
SEER Project

679 Final

SEER Analysis on Medical Discrimination

Group12 Zhiwei Liang, Fan Feng, Xiaozhou Lu, Yinfeng Zhou

Goal

→ Primary Goal

Analyze if there is significant discrimination in diagnosis

→ Secondary Goal and Steps

- ◆ Detect if there is significant difference between standard therapy and actual given therapy
- ◆ Analyse if deviation from standard therapy causes worse outcomes
- ◆ Analyse what kind of people are more likely to be given worse therapy

Data Processing

In this project, we used two datasets : SEER dataset & NCCN guidelines for head and neck cancers.

We transformed the code of '**tumor size**' , 'cs extension' into its corresponding **T stage**, and '**lymph node**' into its corresponding **N stage** based on CS Manual (Collaborative Stage Data Collection System Coding Instructions)

Nasopharynx

CS Extension

- Note 1: Parapharyngeal involvement denotes postero-lateral infiltration of tumor beyond the pharyngobasilar fascia. The pharyngobasilar fascia is the fibrous layer of the pharyngeal wall between the mucosa and the muscular layer, attached superiorly to the basilar part of the occipital bone and diminishing in thickness as it descends.
- Note 2: The masticator space primarily consists of the muscles of mastication, the medial and lateral pterygoid, masseter, and temporalis muscles. The space also includes the ramus of the mandible and the third division of cranial nerve V as it passes through the foramen ovale into the suprahyoid neck.
- Note 3: Use code 300 for localized tumor ONLY if no information is available to assign codes 105 or 205.
- Note 4: Use code 510, 585, 645, or 810 if the physician's assignment of T category is the ONLY information available about the extent of the tumor.

Code	Description	TNM 7 Map	TNM 6 Map
000	In situ, intraepithelial, noninvasive	Tis	Tis
100	OBSOLETE DATA RETAINED V 0200 Invasive tumor confined to one of the following subsites: Inferior wall (superior surface of soft palate) One lateral wall Posterior superior wall (vault)	ERROR	T1
105	Invasive tumor confined to one of the following subsites: Inferior wall (superior surface of soft palate) One lateral wall Posterior superior wall (vault)	T1	T1
200	OBSOLETE DATA RETAINED V 0200 Involvement of two or more subsites: Lateral wall extending into eustachian tube/middle ear Posterior, inferior, or lateral wall(s)	ERROR	T1
205	Involvement of two or more subsites: Lateral wall extending into eustachian tube/middle ear Posterior, inferior, or lateral wall(s)	T1	T1

Data Processing

We managed to **join NCCN guidelines standard therapy** to the seer dataset, and **compared them with recommended therapy**.

From these information, we were able to create several binary response variables:

1. **Diff_Reco_Sur** : 1 indicates difference in surgery, 0 indicates not.
2. **Diff_Reco_Rad**:1 indicates difference in radiation, 0 indicates not.
3. **Diff_Reco_Chem**:1 indicates difference in chemotherapy, 0 indicates not.
4. **surgery_refused**: 1 indicates the patient refused recommended surgery
5. **radiation_refused**: 1 indicates the patient refused recommended radiation

Diff_Reco_Sur	Diff_Reco_Rad	Diff_Reco_Chem	surgery_refused	radiation_refused
0	0	0	0	0

EDA

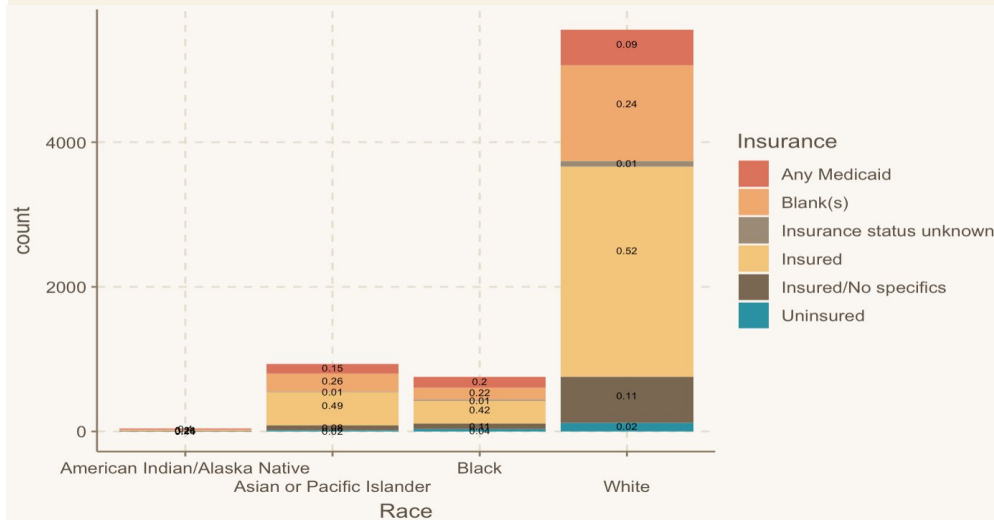
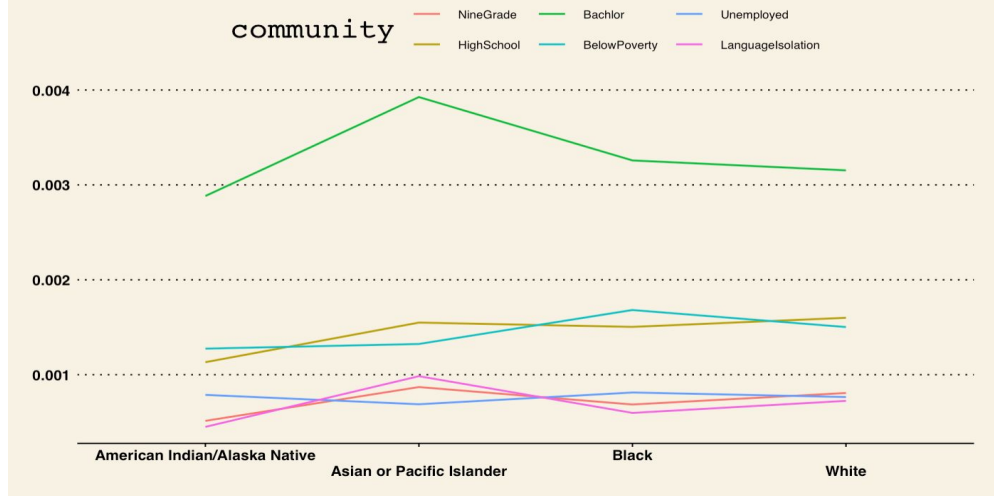
Community Information

- Four educational variables show similar pattern
- Poverty and unemployed rates show similar pattern

Insurance Information

- different races have different distributions of Insurance

Community Information to different Races



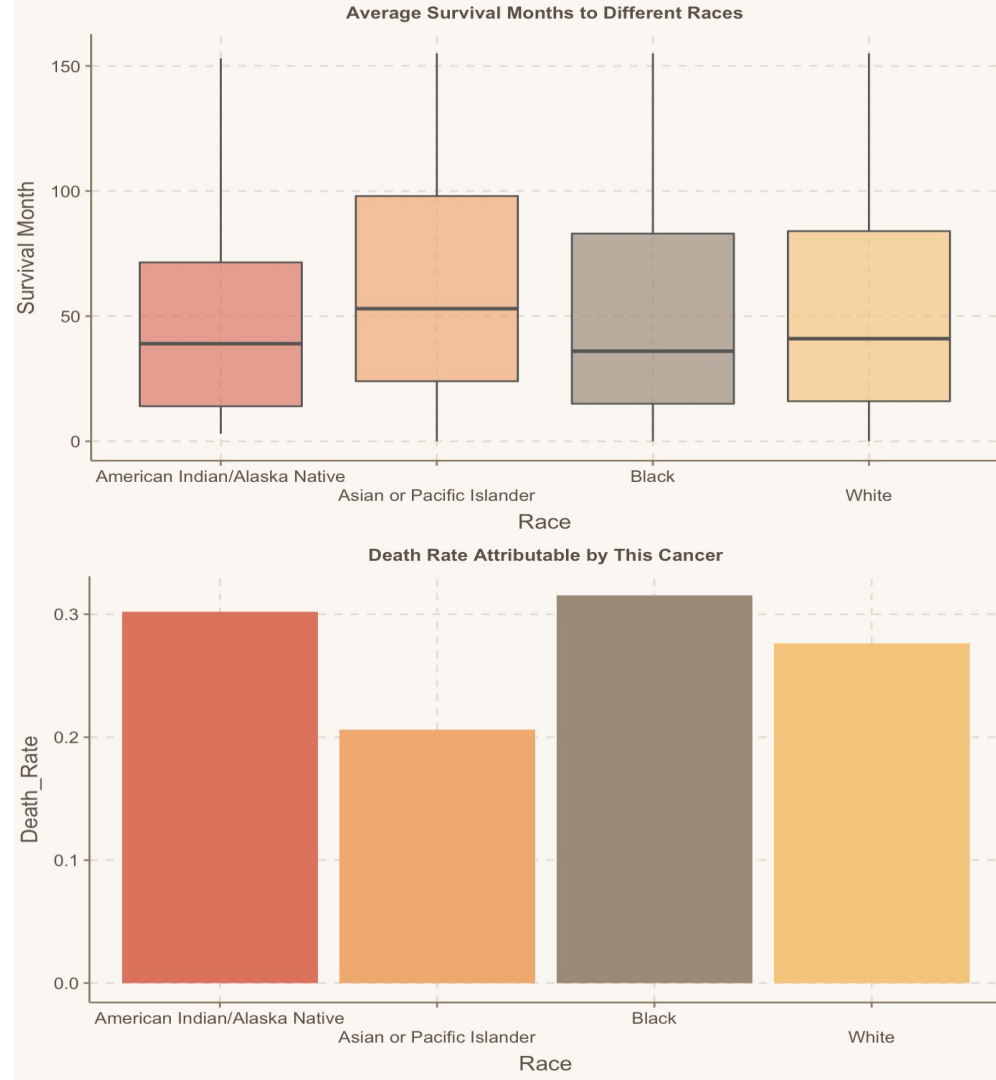
EDA

- **Survival Months**

- Survival months between races also shows similar pattern to educational level

- **Death Rate Directly Caused by This Cancer**

- Death rate shows similar pattern to poverty and unemployed rates



Modeling

Feature Extraction

Race

Age, Nine_Grade ,Atleast_Bachlor,

Person_Below_Poverty, Unemployed,

Median_Household_Income,

Language_Isolation

Model Select

Logistics Regression

Mediation Analysis

Random Forest



Topic1

difference between the **doctors' recommendation** and **NCCN standard therapy**

→ Logistic Regression

◆ For Surgery (*Diff_Reco_Sur*)

The coefficients of **Age** and **Language_Isolation** are statistically significant

◆ For Radiation & Chemotherapy (*Diff_Reco_Rad & Diff_Reco_Chem*)

The coefficients of **Age** is statistically significant

- doctors are **more likely to make different** surgery decisions from the NCCN guidelines standard on **older patients**
- the **lower** percent of households in **language isolation** is, the **higher** possibility for doctors to make **biased** surgery decisions



Topic2

difference between **doctors' recommendation** and **patients' attitude**

→ Logistic Regression

◆ For Surgery & Radiation (surgery_refuse & radiation_refuse)

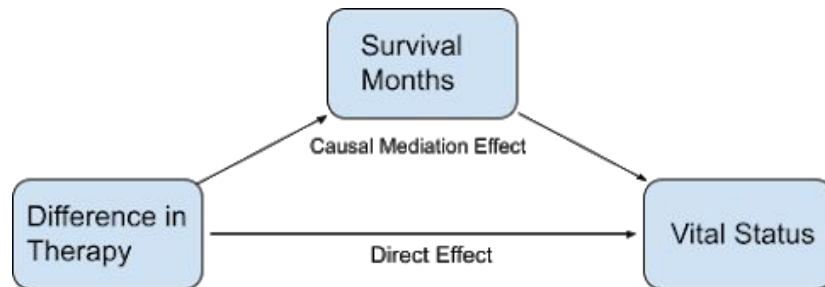
The coefficients of **Age** is statistically significant

→ **Older patients** are **more likely to refuse** the doctors' recommendations of surgery and radiation

Topic3

vital status

→ Causal Mediation Analysis



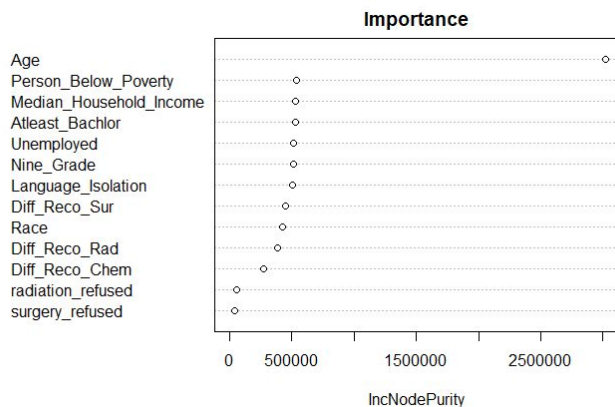
→ **Difference in therapy** does have effects on the patient's **survival months** and **vital status**

Treatment	ACME (Indirect Effect)	ADE(Direct Effect)	Total Effect
Diff_Reco_Sur	significant	significant	significant
Diff_Reco_Rad	significant	non-significant	significant
Diff_Reco_Chem	significant	significant	significant

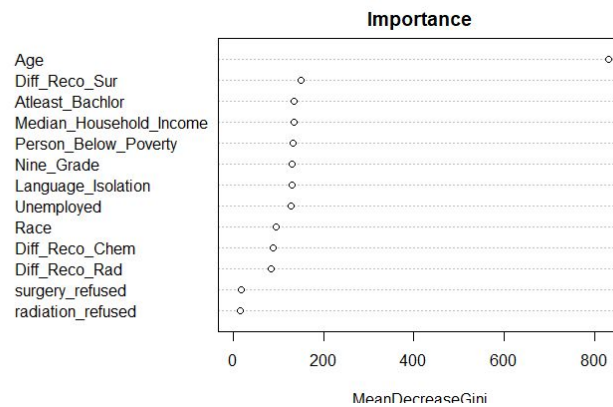
Topic3

vital status

→ Inference: Age might have affected everything!



Variable Importance Plot on Survival_Months



Variable Importance Plot on Vital_Status

→ Age is the most important feature affecting **survival months** and **vital status**

Conclusion&Discussion

- No significant discrimination on races, genders, ages or income levels during the treatment of head and neck cancers.
- The bias of therapy decisions from standard is more likely related to the patients' age. Older people are more likely to be given different therapy from standard, and eventually have worse outcomes.
- No obvious evidence to prove that the lower socioeconomic status, the lower the bias of therapy decisions will be.

Thank you

Group12 Zhiwei Liang, Fan Feng, Xiaozhou Lu, Yinfeng Zhou
