Research Projects in Secondary Cancer Prevention

Yan Yuan, Assistant Professor, SPH

Presentation for CRINA

July 16, 2015

UNIVERSITY OF ALBERTA

Brief Biography

Yan Yuan

- 2011-present, Assistant Prof; School of Public Health, University of Alberta.
- 2008-2011, Biostatistician, Population Health Research, Cancer Control, Alberta Health Services
- 2003 MMath-Biostatistics, 2008 PhD-Statistics, U of Waterloo
- 1999-2001 Lab manager in an Animal behavior lab, University of Guelph, Canada
- 1999, MSc in Animal Behavior, Michigan State University, USA
- 1996 BSc in Biochemistry, Nanjing University, China

Outline

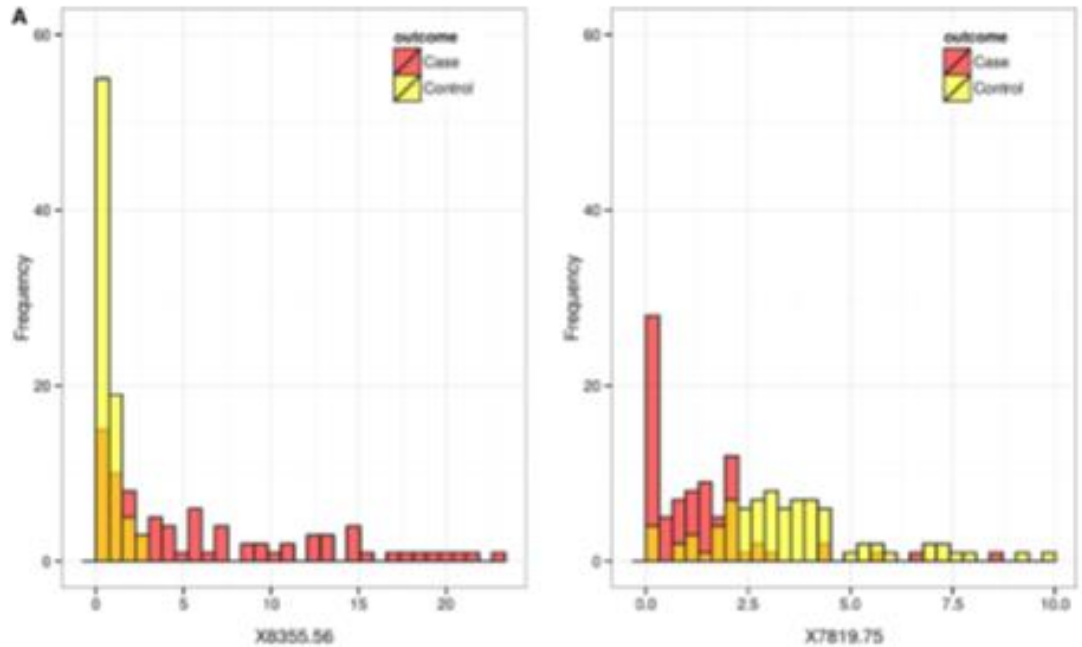
1. Predicting/Detecting the **Rare** Events such as cancer

10-year cancer diagnosis per 1000 person	Colorectal cancer		Breast cancer	Prostate cancer
	Male	Female		
Age 50	6.8	5.2	23	22
Age 60	13	9	35	63

2. Secondary Cancer Prevention – health services research

1.1 Motivating Data

779 potential biomarkers were assessed in 83 late-stage prostate cancer patients and 82 normal subjects. (Adam *et al*. 2002 *Cancer Research*)



1.2 Predicting the Rare Events

- Cancer screening
- Risk prediction – adverse birth outcomes, diabetes, cancer, cardiovascular disease etc.

1.3 Evaluating Prediction Performance for Rare Events

- Threshold Dependent Measure
  - Misclassification rate
  - Sensitivity and Specificity
  - Positive and Negative Predictive Value
- Threshold Independent Measure (Pre-clinical or pre-application stage)
  - Area Under the Receiver Operating Characteristic Curve (AUC or aROC)
  - **Average Positive Predictive Value (AP)**

Score	x_1	$>$	x_2	$> \cdots >$	x_k	$>$	x_{k+1}	$> \cdots >$	x_K	
Partition	R_1		R_2	$\cdots$	R_k		R_{k+1}	$\cdots$	R_K	Total
Class-1	Z_1		Z_2	$\cdots$	Z_k		Z_{k+1}	$\cdots$	Z_K	n_1
Class-0	$\bar{Z}_1$		$\bar{Z}_2$	$\cdots$	$\bar{Z}_k$		$\bar{Z}_{k+1}$	$\cdots$	$\bar{Z}_K$	n_0
Total	S_1		S_2	$\cdots$	S_k		S_{k+1}	$\cdots$	S_K	n

$$\widehat{AP} = \underbrace{\left[\frac{Z_1}{S_1}\right]}_{w_1}\left[\frac{Z_1}{n_1}\right] + \underbrace{\left[\frac{Z_1+Z_2}{S_1+S_2}\right]}_{w_2}\left[\frac{Z_2}{n_1}\right] + \cdots + \underbrace{\left[\frac{Z_1+Z_2+\cdots+Z_K}{S_1+S_2+\cdots+S_K}\right]}_{w_K}\left[\frac{Z_K}{n_1}\right]$$

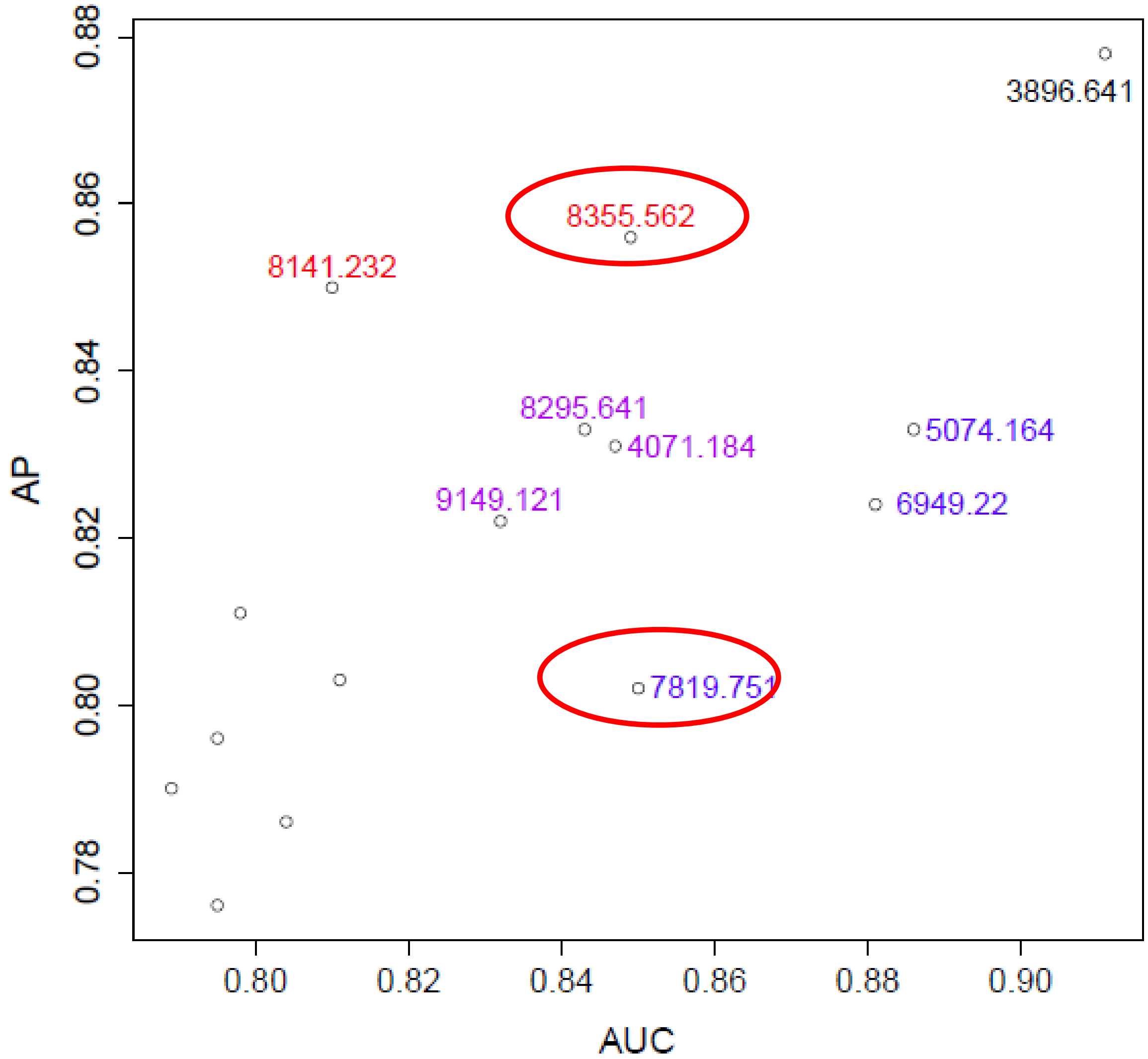
$$= \sum_{k=1} w_k \left[\frac{Z_k}{n_1}\right].$$

$$\widehat{AUC} = \frac{n}{n_0}\left\{\underbrace{\left[\frac{S_1+S_2+\ldots+S_K}{n}\right]}_{w_1'}\left[\frac{Z_1}{n_1}\right] + \underbrace{\left[\frac{S_2+\ldots+S_K}{n}\right]}_{w_2'}\left[\frac{Z_2}{n_1}\right] + \ldots + \underbrace{\left[\frac{S_K}{n}\right]}_{w_K'}\left[\frac{Z_K}{n_1}\right] - \frac{1}{2}\left(\frac{n_1}{n_0}\right)\right\} - \frac{1}{2}\left(\frac{n_1}{n_0}\right)$$

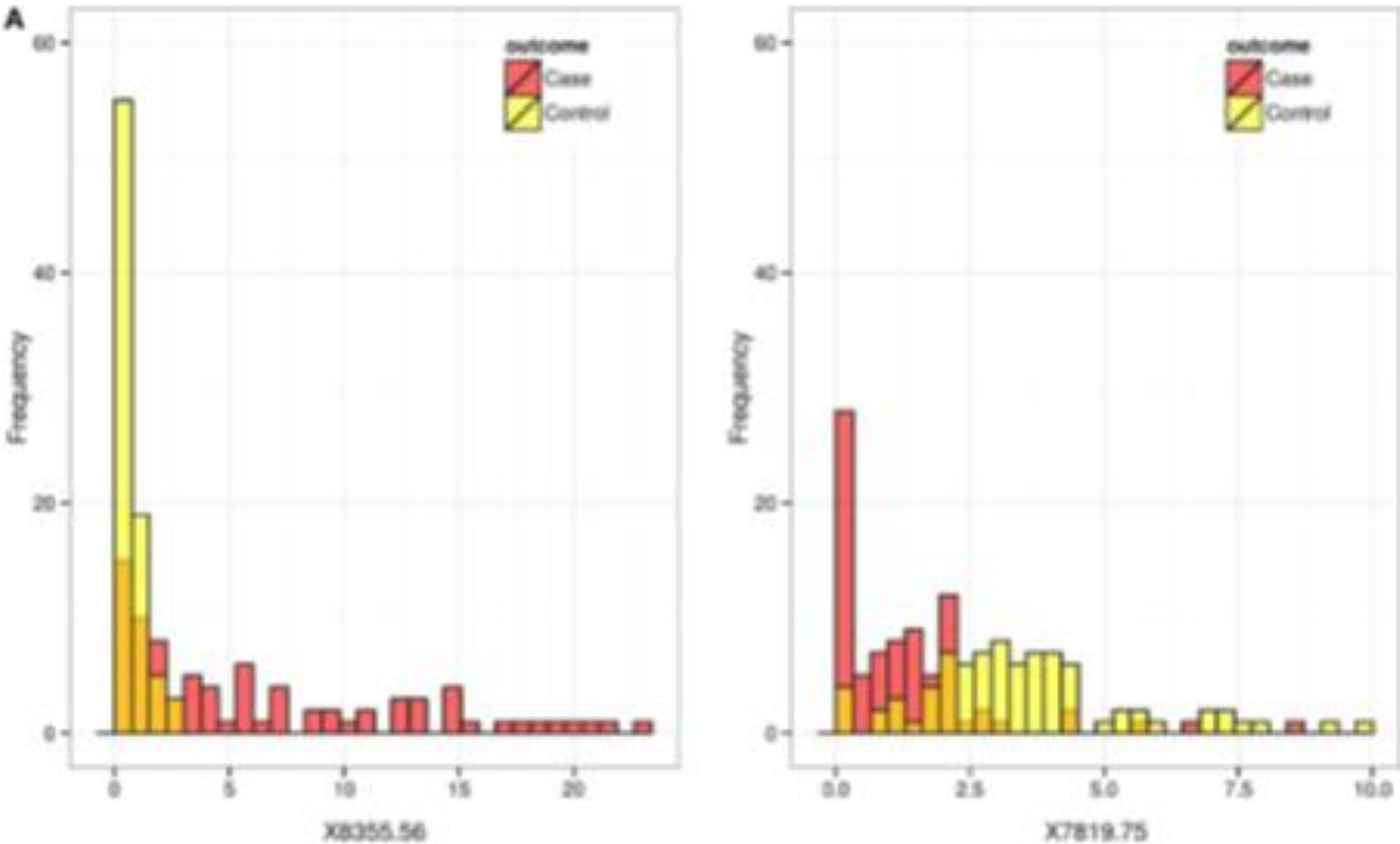
$$= \frac{n}{n_0}\sum_{k=1} w_k'\left[\frac{Z_k}{n_1}\right] - \frac{1}{2}\left(\frac{n_1}{n_0}\right)$$

Example A: Biomarkers for prostate cancer screening

779 potential biomarkers were assessed in 83 late-stage prostate cancer patients and 82 normal subjects. (Adam *et al*. 2002 *Cancer Research*)



Example A: Two biomarker similar on AUC scale for prostate cancer screening



1.4 An Experiment and Results

- The biomarker study is based on a case-control study (# disease ≈ # non-disease); its goal is to identify potential screening markers.
- How AP and the ranking of biomarkers is affected when the incidence is much lower as in a screening setting?

Inflate the controls by replicating them

Biomarker	AUC	AP		
	n_0 X 1 $\pi = 0.5$	n_0 X 1 $\pi = 0.5$	n_0 X 10 $\pi = 0.1$	n_0 X 100 $\pi = 0.01$
8355.562	**0.849**	0.856	0.606	0.571
7819.751	**0.850**	0.802	0.370	0.062

Example B: Two technology for Breast cancer screening

42,760 screening participants underwent two screening technology, 335 were diagnosed with breast cancer at 15 months follow-up. (Pisano et al. 2005 *New England Journal of Medicine*)

Malignancy score		**7**	**6**	**5**	**4**	**3**	**2**	**1**	**Total**
Digital M	Category Total	11	29	69	1061	2224	6588	32588	**42570**
	Cancers	10	18	25	85	49	25	122	**334**
Film M	Category Total	17	29	70	942	2291	6910	32486	**42745**
	Cancers	13	24	25	74	35	33	131	**335**

1.4B Results

Given that 335 breast cancer diagnosed in 42,760 screening participants at 15 months follow-up, the cumulative incidence π is 0.783%.

	Seven-point Malignancy Scale	
	$\widehat{AUC}$ (s.e.)	$\widehat{AP}$ (s.e.)
Film mammography	0.735 (0.012)	0.166 (0.022)
Digital mammography	0.753 (0.012)	0.144 (0.021)

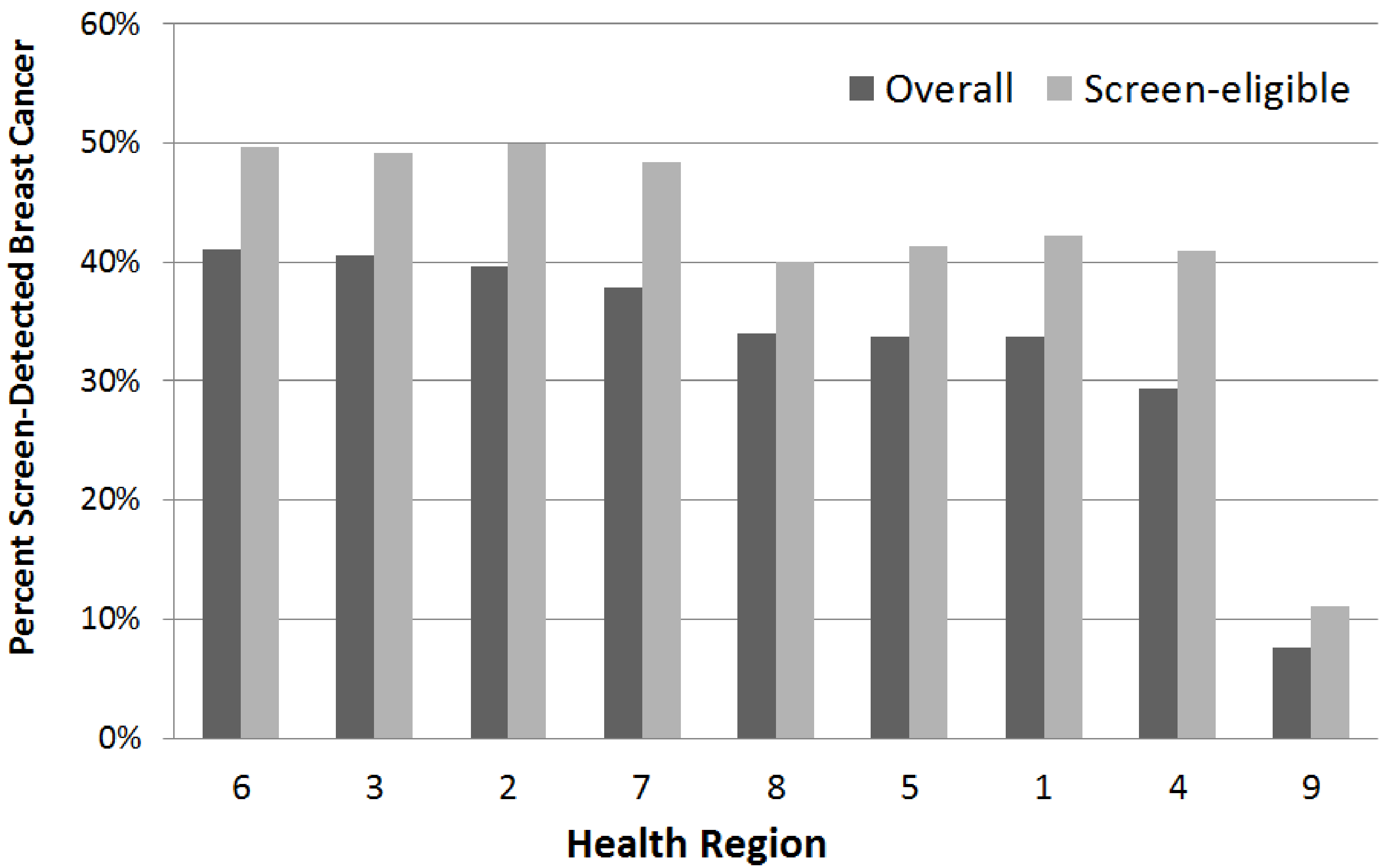
Remark: Resampling method can be used for the inference of the difference in AP when we have paired data.

2. Breast Cancer Diagnostic Care in Alberta

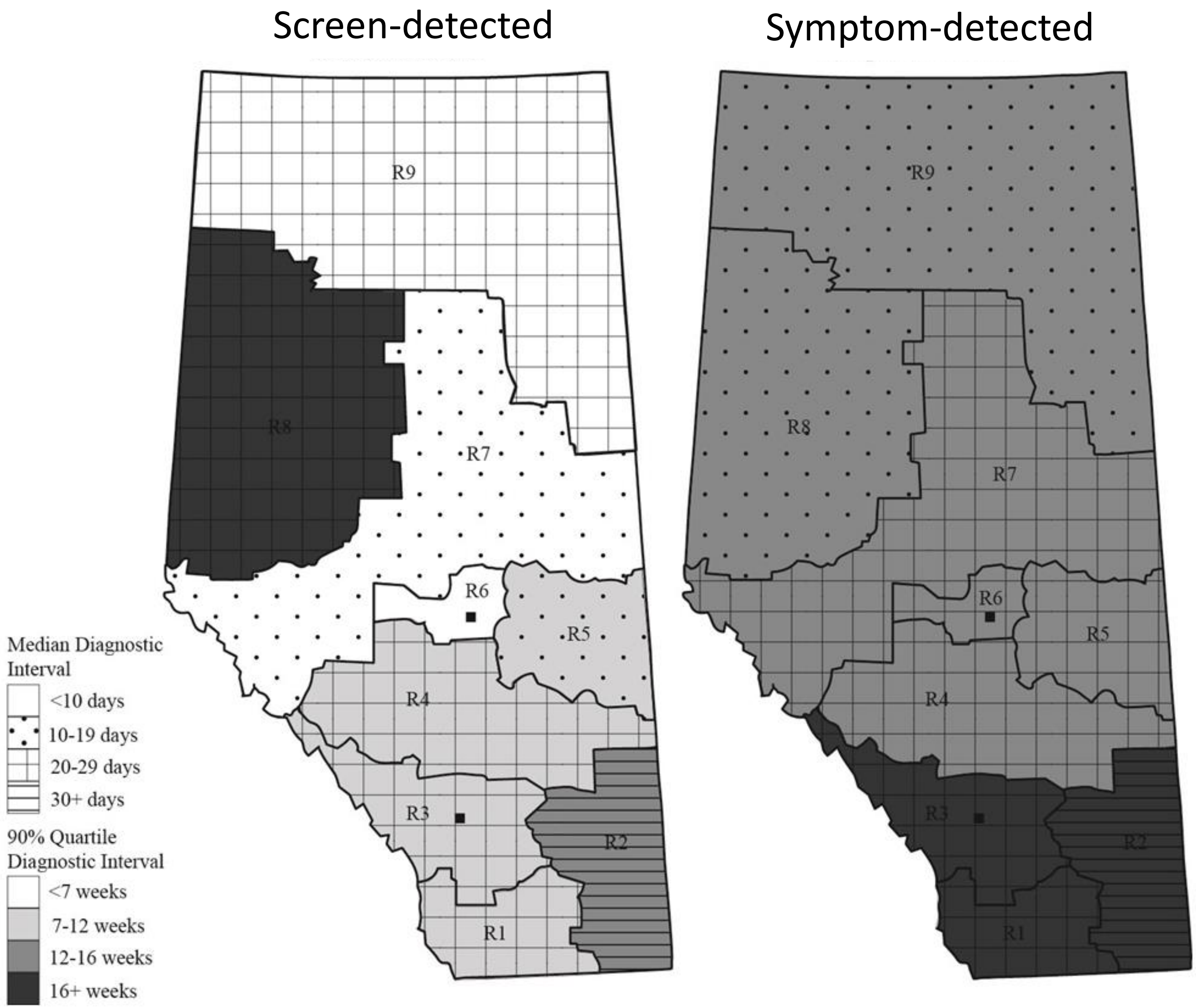
- Objectives
  - The proportion of screen vs. symptom-detected breast cancers
  - Time to diagnosis stratified by mode of detection
  - Assess the relationship of several demographic, clinical, and healthcare system factors to the first two objectives
- Study Population

  Female residents of Alberta with histologically-confirmed first primary breast cancer, diagnosed between 2004-2010.

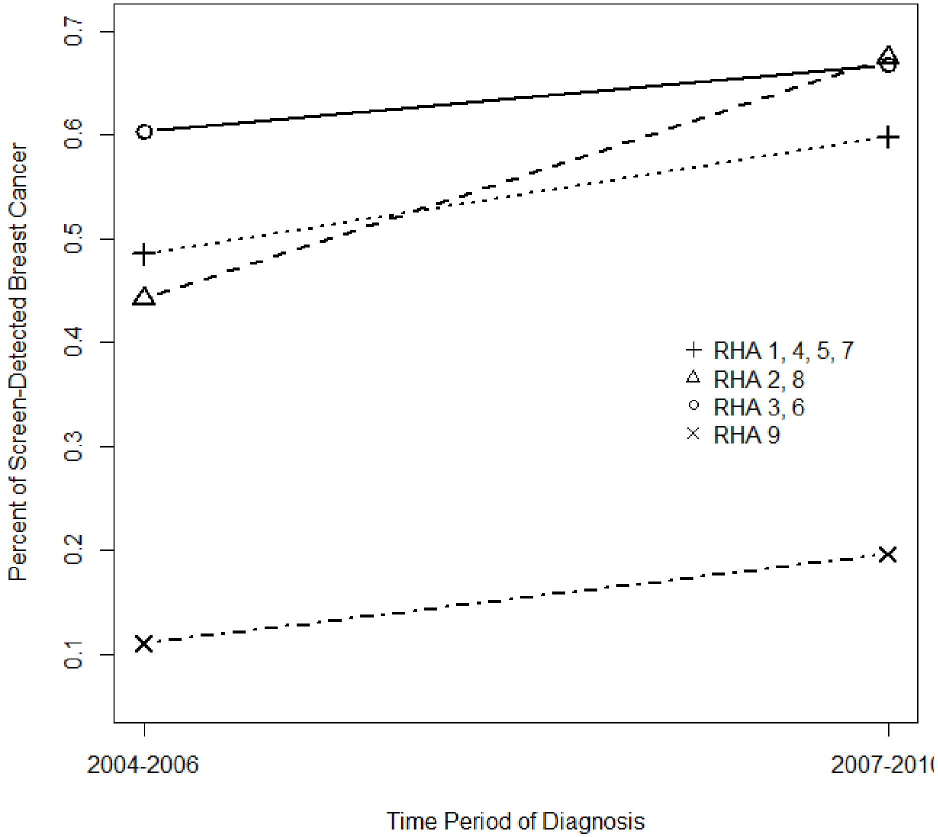
2.1 Detection Mode by age and RHA

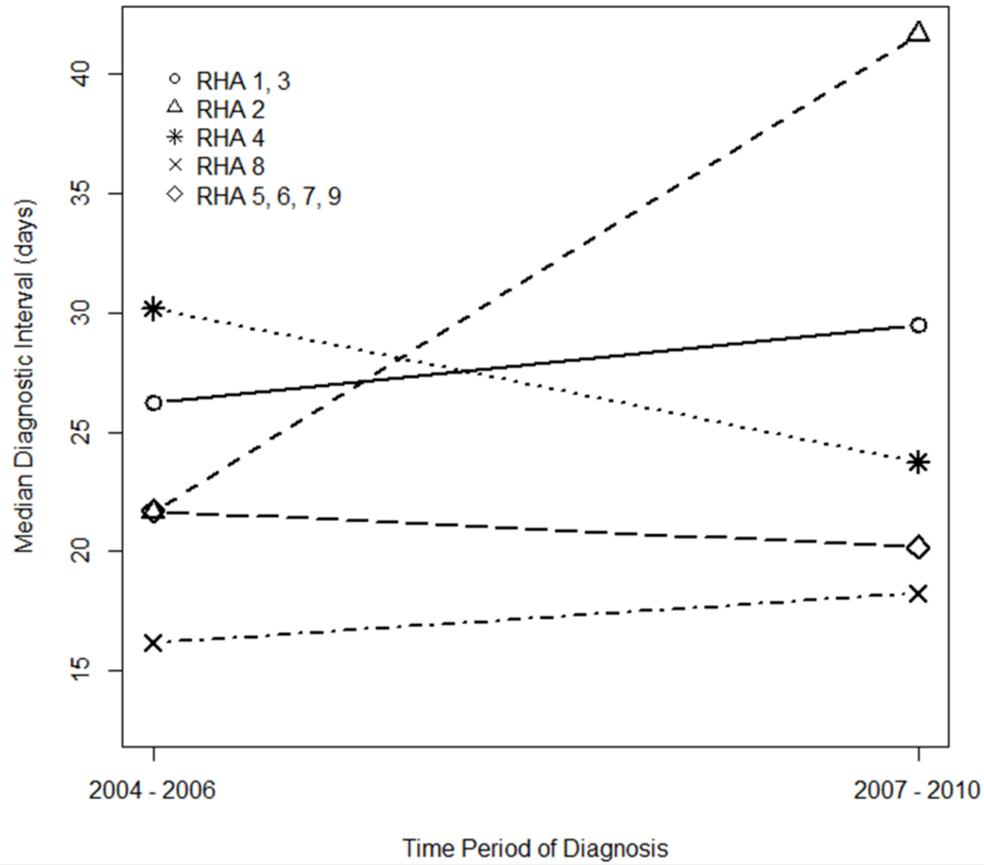


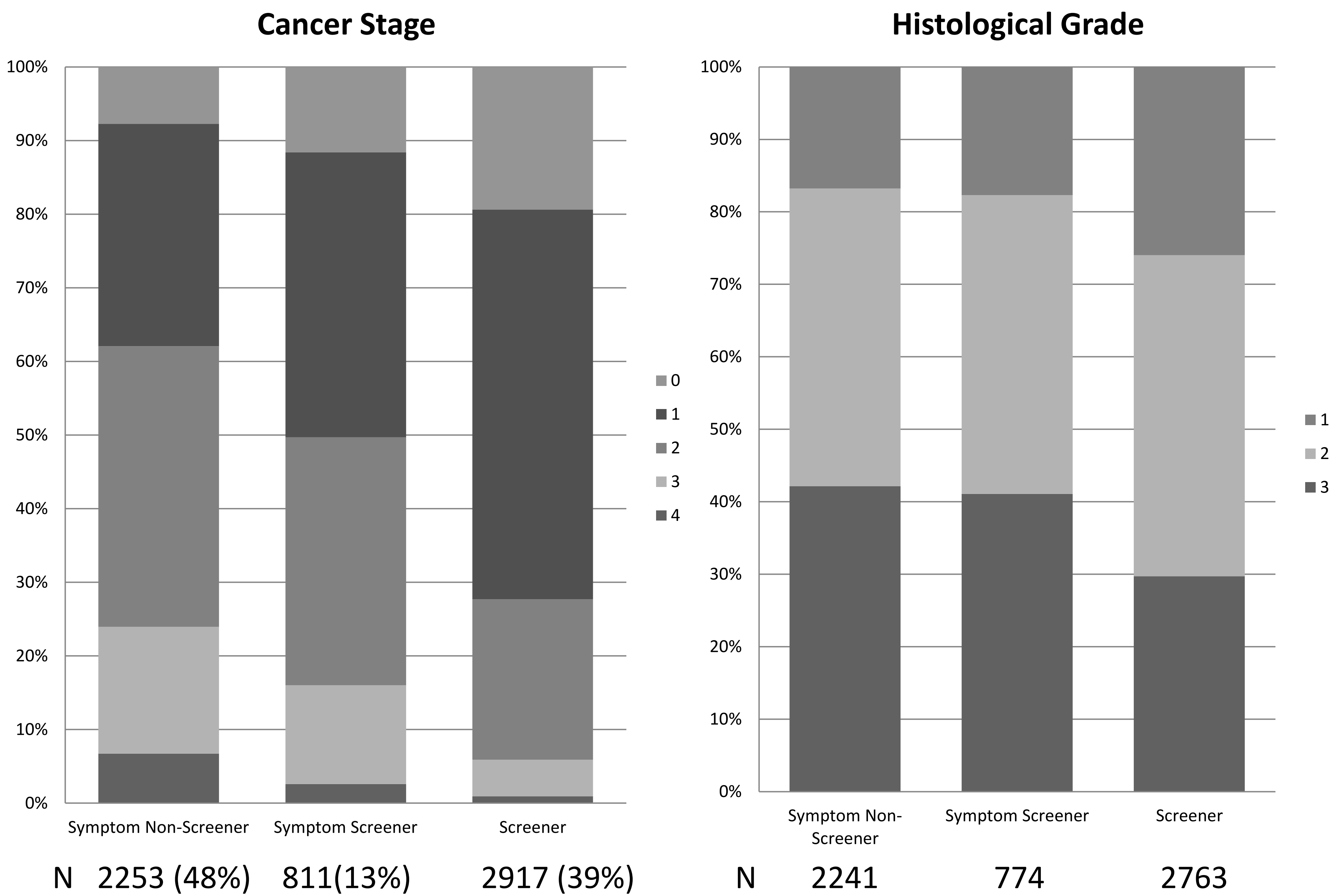
2.2 Diagnostic interval by detection mode and RHA



2.3 RHA Interact with time period






Cancer Stage
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0
1
2
3
4
Symptom Non-Screener
Symptom Screener
Screener
N 2253 (48%) 811(13%) 2917 (39%)
Histological Grade
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1
2
3
Symptom Non-Screener
Symptom Screener
Screener
N 2241 774 2763