

OUTPUT FROM MULTIPLE LOGISTIC MODELS

The following generated output represents the multiple logistic regression runs for the results presented in Section 4.3 under EFFECT OF SYMPTOMS.

They are presented in the order mentioned below in the next page:

- a. In-hospital mortality using vital signs
- b. Hospital admission using vital signs
- c. Hospital admission using UMLS-based features
- d. In-hospital mortality using UMLS-based features
- e. Hospital admission using Gazetteer-based features
- f. In-hospital mortality using Gazetteer-based features

a. In-hospital mortality using vital signs

```

name: <unnamed>
log: Y:\DataStageData\ForChris2\Update\OutputV9.smc1
log type: smc1
opened on: 4 Jan 2021, 15:05:42

```

```

.
end of do-file

```

```

. do "C:\Users\gsilver1\AppData\Local\Temp\2\STDc90_000000.tmp"

```

```

. logistic Mortality age_raw elixsum i.race Non_English_speaking male SBP_min24h Temp_max24h RR_max24h HR_max24h

```

```

Logistic regression                                Number of obs    =    4,757
                                                    LR chi2(13)      =    558.19
                                                    Prob > chi2       =    0.0000
Log likelihood = -683.26707                        Pseudo R2        =    0.2900

```

Mortality	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.048624	.0056249	8.85	0.000	1.037657	1.059707
elixsum	1.162229	.0207643	8.41	0.000	1.122236	1.203648
race						
Black	.6981624	.2103871	-1.19	0.233	.3867693	1.260262
Asian	1.460611	.4567242	1.21	0.226	.7913509	2.695877
Hispanic	.6410368	.3154294	-0.90	0.366	.2443653	1.681614
Declined	1.268845	.3948372	0.77	0.444	.6894984	2.334986
Other	2.543778	1.297586	1.83	0.067	.9360085	6.913194
Non_English_speak~g	.952669	.2385439	-0.19	0.846	.5831833	1.556249
male	1.547085	.2325314	2.90	0.004	1.152328	2.077075
SBP_min24h	.9817577	.0032595	-5.55	0.000	.97539	.9881669
Temp_max24h	1.147591	.0512533	3.08	0.002	1.051407	1.252573
RR_max24h	1.030051	.0050293	6.06	0.000	1.020241	1.039956
HR_max24h	1.010845	.0033344	3.27	0.001	1.004331	1.017401
_cons	7.88e-10	3.60e-09	-4.59	0.000	1.03e-13	6.04e-06

Note: _cons estimates baseline odds.

```

. lroc

```

Logistic model for Mortality

```

number of observations =    4757
area under ROC curve   =    0.8980

```

```

. predict mort_pr8, pr
(249 missing values generated)

```

```

. roctab Mortality mort_pr8, summary

```

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
4,757	0.8980	0.0079	0.88259	0.91338

.b. Hospital admission using vital signs

```
. logistic Inpatient age_raw elixsum i.race Non_English_speaking male SBP_min24h Temp_max24h RR_max24h HR_max24h
```

Logistic regression

Number of obs = 4,757

LR chi2(13) = 2359.10

Prob > chi2 = 0.0000

Pseudo R2 = 0.3694

Log likelihood = -2013.7665

Inpatient	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.02688	.0024574	11.08	0.000	1.022075	1.031708
elixsum	1.290513	.0165898	19.84	0.000	1.258404	1.323442
race						
Black	.8522927	.1072006	-1.27	0.204	.6660785	1.090566
Asian	1.481206	.2348056	2.48	0.013	1.085626	2.020927
Hispanic	1.28889	.2198611	1.49	0.137	.9226065	1.800592
Declined	1.22383	.1945166	1.27	0.204	.8962526	1.671136
Other	1.333478	.4083705	0.94	0.347	.7316613	2.43031
Non_English_speak~g	.8341411	.0975645	-1.55	0.121	.6632536	1.049058
male	.9852169	.0788614	-0.19	0.852	.842166	1.152567
SBP_min24h	.9777952	.0020985	-10.46	0.000	.9736909	.9819168
Temp_max24h	1.44438	.0406849	13.05	0.000	1.3668	1.526363
RR_max24h	1.034609	.0048548	7.25	0.000	1.025138	1.044168
HR_max24h	1.012552	.0023652	5.34	0.000	1.007927	1.017199
_cons	1.13e-17	3.19e-17	-13.88	0.000	4.59e-20	2.80e-15

Note: _cons estimates baseline odds.

```
.  
. lroc
```

Logistic model for Inpatient

number of observations = 4757

area under ROC curve = 0.8885

```
. predict IP_Comp_pr6, pr  
(249 missing values generated)
```

```
. roctab Inpatient IP_Comp_pr6, summary
```

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
4,757	0.8885	0.0047	0.87924	0.89770

.c. Hospital admission using UMLS-based features

```
. logistic Inpatient age_raw elixsum i.race Non_English_speaking male i.cdc_aches3 i.cdc_fatigue3 i.cdc_fever3 i.cdc_cough3  
> i.cdc_rhinitis_congestion3 i.cdc_nausea_vomiting3 i.cdc_diarrhea3
```

Logistic regression

Number of obs = 5,006

LR chi2(20) = 1824.93

Prob > chi2 = 0.0000

Pseudo R2 = 0.2745

Log likelihood = -2411.2108

Inpatient	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.020969	.0022016	9.62	0.000	1.016663	1.025293
elixsum	1.317334	.0156432	23.21	0.000	1.287028	1.348354
race						
Black	.8326897	.0954238	-1.60	0.110	.6651783	1.042386
Asian	1.996825	.2809742	4.91	0.000	1.515539	2.630952
Hispanic	1.497361	.2271517	2.66	0.008	1.112239	2.015835
Declined	1.313212	.1885255	1.90	0.058	.9911421	1.739938
Other	1.7353	.4806514	1.99	0.047	1.008335	2.986375
Non_English_speak~g	.9432035	.0988966	-0.56	0.577	.7679897	1.158392
male	1.101916	.0801622	1.33	0.182	.9554878	1.270784
1.cdc_aches3	.7981973	.0737103	-2.44	0.015	.6660475	.9565669
1.cdc_fatigue3	1.288411	.0988381	3.30	0.001	1.108552	1.497452
1.cdc_fever3	1.3002	.100936	3.38	0.001	1.116685	1.513875
1.cdc_cough3	.9360905	.0728134	-0.85	0.396	.803725	1.090255
1.cdc_dyspnea3	1.430492	.1077709	4.75	0.000	1.23412	1.658111
1.cdc_headaches3	.6069915	.0567572	-5.34	0.000	.5053479	.7290793
1.cdc_taste_smell~3	.6187076	.1018326	-2.92	0.004	.4481123	.8542482
1.cdc_sore_throat3	.6161481	.0741033	-4.03	0.000	.4867571	.7799341
1.cdc_rhinitis_co~3	.6814729	.0759973	-3.44	0.001	.5476755	.8479571
1.cdc_nausea_vomi~3	1.097397	.09015	1.13	0.258	.9341971	1.289107
1.cdc_diarrhea3	1.15247	.107521	1.52	0.128	.9598778	1.383705
_cons	.0485083	.0077597	-18.92	0.000	.0354528	.0663713

Note: **_cons** estimates baseline odds.

```
. lroc
```

Logistic model for Inpatient

```
number of observations =    5006
area under ROC curve   =    0.8436
```

```
. predict IP_Comp_pr4, pr
```

```
. roctab Inpatient IP_Comp_pr4, summary
```

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
5,006	0.8436	0.0055	0.83278	0.85442

d. In-hospital mortality using UMLS-based features

```
. logistic Mortality age_raw elixsum i.race Non_English_speaking male i.cdc_aches3 i.cdc_fatigue3 i.cdc_fever3 i.cdc_cough3
> i.cdc_rhinitis_congestion3 i.cdc_nausea_vomiting3 i.cdc_diarrhea3
```

```
Logistic regression      Number of obs    =    5,006
                        LR chi2(20)         =   453.43
                        Prob > chi2          =    0.0000
Log likelihood = -766.1683      Pseudo R2       =    0.2283
```

Mortality	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.04632	.0055468	8.54	0.000	1.035505	1.057248
elixsum	1.178887	.0197193	9.84	0.000	1.140865	1.218177
race						
Black	.8147901	.2293112	-0.73	0.467	.4693394	1.414505
Asian	2.168257	.63155	2.66	0.008	1.225125	3.837434
Hispanic	.880157	.3881901	-0.29	0.772	.3708008	2.089198
Declined	1.443089	.4254991	1.24	0.214	.8096808	2.572009
Other	2.858513	1.38488	2.17	0.030	1.105997	7.387993
Non_English_speak~g	1.015627	.2373616	0.07	0.947	.642392	1.605713
male	1.66481	.2409594	3.52	0.000	1.253616	2.210879
1.cdc_aches3	.9273786	.1895922	-0.37	0.712	.62121	1.384445
1.cdc_fatigue3	1.007941	.1462545	0.05	0.957	.7584444	1.339511
1.cdc_fever3	.982856	.1454943	-0.12	0.907	.7353341	1.313697
1.cdc_cough3	.9568709	.1444571	-0.29	0.770	.7117852	1.286346
1.cdc_dyspnea3	2.384961	.372015	5.57	0.000	1.756741	3.237836
1.cdc_headaches3	.7903849	.1851002	-1.00	0.315	.4994543	1.250782
1.cdc_taste_smell~3	1.020285	.421063	0.05	0.961	.4543993	2.290895
1.cdc_sore_throat3	.5881215	.2018837	-1.55	0.122	.300107	1.152545
1.cdc_rhinitis_co~3	.8752136	.2130792	-0.55	0.584	.5431012	1.410416
1.cdc_nausea_vomi~3	.6485158	.119147	-2.36	0.018	.4524135	.9296202
1.cdc_diarrhea3	1.094391	.198735	0.50	0.619	.7666539	1.562233
_cons	.0005045	.0002295	-16.69	0.000	.0002069	.0012303

Note: **_cons** estimates baseline odds.

. lroc

Logistic model for Mortality

number of observations = **5006**

area under ROC curve = **0.8658**

. predict mort_pr4, pr

. roctab Mortality mort_pr4, summary

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
5,006	0.8658	0.0089	0.84829	0.88329

.
end of do-file

.

e. Hospital admission using gazetteer-base features

```

name: <unnamed>
log: Y:\DataStageData\ForChris2\Update\OutputV10.smc1
log type: smc1
opened on: 4 Jan 2021, 15:11:14

```

```

.
end of do-file

```

```

. do "C:\Users\gsilver1\AppData\Local\Temp\2\STD2d98_000000.tmp"

```

```

. logistic Inpatient age_raw elixsum i.race Non_English_speaking male i.cdc_aches3 i.cdc_fatigue3 i.cdc_fever3 i.cdc_cough3
> i.cdc_rhinitis_congestion3 i.cdc_nausea_vomiting3 i.cdc_diarrhea3

```

```

Logistic regression               Number of obs   =      5,006
                                LR chi2(20)        =     1771.58
                                Prob > chi2         =      0.0000
                                Pseudo R2          =      0.2665
Log likelihood = -2437.8866

```

Inpatient	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.022782	.0021709	10.61	0.000	1.018536	1.027046
elixsum	1.323338	.0157023	23.61	0.000	1.292917	1.354474
race						
Black	.783832	.0892107	-2.14	0.032	.6271114	.9797185
Asian	2.009845	.281751	4.98	0.000	1.526992	2.645384
Hispanic	1.454841	.2192577	2.49	0.013	1.082759	1.954786
Declined	1.329052	.1900652	1.99	0.047	1.004184	1.75902
Other	1.713019	.4709804	1.96	0.050	.9993818	2.93625
Non_English_speak~g	.9554307	.0996403	-0.44	0.662	.7788052	1.172113
male	1.113001	.0801137	1.49	0.137	.9665536	1.281636
1.cdc_aches3	.8832033	.0722991	-1.52	0.129	.752283	1.036908
1.cdc_fatigue3	1.0979	.0821109	1.25	0.212	.9482047	1.271227
1.cdc_fever3	1.209431	.0974641	2.36	0.018	1.032727	1.41637
1.cdc_cough3	.9496626	.0810959	-0.60	0.545	.8033068	1.122683
1.cdc_dyspnea3	1.349678	.1097498	3.69	0.000	1.150839	1.582873
1.cdc_headaches3	.6690779	.0522643	-5.14	0.000	.574098	.7797716
1.cdc_taste_smell~3	.7548059	.0992174	-2.14	0.032	.5833739	.9766155
1.cdc_sore_throat3	.7433359	.0683225	-3.23	0.001	.6207952	.8900653
1.cdc_rhinitis_co~3	.7969962	.0778992	-2.32	0.020	.6580501	.9652805
1.cdc_nausea_vomi~3	.9725948	.0762973	-0.35	0.723	.8339838	1.134243
1.cdc_diarrhea3	1.037287	.0866239	0.44	0.661	.8806738	1.221751
_cons	.0500902	.0081024	-18.51	0.000	.0364809	.0687765

Note: _cons estimates baseline odds.

```

. lroc

```

Logistic model for Inpatient

```

number of observations =      5006
area under ROC curve   =      0.8403

```

```
. predict IP_Comp_pr4, pr
. roctab Inpatient IP_Comp_pr4, summary
```

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
5,006	0.8403	0.0056	0.82933	0.85121

. f. In-hospital mortality using gazetteer-base features

```
.
.
. logistic Mortality age_raw elixsum i.race Non_English_speaking male i.cdc_aches3 i.cdc_fatigue3 i.cdc_fever3 i.cdc_cough3
> i.cdc_rhinitis_congestion3 i.cdc_nausea_vomiting3 i.cdc_diarrhea3
```

```
Logistic regression                                Number of obs    =      5,006
                                                    LR chi2(20)        =     444.48
                                                    Prob > chi2         =     0.0000
Log likelihood = -770.64521                        Pseudo R2          =     0.2238
```

Mortality	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.046759	.0053558	8.93	0.000	1.036315	1.057309
elixsum	1.186811	.0198063	10.26	0.000	1.14862	1.226273
race						
Black	.750071	.2100726	-1.03	0.304	.4332176	1.29867
Asian	2.150846	.6237407	2.64	0.008	1.218324	3.797134
Hispanic	.820113	.3644646	-0.45	0.655	.3432362	1.95954
Declined	1.30447	.3829661	0.91	0.365	.7337338	2.319155
Other	2.864565	1.393573	2.16	0.031	1.103978	7.432877
Non_English_speak~g	1.089571	.2533445	0.37	0.712	.690774	1.7186
male	1.667976	.2403234	3.55	0.000	1.257615	2.212237
1.cdc_aches3	.8746948	.1595837	-0.73	0.463	.611729	1.250703
1.cdc_fatigue3	.7714079	.112255	-1.78	0.075	.5799861	1.026008
1.cdc_fever3	1.099097	.1701674	0.61	0.542	.8114257	1.488754
1.cdc_cough3	1.008763	.1623538	0.05	0.957	.7358579	1.38288
1.cdc_dyspnea3	1.620408	.2726329	2.87	0.004	1.165226	2.253402
1.cdc_headaches3	.6864618	.1138651	-2.27	0.023	.4959359	.950183
1.cdc_taste_smell~3	.7921279	.2872633	-0.64	0.520	.3891429	1.612433
1.cdc_sore_throat3	.7843541	.169253	-1.13	0.260	.5138476	1.197264
1.cdc_rhinitis_co~3	.8627713	.1909717	-0.67	0.505	.559094	1.331394
1.cdc_nausea_vomi~3	.636438	.0994506	-2.89	0.004	.4685396	.8645017
1.cdc_diarrhea3	1.029108	.1726404	0.17	0.864	.7407396	1.429737
_cons	.0007552	.000335	-16.21	0.000	.0003166	.0018015

Note: **_cons** estimates baseline odds.

```
. lroc
```

Logistic model for Mortality

```
number of observations =      5006
area under ROC curve   =     0.8621
```

```
. predict mort_pr4, pr
. roctab Mortality mort_pr4, summary
```

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
5,006	0.8621	0.0090	0.84453	0.87969

```
.
end of do-file
.
```


Mortality	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age_raw	1.04632	.0055468	8.54	0.000	1.035505	1.057248
elixsum	1.178887	.0197193	9.84	0.000	1.140865	1.218177
race						
Black	.8147901	.2293112	-0.73	0.467	.4693394	1.414505
Asian	2.168257	.63155	2.66	0.008	1.225125	3.837434
Hispanic	.880157	.3881901	-0.29	0.772	.3708008	2.089198
Declined	1.443089	.4254991	1.24	0.214	.8096808	2.572009
Other	2.858513	1.38488	2.17	0.030	1.105997	7.387993
Non_English_speak~g	1.015627	.2373616	0.07	0.947	.642392	1.605713
male	1.66481	.2409594	3.52	0.000	1.253616	2.210879
1.cdc_aches3	.9273786	.1895922	-0.37	0.712	.62121	1.384445
1.cdc_fatigue3	1.007941	.1462545	0.05	0.957	.7584444	1.339511
1.cdc_fever3	.982856	.1454943	-0.12	0.907	.7353341	1.313697
1.cdc_cough3	.9568709	.1444571	-0.29	0.770	.7117852	1.286346
1.cdc_dyspnea3	2.384961	.372015	5.57	0.000	1.756741	3.237836
1.cdc_headaches3	.7903849	.1851002	-1.00	0.315	.4994543	1.250782
1.cdc_taste_smell~3	1.020285	.421063	0.05	0.961	.4543993	2.290895
1.cdc_sore_throat3	.5881215	.2018837	-1.55	0.122	.300107	1.152545
1.cdc_rhinitis_co~3	.8752136	.2130792	-0.55	0.584	.5431012	1.410416
1.cdc_nausea_vomi~3	.6485158	.119147	-2.36	0.018	.4524135	.9296202
1.cdc_diarrhea3	1.094391	.198735	0.50	0.619	.7666539	1.562233
_cons	.0005045	.0002295	-16.69	0.000	.0002069	.0012303

Note: **_cons** estimates baseline odds.

. lroc

Logistic model for Mortality

number of observations = **5006**

area under ROC curve = **0.8658**

. predict mort_pr4, pr

. roctab Mortality mort_pr4, summary

Obs	ROC Area	Std. Err.	—Asymptotic Normal— [95% Conf. Interval]	
5,006	0.8658	0.0089	0.84829	0.88329

.
end of do-file

.