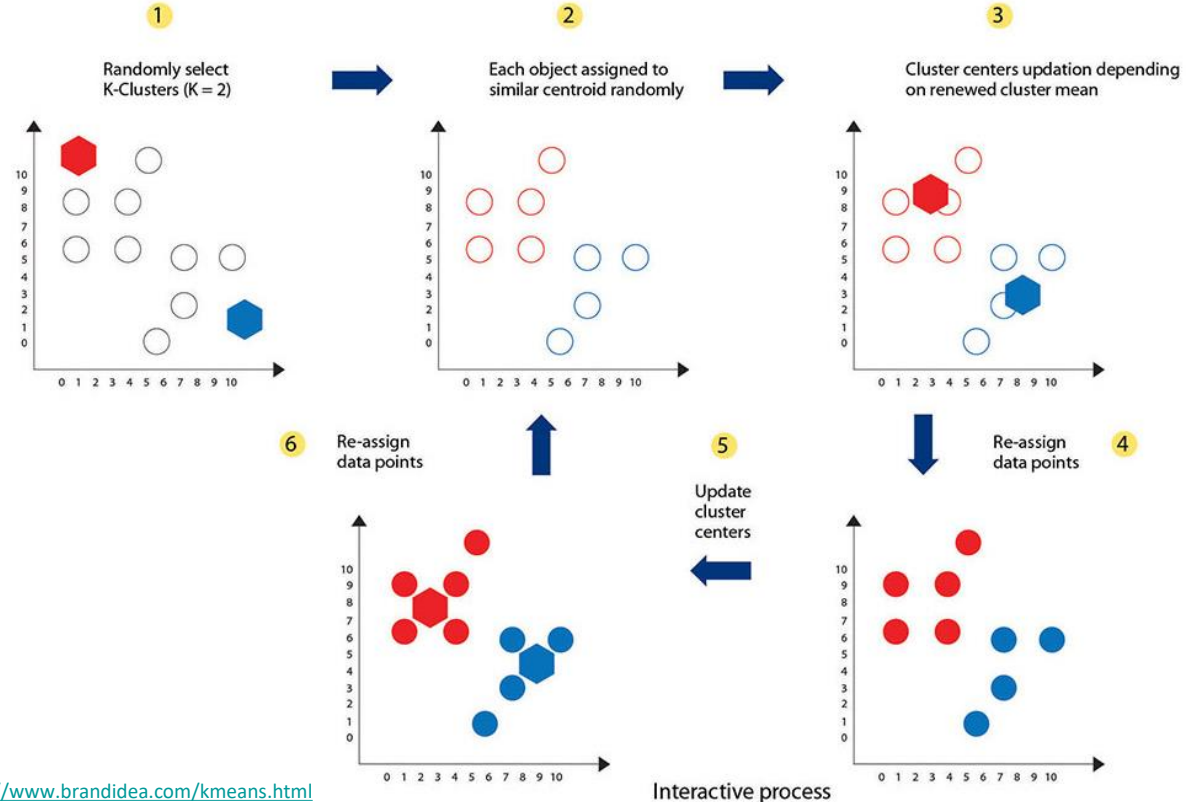
Objectives

- Find subpopulations of patients with "modifiable" factors to inform future quality improvement efforts

- Use unsupervised learning algorithms such as clustering to identify heterogeneous groups within the population of colon cancer patients

- Assess the additional benefit of incorporating a concordance metric in the analysis

K-Means Clustering

Methods

- Partition data into distinct groups so that the observations within each group are similar, while observations in other groups are different

- Similarity is measured by squared Euclidian distance:

$$\frac{1}{|C_K|} \sum_{i,i' \in C_K} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2$$

1 Randomly select K-Clusters (K = 2)

2 Each object assigned to similar centroid randomly

3 Cluster centers updation depending on renewed cluster mean

6 Re-assign data points

5 Update cluster centers

4 Re-assign data points

Interactive process

Clustering Analysis – What Did We Find?
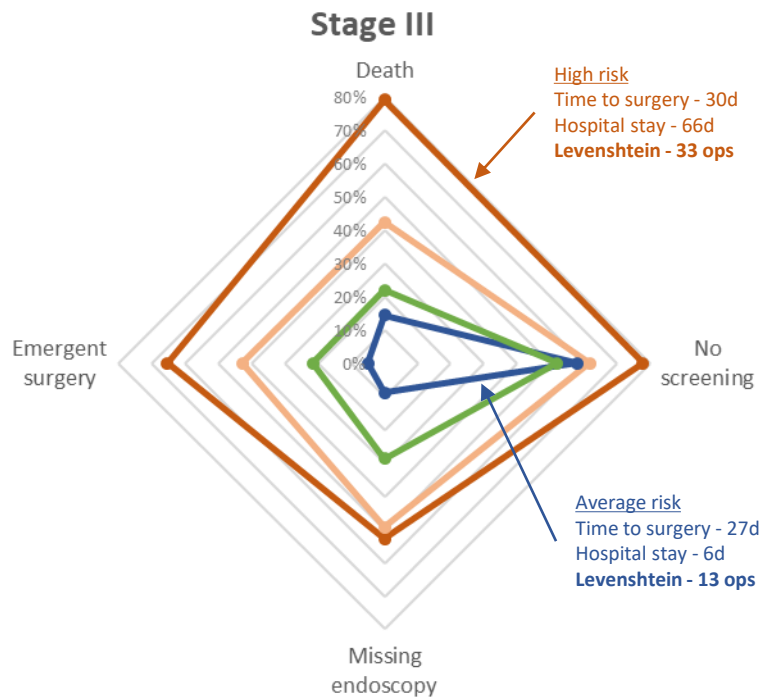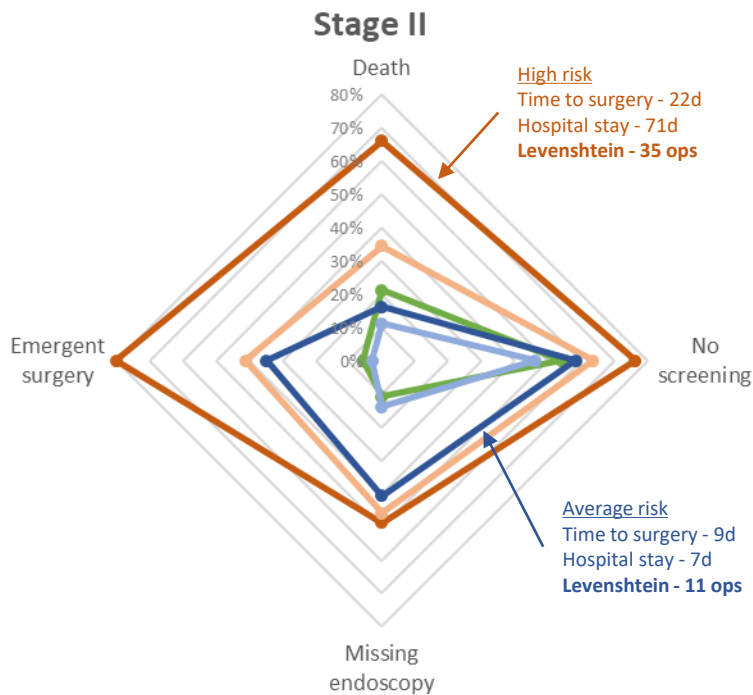
Variables used for analysis

- Time from diagnosis date to surgical resection
- Length of hospital stay
- Levenshtein concordance metric (un-normalized)

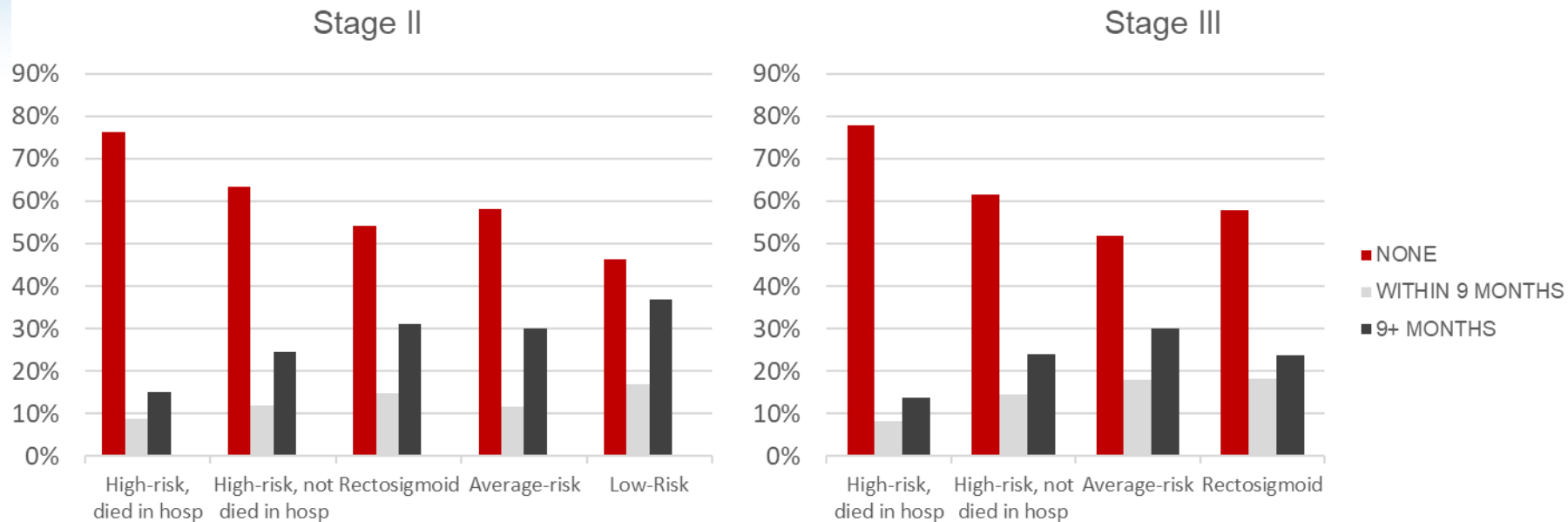**Found clusters to target**

- High mortality with variation in time-to-events
- Identified subgroups with high prevalence of radiation treatment (cancer in rectosigmoid junction)
- Identified lack of "early identification" activity (screening, endoscopy)

Adding the Levenshtein metric resulted in refined data segmentation

Clusters Profiles

Stage II

- Death
- No screening
- Missing endoscopy
- Emergent surgery

High risk
Time to surgery - 22d
Hospital stay - 71d
Levenshtein - 35 ops

Average risk
Time to surgery - 9d
Hospital stay - 7d
Levenshtein - 11 ops

Stage III

- Death
- No screening
- Missing endoscopy
- Emergent surgery

High risk
Time to surgery - 30d
Hospital stay - 66d
Levenshtein - 33 ops

Average risk
Time to surgery - 27d
Hospital stay - 6d
Levenshtein - 13 ops

Legend:
- High-risk, died in hosp
- High-risk, not died in hosp
- Rectosigmoid
- Average-risk
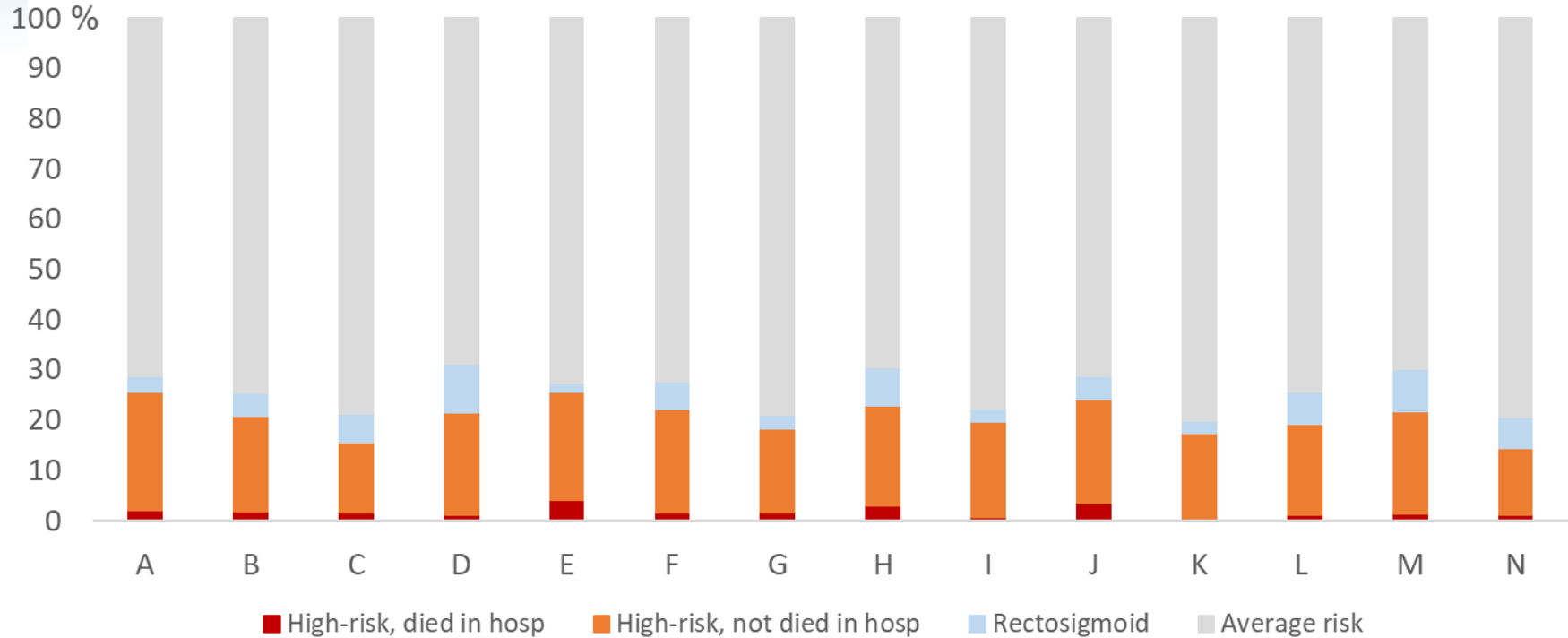- Low-Risk

24

Cancer Screening by Cluster Groups

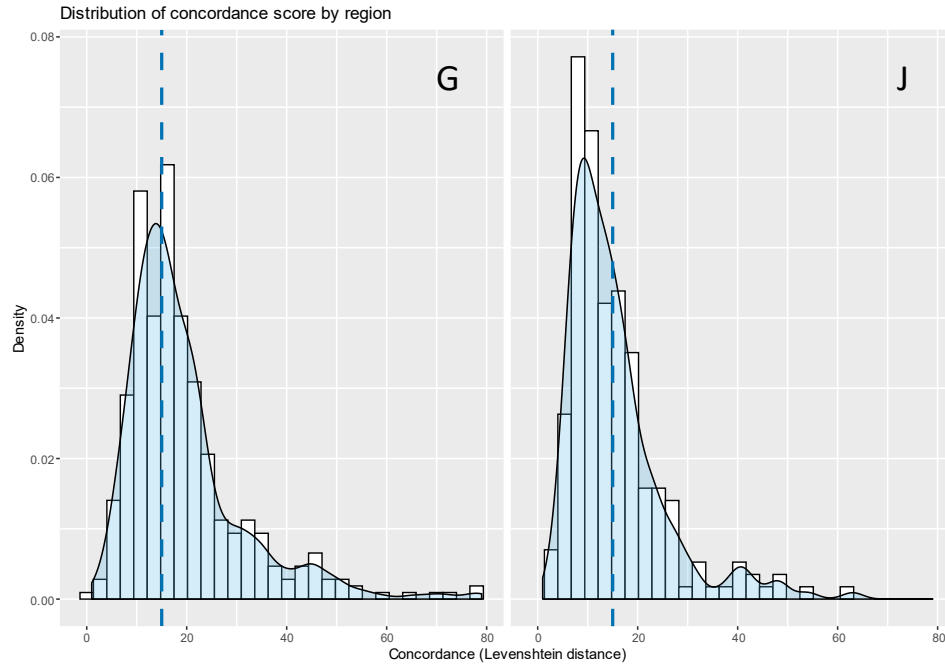

High-mortality groups more likely to have no screening.

Regional Analysis

Stage III – Regional Variation in Cluster Volumes

Stage III – Regional Variation in Concordance

Distribution of concordance score by region

Region	% Above provincial median
G	61.0%
J	41.0%

Ontario Health
Cancer Care Ontario

Conclusion

Conclusion

- Demonstrated feasibility of measuring pathway concordance across the cancer continuum at a population based level

- Novel application of measurement shows promising results in demonstrating:

  - Regional and practice variation

  - Association with improved survival

- Approach could be used to identify opportunities for quality improvement and measuring system performance

Ontario Health
Cancer Care Ontario

Next Steps

Evaluation of metrics

- To determine the optimal metric – Ontario Health (CCO) perspective

- Currently drafting multiple criteria to rank metrics with expert group(s)

- Goal is to endorse one metric, or multiple metrics based on strengths/weaknesses

Extension to breast cancer pathways

- Developing metrics for breast cancer pathway concordance analysis

- Continuing to work with the University of Toronto (MIE) to define reference pathways and assess metric performance

Ontario Health
Cancer Care Ontario

Acknowledgements

Questions / Discussion