

The Infection Rate Prediction Based on CDC Transmision Risk Level

Sina Mokhtar

AUROC

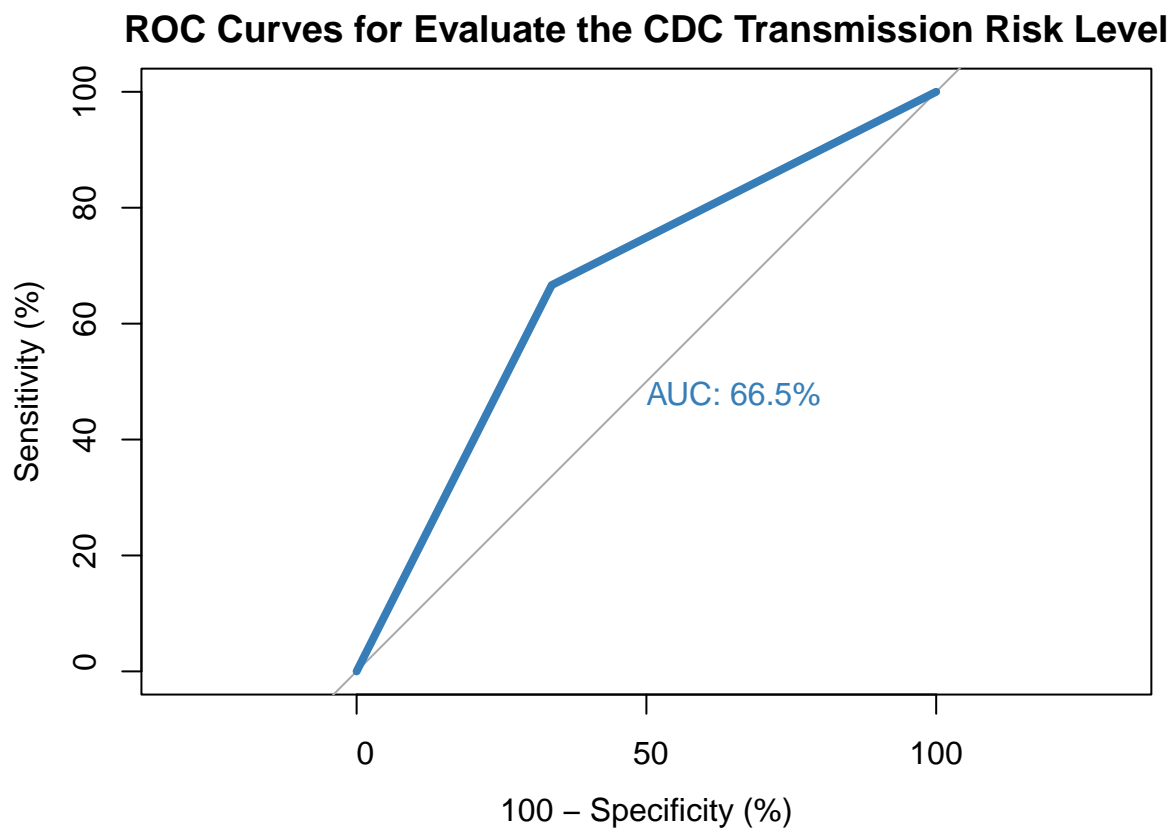
The AUROC (Area Under the Receiver Operating Characteristic) is a metric to evaluate the prediction performance. A value closer to 1 suggests a better prediction ability, while a value close to 0.5 indicates random prediction.

Method

To compare the next 21 days expected infection rates with the actual infection rates , we randomly sampled atleast %20 of counties from each UR category risk level for each day. We then compared their infection rates number with other counties of different risk levels in three weeks later.

The expected outcome (0 or 1) was determined based on the following revised criteria:

- If a county is categorized as low-risk, we expect its infection rate 3 weeks later to be lower than that of other counties.
- If a county is categorized as medium-risk, we expect its infection rate 3 weeks later to be higher than low-risk counties and lower than substantial and high-risk counties.
- If a county is categorized as substantial-risk, we expect its infection rate 3 weeks later to be lower than high-risk counties and higher than low and medium-risk counties.
- If a county is categorized as high-risk, we expect its infection rate 3 weeks later to be higher than that of other counties.



```
##
## Call:
## roc.data.frame(data = merged_counties, response = actual_higher_rt,      predictor = expected_higher_rt)
##
## Data: expected_higher_rt in 9207687 controls (actual_higher_rt 0) > 9034867 cases (actual_higher_rt 1)
## Area under the curve: 66.5%
## [1] 0.3350304
```

Confusion Matrix for Large central metro

```
## [1] 94741
```

Confusion Matrix for Large fringe metro

```
## [1] 572542
```

Confusion Matrix for Medium metro

```
## [1] 622657
```

Confusion Matrix for Small metro

```
## [1] 578699
```

Confusion Matrix for Micropolitan

```
## [1] 1056396
```

Confusion Matrix for Noncore

[1] 3096997