1 CRM Causal Inference Challenge: Numerical Results

The Mean Squared (Prediction) Error of each team for each of the 5 knock-out conditions is shown below. Formally, for each censored variable j and censored observation i, the table displays-

$$\text{MSE} = \sum_i (y_{ji} - \hat{y}_{ji})^2$$

The baseline method simply uses the mean of the uncensored data (all wild-type and other knock-out conditions) to predict the censored values. The "Roonil Wazlib" team (Arjun Sondhi and Ali Shojaie) were not able to attend the conference and thus do not have a poster. However, they used Invariant Causal Prediction of Peters et al (2015), and their results are included as an additional point of comparison.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
AI-ISU	29.85	2.48E+00	1.54E+00	1.57E-04	1.51E-02
BeschuitMet	23.54	2.36E-03	4.26E-03	1.92E-05	9.67E-03
Muisjes	22.78	4.37E-03	6.51E-03	1.99E-05	7.45E-03
Noncompliers	24.35	3.36E-03	4.26E-03	6.71E-05	5.31E-03
Roonil Wazlib	29.00	2.32E-02	6.15E-03	4.09E-04	2.63E-02
UC Berkeley	25.39	8.19E-01	5.72E-01	4.68E-04	6.83E-03
Baseline	25.89	6.14E+00	9.08E+00	1.43E-03	1.49E-02

Table 1: Mean Squared Prediction Error for each knock-out condition

The ratio of MSE for each team relative to the baseline method is also shown for a somewhat easier interpretation of performance.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
AI-ISU	1.15299	0.40414	0.16950	0.10979	1.01132
BeschuitMet	0.90905	0.00038	0.00047	0.01341	0.64741
Muisjes	0.87972	0.00071	0.00072	0.01393	0.49890
Noncompliers	0.94046	0.00055	0.00047	0.04689	0.35575
Roonil Wazlib	1.11997	0.00378	0.00068	0.28587	1.75963
UC Berkeley	0.98054	0.13344	0.06299	0.32674	0.45762
Baseline	1.00000	1.00000	1.00000	1.00000	1.00000

Table 2: Mean Squared Prediction Error for each knock-out condition relative to Baseline method

The rank (1 being smallest MSE, 7 being the largest) for each team in each knock-out condition is also shown below. Note that for knock-out condition 3, BeschuitMet and The Noncompliers predicted the same values up to 8 digits, so they both have a rank of 1.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
AI-ISU	7	6	6	4	6
BeschuitMet	2	1	1	1	4
Muisjes	1	3	4	2	3
Noncompliers	3	2	1	3	1
Roonil Wazlib	6	4	3	5	7
UC Berkeley	4	5	5	6	2
Baseline	5	7	7	7	5

Table 3: Rank of Mean Squared Prediction Error for each knock-out condition

The large predictive errors for knock-out condition 1 are due to the fact that the distribution of the censored variable (IMPC_HEM_027_001) looks very different in the knock-out condition than the distribution

for all uncensored data (wild-type and other uncensored knock-out conditions). Histograms of the censored values and uncensored values are shown below.

Censored Data for Knock-out Condition

Non-censored data