

Race, Ethnicity, Gender, and Age in Covid

Mahesh Shukla

A3.AI

4/29/2021

*Work-in-progress
These are findings from the analysis of unlinked datasets
Deeper analysis using linked dataset will start soon

Race and Ethnicity in Covid

Race and Ethnicity in Covid

Descriptive analysis

N= 20,449,202 patients

Racial / Ethnic group	Risk of Covid infection in 21-25 age group	Risk of Covid death in 21-25 age group	Risk of Covid infection in 46-50 age group	Risk of Covid death in 46-50 age group
Caucasian	0.94%	1 in 100,000	1.08%	5 in 100,000
African American	1.27%	7 in 100,000	1.75%	16 in 100,000
Hispanic	3.11%	6 in 100,000	4.00%	32 in 100,000
Asian	16.00%	0 in 100,000	1.77%	6 in 100,000
Native Americans and Pacific Islander	2.03%	0 in 100,000	2.58%	10 in 100,000

The numbers of Asian and Native Americans and Pacific Islander group in the data are small For the rest, all the risk differences mentioned in the table are statistically significant

Race and Ethnicity in Covid

Logistic regression modeling Covid infection

N= 20,004,824 patients

Covid patients out of the above: 363,279

Controlling for age, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions and when compared to Caucasians,

- there is 97% increase in the odds of contracting Covid for Hispanics
- there is 73% increase in the odds of contracting Covid for Asians
- there is 27% increase in the odds of contracting Covid for Native Americans and Pacific Islanders
- there is 19% increase in the odds of contracting Covid for African Americans

Logistic regression modeling death in the face of Covid, or death given Covid

N= 363,279 Covid patients

Deaths in 2020: 23,329

Controlling for age, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions and when compared to Caucasians,

- there is 27% increase in the odds of death in the face of Covid for Hispanics

Race and Ethnicity in Covid infections

	DF	Estimate	Std err	Chi-Sq	p-value	Effect	OR	CI	CI
3	1	0.3839	0.00778	2508.6124	<.0001	RACE 3 vs 1	1.969	1.937	2.002
4	1	0.2536	0.0153	266.0811	<.0001	RACE 4 vs 1	1.729	1.666	1.794
5	1	-0.0478	0.00861	30.8128	<.0001	RACE 5 vs 1	1.271	1.248	1.295
2	1	0.2041	0.000713	81963.6325	<.0001	UMC_30	1.226	1.225	1.228
2	1	-0.11	0.0074	221.0788	<.0001	RACE 2 vs 1	1.195	1.176	1.213
1	1	0.1038	0.00424	599.0384	<.0001	EDUCATION 1 vs 4	1.182	1.163	1.202
10	1	0.1013	0.00897	127.5259	<.0001	JOB 10 vs 7	1.094	1.068	1.12
1	1	0.068	0.00457	221.3371	<.0001	HOMEVALUE 1 vs 4	1.093	1.073	1.114
0	1	-0.2038	0.0106	366.9121	<.0001	RACE 0 vs 1	1.088	1.062	1.114
2	1	0.0415	0.00257	260.9582	<.0001	GENDER 2 vs 1	1.087	1.076	1.098
	1	0.0637	0.0019	1338.3051	<.0001	AIG_EMPLOYMENT	1.072	1.068	1.076
2	1	0.1646	0.00842	382.7539	<.0001	RELIGION 2 vs 1	1.067	1.046	1.089
2	1	-0.00665	0.00284	5.4584	0.0195	EDUCATION 2 vs 4	1.059	1.044	1.073
4	1	0.0614	0.00745	67.9348	<.0001	JOB 4 vs 7	1.051	1.03	1.073
	1	0.0418	0.00146	818.8837	<.0001	OS_RISKTAKING_HEALTH	1.043	1.04	1.046
0	1	0.021	0.00315	44.4899	<.0001	DAILY_SMOKER 0 vs 1	1.043	1.03	1.056
14	1	0.0528	0.00656	64.8439	<.0001	JOB 14 vs 7	1.042	1.022	1.062
2	1	0.0516	0.00318	263.6179	<.0001	POLITICS 2 vs 1	1.036	1.024	1.047
0	1	0.0178	0.00481	13.6411	0.0002	HISPANIC_PROF 0 vs 1	1.036	1.017	1.056
0	1	0.0172	0.00252	46.5937	<.0001	HOMEOWNER 0 vs 1	1.035	1.025	1.045
3	1	-0.0336	0.00325	106.5438	<.0001	EDUCATION 3 vs 4	1.031	1.017	1.045
0	1	0.0138	0.00579	5.6541	0.0174	HEALTH_INSURANCE 0 vs 1	1.028	1.005	1.051
0	1	0.011	0.00246	19.9619	<.0001	PRIMARY_CARE 0 vs 1	1.022	1.012	1.032
	1	0.015	0.00173	75.0647	<.0001	HS_PRIMARY_LANG_SPAN	1.015	1.012	1.019
	1	0.0117	0.000225	2702.3017	<.0001	AGE	1.012	1.011	1.012
2	1	-0.00914	0.00398	5.2641	0.0218	HOMEVALUE 2 vs 4	1.012	0.996	1.029
0	1	0.00514	0.00248	4.2866	0.0384	POOR_SLEEP_QUALITY 0 vs 1	1.01	1.001	1.02
0	1	0.00229	0.00271	0.714	0.3981	HEAVY_DRINKER 0 vs 1	1.005	0.994	1.015
0	1	-0.00235	0.00235	0.9999	0.3173	OVERWEIGHT 0 vs 1	0.995	0.986	1.005
0	1	-0.0392	0.00597	42.9928	<.0001	HOMEVALUE 0 vs 4	0.982	0.962	1.003
0	1	-0.00944	0.00287	10.8431	0.001	JUNK_DIET 0 vs 1	0.981	0.97	0.992
3	1	-0.0457	0.0087	27.5908	<.0001	JOB 3 vs 7	0.944	0.922	0.966
3	1	-0.0683	0.00316	465.754	<.0001	POLITICS 3 vs 1	0.919	0.908	0.929
0	1	-0.0474	0.00241	387.741	<.0001	MED_UTILIZATION 0 vs 1	0.91	0.901	0.918
0	1	-0.0554	0.00246	504.9846	<.0001	DAILY_EXERCISE 0 vs 1	0.895	0.887	0.904
0	1	-0.0743	0.00295	635.3443	<.0001	STRESSED 0 vs 1	0.862	0.852	0.872
4	1	-0.0597	0.0051	136.8815	<.0001	RELIGION 4 vs 1	0.852	0.844	0.861
5	1	-0.0737	0.00535	189.5643	<.0001	RELIGION 5 vs 1	0.841	0.831	0.85
6	1	-0.1945	0.0268	52.8428	<.0001	JOB 6 vs 7	0.814	0.764	0.867
3	1	-0.1311	0.0126	107.7047	<.0001	RELIGION 3 vs 1	0.794	0.769	0.819
0	1	0.00644	0.00994	0.4191	0.5174	AFRICAN_AMERICAN_PRO 0 vs 1	1.013	0.974	1.053
5	1	0.013	0.0241	0.2877	0.5917	JOB 5 vs 7	1.001	0.946	1.06
3	1	0.00137	0.00429	0.1016	0.7499	HOMEVALUE 3 vs 4	1.023	1.011	1.035
1	1	0.9056	7.1928	0.0159	0.8998	INCOME 1 vs 4	0.985	0.967	1.004
0	1	-3.5281	28.7713	0.015	0.9024	INCOME 0 vs 4	0.012	<0.001	>999.999
3	1	0.8562	7.1928	0.0142	0.9052	INCOME 3 vs 4	0.938	0.927	0.949
2	1	0.8458	7.1928	0.0138	0.9064	INCOME 2 vs 4	0.928	0.913	0.943

Race and Ethnicity in Covid deaths

Parameter	DF	Estimate	Std err	Chi-Sq	p-value	Effect	OR	CI	CI
HOMEVALUE	1 1	0.2682	0.0183	214.6745	<.0001	HOMEVALUE 1 vs 4	1.729	1.596	1.873
GENDER	2 1	0.2387	0.00987	585.2399	<.0001	GENDER 2 vs 1	1.612	1.551	1.676
HOMEVALUE	2 1	0.083	0.0167	24.6031	<.0001	HOMEVALUE 2 vs 4	1.437	1.335	1.545
RACE	3 1	0.3484	0.0366	90.7173	<.0001	RACE 3 vs 1	1.27	1.176	1.372
HOMEVALUE	3 1	-0.0486	0.0187	6.7587	0.0093	HOMEVALUE 3 vs 4	1.259	1.19	1.333
RELIGION	4 1	0.1961	0.0271	52.2108	<.0001	RELIGION 4 vs 1	1.219	1.17	1.269
INCOME	3 1	0.0436	0.0159	7.5079	0.0061	INCOME 3 vs 4	1.162	1.095	1.234
HOMEOWNER	0 1	0.0722	0.01	51.7845	<.0001	HOMEOWNER 0 vs 1	1.155	1.111	1.202
INCOME	2 1	0.0333	0.0165	4.0788	0.0434	INCOME 2 vs 4	1.15	1.066	1.241
HEAVY_DRINKER	0 1	0.0625	0.0168	13.8171	0.0002	HEAVY_DRINKER 0 vs 1	1.133	1.061	1.21
RELIGION	5 1	0.101	0.0292	11.9326	0.0006	RELIGION 5 vs 1	1.108	1.053	1.167
HEALTH_INSURANCE	0 1	0.0474	0.0228	4.2957	0.0382	HEALTH_INSURANCE 0 vs 1	1.099	1.005	1.202
AGE	1	0.0726	0.00107	4611.4887	<.0001	AGE	1.075	1.073	1.078
EDUCATION	2 1	0.0269	0.012	4.9917	0.0255	EDUCATION 2 vs 4	1.072	1.009	1.138
MED_UTILIZATION	0 1	0.0303	0.0111	7.4412	0.0064	MED_UTILIZATION 0 vs 1	1.062	1.017	1.11
JOB	4 1	0.0789	0.0365	4.6704	0.0307	JOB 4 vs 7	1.044	0.945	1.154
UMC_30	1	0.0296	0.00254	135.5258	<.0001	UMC_30	1.03	1.025	1.035
HS_PRIMARY_LANG_SPAN	1	-0.0381	0.00814	21.9607	<.0001	HS_PRIMARY_LANG_SPAN	0.963	0.947	0.978
POLITICS	2 1	-0.0405	0.0132	9.3812	0.0022	POLITICS 2 vs 1	0.93	0.89	0.973
AIQ_EMPLOYMENT	1	-0.0772	0.01	59.2841	<.0001	AIQ_EMPLOYMENT	0.926	0.908	0.944
JUNK_DIET	0 1	-0.0653	0.0112	33.887	<.0001	JUNK_DIET 0 vs 1	0.878	0.84	0.917
OVERWEIGHT	0 1	-0.0672	0.00938	51.3774	<.0001	OVERWEIGHT 0 vs 1	0.874	0.843	0.907
DAILY_SMOKER	0 1	-0.0688	0.0157	19.2584	<.0001	DAILY_SMOKER 0 vs 1	0.871	0.82	0.927
HISPANIC_PROF	0 1	-0.0715	0.0263	7.3765	0.0066	HISPANIC_PROF 0 vs 1	0.867	0.782	0.961
RACE	0 1	-0.2023	0.0492	16.9078	<.0001	RACE 0 vs 1	0.732	0.657	0.816
RELIGION	2 1	-0.3548	0.0496	51.2609	<.0001	RELIGION 2 vs 1	0.703	0.627	0.787
RACE	5 1	-0.2928	0.0441	43.9912	<.0001	RACE 5 vs 1	0.669	0.608	0.737

Odds of Covid infection for Black, Indigenous, and People of Color (BIPOCs)

Logistic regression modeling Covid infection

N= 20,004,824 patients

Covid patients out of the above: 363,279

Caucasians out of the above: 13,301,754

Controlling for age, income, education, employment, job, politics, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions and when compared with Caucasians,

- There is 37.5% increase in the odds of contracting Covid for BIPOCs
- This increase is statistically highly significant

Odds of Covid infection for BIPOCs

Parameter		DF	Estimate	Std err	Chi-Sq	p-value	Effect	OR	CI	CI
WHITE	0	1	0.1591	0.00253	3954.6498	<.0001	WHITE 0 vs 1	1.375	1.361	1.388
UMC_30		1	0.2053	0.000713	83052.744	<.0001	UMC_30	1.228	1.226	1.23
EDUCATION	1	1	0.1164	0.00424	754.3115	<.0001	EDUCATION 1 vs 4	1.214	1.195	1.234
AFRICAN_AMERICAN_PRO	0	1	0.0547	0.00366	32.0583	<.0001	AFRICAN_AMERICAN_PRO 0 vs 1	1.116	1.074	1.159
HOMEVALUE	1	1	0.0753	0.00456	271.3493	<.0001	HOMEVALUE 1 vs 4	1.105	1.085	1.126
GENDER	2	1	0.0432	0.00256	363.1683	<.0001	GENDER 2 vs 1	1.103	1.092	1.115
JOB	10	1	0.1101	0.00836	150.3863	<.0001	JOB 10 vs 7	1.096	1.071	1.123
HS_PRIMARY_LANG_SPAN		1	0.0808	0.00118	4653.0484	<.0001	HS_PRIMARY_LANG_SPAN	1.084	1.082	1.087
AIG_EMPLOYMENT		1	0.07	0.0019	1359.4198	<.0001	AIG_EMPLOYMENT	1.073	1.063	1.077
DAILY_SMOKER	0	1	0.0312	0.00314	98.8493	<.0001	DAILY_SMOKER 0 vs 1	1.064	1.051	1.078
JOB	4	1	0.0736	0.00745	97.6002	<.0001	JOB 4 vs 7	1.057	1.036	1.079
POOR_SLEEP_QUALITY	0	1	0.0206	0.00244	71.0302	<.0001	POOR_SLEEP_QUALITY 0 vs 1	1.042	1.032	1.052
EDUCATION	3	1	-0.0376	0.00325	133.5396	<.0001	EDUCATION 3 vs 4	1.041	1.027	1.055
OS_RISKTAKING_HEALTH		1	0.0401	0.00146	758.5047	<.0001	OS_RISKTAKING_HEALTH	1.041	1.038	1.044
HEALTH_INSURANCE	0	1	0.0184	0.00578	10.1307	0.0015	HEALTH_INSURANCE 0 vs 1	1.038	1.014	1.061
POLITICS	2	1	0.0482	0.00309	242.9776	<.0001	POLITICS 2 vs 1	1.033	1.022	1.045
HOMEOWNER	0	1	0.0164	0.00252	42.5277	<.0001	HOMEOWNER 0 vs 1	1.033	1.023	1.044
JOB	14	1	0.0478	0.00656	53.1032	<.0001	JOB 14 vs 7	1.03	1.011	1.05
HEAVY_DRINKER	0	1	0.0137	0.0027	25.9295	<.0001	HEAVY_DRINKER 0 vs 1	1.028	1.017	1.039
RELIGION	2	1	0.1731	0.00836	458.6891	<.0001	RELIGION 2 vs 1	1.023	1.002	1.043
PRIMARY_CARE	0	1	0.00718	0.00245	8.5668	0.0034	PRIMARY_CARE 0 vs 1	1.014	1.005	1.024
AGE		1	0.0115	0.000224	2639.0324	<.0001	AGE	1.012	1.011	1.012
JUNK_DIET	0	1	-0.0121	0.00287	17.6398	<.0001	JUNK_DIET 0 vs 1	0.976	0.965	0.987
HOMEVALUE	0	1	-0.0517	0.00597	75.1617	<.0001	HOMEVALUE 0 vs 4	0.974	0.954	0.994
HISPANIC_PROF	0	1	-0.0188	0.00475	15.6338	<.0001	HISPANIC_PROF 0 vs 1	0.963	0.945	0.981
JOB	3	1	-0.0543	0.00863	38.368	<.0001	JOB 3 vs 7	0.93	0.909	0.952
POLITICS	3	1	-0.0638	0.00315	408.9441	<.0001	POLITICS 3 vs 1	0.924	0.914	0.934
MED_UTILIZATION	0	1	-0.0451	0.0024	353.9969	<.0001	MED_UTILIZATION 0 vs 1	0.914	0.905	0.922
DAILY_EXERCISE	0	1	-0.0571	0.00247	536.5668	<.0001	DAILY_EXERCISE 0 vs 1	0.892	0.883	0.901
STRESSED	0	1	-0.0712	0.00295	583.2961	<.0001	STRESSED 0 vs 1	0.867	0.857	0.877
JOB	6	1	-0.2104	0.0268	61.8649	<.0001	JOB 6 vs 7	0.796	0.747	0.848
RELIGION	4	1	-0.0724	0.00462	245.7476	<.0001	RELIGION 4 vs 1	0.795	0.788	0.803
RELIGION	5	1	-0.0763	0.00509	224.4416	<.0001	RELIGION 5 vs 1	0.792	0.783	0.801
RELIGION	3	1	-0.1871	0.011	289.4733	<.0001	RELIGION 3 vs 1	0.709	0.69	0.729
HOMEVALUE	3	1	0.00467	0.00429	1.1857	0.2762	HOMEVALUE 3 vs 4	1.03	1.018	1.042
OVERWEIGHT	0	1	-0.00213	0.00235	0.8223	0.3645	OVERWEIGHT 0 vs 1	0.996	0.987	1.005
HOMEVALUE	2	1	-0.00322	0.00398	0.6544	0.4185	HOMEVALUE 2 vs 4	1.022	1.006	1.039
JOB	5	1	0.0152	0.0241	0.394	0.5302	JOB 5 vs 7	0.997	0.942	1.056
EDUCATION	2	1	-0.00133	0.00284	0.2213	0.638	EDUCATION 2 vs 4	1.079	1.065	1.094
INCOME	1	1	0.8855	7.5387	0.0138	0.9065	INCOME 1 vs 4	0.971	0.954	0.989
INCOME	0	1	-3.4813	30.1549	0.0133	0.9081	INCOME 0 vs 4	0.012	<0.001	>999.999
INCOME	3	1	0.8476	7.5387	0.0126	0.9105	INCOME 3 vs 4	0.935	0.924	0.946
INCOME	2	1	0.8337	7.5387	0.0122	0.9119	INCOME 2 vs 4	0.922	0.908	0.937

Tamara Gilkes Borr of the Economist summarized our findings in her article, as follows:

“NON-WHITE Americans have suffered disproportionately from covid-19. According to A3.AI, a research group, black Americans, Hispanic Americans, Asian Americans, Native Americans, and Pacific Islanders are 38% likelier than white people to be infected with covid-19, even after controlling for income, underlying health conditions and other variables that potentially affect the disease. Its analysis suggests that a non-white 18-year-old is indeed likelier to contract covid-19 than a white 50-year-old. Vaccinating America’s non-white population is therefore crucial

Gender and Covid

Gender in Covid

Logistic regression modeling Covid infection

N= 20,004,824 patients

Covid patients out of the above: 363,279

Controlling for age, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions and when compared to females,

- there is 9% increase in the odds of contracting Covid for males

Logistic regression modeling death in the face of Covid, or death given Covid

N= 363,279 Covid patients

Deaths: 23,329

Controlling for age, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions and when compared to females,

- there is 61% increase in the odds of death in the face of Covid for males

Gender in all cause deaths

Parameter		DF	Estimate	Std err	Chi-Sq	p-value	Effect	OR	CI	CI
AGE		1	0.0871	0.0003	85270.544	<.0001	AGE	1.091	1.09	1.092
GENDER	2	1	0.1339	0.0026	2904.1594	<.0001	GENDER 2 vs 1	1.323	1.309	1.336
RACE	0	1	-0.128	0.0134	91.5552	<.0001	RACE 0 vs 1	0.666	0.648	0.685
RACE	2	1	0.236	0.00927	648.0641	<.0001	RACE 2 vs 1	0.959	0.944	0.974
RACE	3	1	0.3047	0.0128	563.8315	<.0001	RACE 3 vs 1	1.027	1	1.055
RACE	4	1	-0.2478	0.0273	82.5779	<.0001	RACE 4 vs 1	0.591	0.555	0.63
RACE	5	1	-0.4429	0.0139	1011.0077	<.0001	RACE 5 vs 1	0.486	0.472	0.501
JOB	3	1	-0.0712	0.00989	51.7738	<.0001	JOB 3 vs 7	0.906	0.881	0.932
JOB	4	1	0.1379	0.0084	269.7196	<.0001	JOB 4 vs 7	1.117	1.09	1.146
JOB	6	1	-0.1644	0.0311	28.0202	<.0001	JOB 6 vs 7	0.826	0.767	0.889
JOB	10	1	0.0836	0.0119	56.8783	<.0001	JOB 10 vs 7	1.065	1.032	1.099
JOB	14	1	0.066	0.00744	78.835	<.0001	JOB 14 vs 7	1.04	1.016	1.064
AIR_EMPLOYMENT		1	0.0177	0.00269	43.3512	<.0001	AIR_EMPLOYMENT	1.018	1.013	1.023
OVERWEIGHT	0	1	-0.0533	0.00248	461.3957	<.0001	OVERWEIGHT 0 vs 1	0.899	0.89	0.908
POLITICS	2	1	-0.0309	0.00347	79.0446	<.0001	POLITICS 2 vs 1	0.961	0.951	0.972
POLITICS	3	1	0.0224	0.00321	48.6641	<.0001	POLITICS 3 vs 1	1.014	1.004	1.025
EDUCATION	1	1	0.0216	0.00482	19.9725	<.0001	EDUCATION 1 vs 4	1.047	1.028	1.065
EDUCATION	2	1	0.0257	0.00311	68.3211	<.0001	EDUCATION 2 vs 4	1.051	1.036	1.066
EDUCATION	3	1	-0.0233	0.00382	37.3782	<.0001	EDUCATION 3 vs 4	1.001	0.985	1.016
RELIGION	2	1	-0.2187	0.0126	299.9695	<.0001	RELIGION 2 vs 1	0.801	0.779	0.824
RELIGION	4	1	0.135	0.00771	306.6585	<.0001	RELIGION 4 vs 1	1.141	1.128	1.154
RELIGION	5	1	0.1188	0.00804	218.1674	<.0001	RELIGION 5 vs 1	1.123	1.108	1.137
HOMEOWNER	0	1	0.1129	0.00262	1849.9595	<.0001	HOMEOWNER 0 vs 1	1.253	1.24	1.266
HOMEVALUE	1	1	0.3046	0.0048	4021.5545	<.0001	HOMEVALUE 1 vs 4	1.903	1.862	1.944
HOMEVALUE	2	1	0.1268	0.00435	848.0094	<.0001	HOMEVALUE 2 vs 4	1.593	1.562	1.624
HOMEVALUE	3	1	-0.0677	0.00436	185.9372	<.0001	HOMEVALUE 3 vs 4	1.311	1.291	1.332
HS_PRIMARY_LANG_SPAN		1	-0.058	0.00301	371.1549	<.0001	HS_PRIMARY_LANG_SPAN	0.944	0.938	0.949
PRIMARY_CARE	0	1	0.039	0.0024	263.1295	<.0001	PRIMARY_CARE 0 vs 1	1.081	1.071	1.091
MED_UTILIZATION	0	1	0.015	0.00303	24.4372	<.0001	MED_UTILIZATION 0 vs 1	1.03	1.018	1.043
OS_RISKTAKING_HEALTH		1	0.0392	0.00167	551.2526	<.0001	OS_RISKTAKING_HEALTH	1.04	1.037	1.043
JUNK_DIET	0	1	-0.1047	0.00288	1320.1223	<.0001	JUNK_DIET 0 vs 1	0.811	0.802	0.82
STRESSED	0	1	-0.1305	0.00552	558.907	<.0001	STRESSED 0 vs 1	0.77	0.754	0.787
HEAVY_DRINKER	0	1	0.0248	0.00454	29.7515	<.0001	HEAVY_DRINKER 0 vs 1	1.051	1.032	1.07
DAILY_SMOKER	0	1	-0.1147	0.00379	916.4555	<.0001	DAILY_SMOKER 0 vs 1	0.795	0.783	0.807
UMC_30		1	0.1116	0.00077	20892.54	<.0001	UMC_30	1.118	1.116	1.12
HOMEVALUE	0	1	-0.0251	0.00666	14.2012	0.0002	HOMEVALUE 0 vs 4	1.368	1.336	1.401
JOB	5	1	-0.0853	0.0255	11.2197	0.0008	JOB 5 vs 7	0.894	0.841	0.95
HEALTH_INSURANCE	0	1	0.0202	0.00613	10.8175	0.001	HEALTH_INSURANCE 0 vs 1	1.041	1.016	1.067
HISPANIC_PROF	0	1	-0.0214	0.00956	5.0292	0.0249	HISPANIC_PROF 0 vs 1	0.958	0.923	0.995
POOR_SLEEP_QUALITY	0	1	-0.0046	0.00233	3.9044	0.0482	POOR_SLEEP_QUALITY 0 vs 1	0.991	0.982	1
RELIGION	3	1	-0.0384	0.0254	2.2748	0.1315	RELIGION 3 vs 1	0.959	0.901	1.021
AFRICAN_AMERICAN_PRO	0	1	0.00629	0.0115	0.2971	0.5857	AFRICAN_AMERICAN_PRO 0 vs 1	1.013	0.968	1.06
DAILY_EXERCISE	0	1	-0.0022	0.00407	0.2932	0.5882	DAILY_EXERCISE 0 vs 1	0.996	0.98	1.012
INCOME	2	1	0.8792	13.8739	0.004	0.9495	INCOME 2 vs 4	1.083	1.061	1.105
INCOME	3	1	0.8626	13.8739	0.0039	0.9504	INCOME 3 vs 4	1.065	1.048	1.082
INCOME	1	1	0.8546	13.8739	0.0038	0.9509	INCOME 1 vs 4	1.056	1.033	1.08
INCOME	0	1	-3.3962	55.4958	0.0037	0.9512	INCOME 0 vs 4	0.015	<.0001	>999.999

Gender in Covid

Odds Ratio

Odds for men dying = Number of men died/Number of men alive = 0.019866

Odds for women dying = Number of women died/Number of women alive = 0.013712

Odds Ratio for men dying = Odds for men/Odds for women = **1.45**

Risk Ratio

Risk for men of dying: Number of men died/Total number of men = 0.019479

Risk for women of dying: Number of women died/Total number of women = 0.013526

Risk Ratio for men vs. women of dying from all causes = Risk for men/Risk for women = **1.44**

Risk for men of dying from COVID = 0.001510

Risk for women of dying from COVID = 0.000905

Risk Ratio for men vs women of dying from COVID: **1.67**

Difference in Risk Ratio (COVID related risk increase; all cause deaths compared to covid deaths)

= 1.67 - 1.44 = **0.23** i.e., 23 percentage points or 15.97%

Gender in Covid

Conclusion

- Men have 44% greater risk of dying than women
 - Men have 67% greater risk of dying of Covid than women
 - The risk difference is 23 percentage points i.e., 15.97%
 - These are raw numbers, not controlling for confounders
 - The outcome prevalence is less than 10%
 - Odds ratio approximates the risk ratio when the outcome is rare
 - In such a case logistic regression works
 - If the outcome were not rare, relative risk regression would be preferred over logistic regression for direct estimation of the risk ratio
 - There is 30% increase in the odds for men of dying of Covid controlling for the relevant confounders in the data
 - One can also say, there is 30% increase in such a risk
-
- The raw risk difference is 23 percentage points or 15.97% when **not** controlling for confounders
 - It is 30% when controlling for the confounders
 - **Men are at a greater risk of dying of Covid than women**

Research Database Consortium Common Profile Linked Dataset

Racial composition of the linked dataset*

RACE	COUNT	%
White	5,856,747	62.0%
Hispanic	2,012,038	21.3%
Black	718,953	7.6%
Unknown	705,440	7.5%
Asians	96,359	1.0%
Native Americans	56,677	0.6%

*Where race is not NULL and patient was active in 2020

Distribution of the U.S. population by race and ethnicity (2019):

- White: 60.1% (*Non-Hispanic*)
- Hispanic: 18.5%
- Black: 12.2%
- Asian: 5.6%
- Multiple Races: 2.8%
- American Indian/Alaska Native: 0.7%
- Native Hawaiian/Other Pacific Islander: 0.2%

Source: US Census Bureau

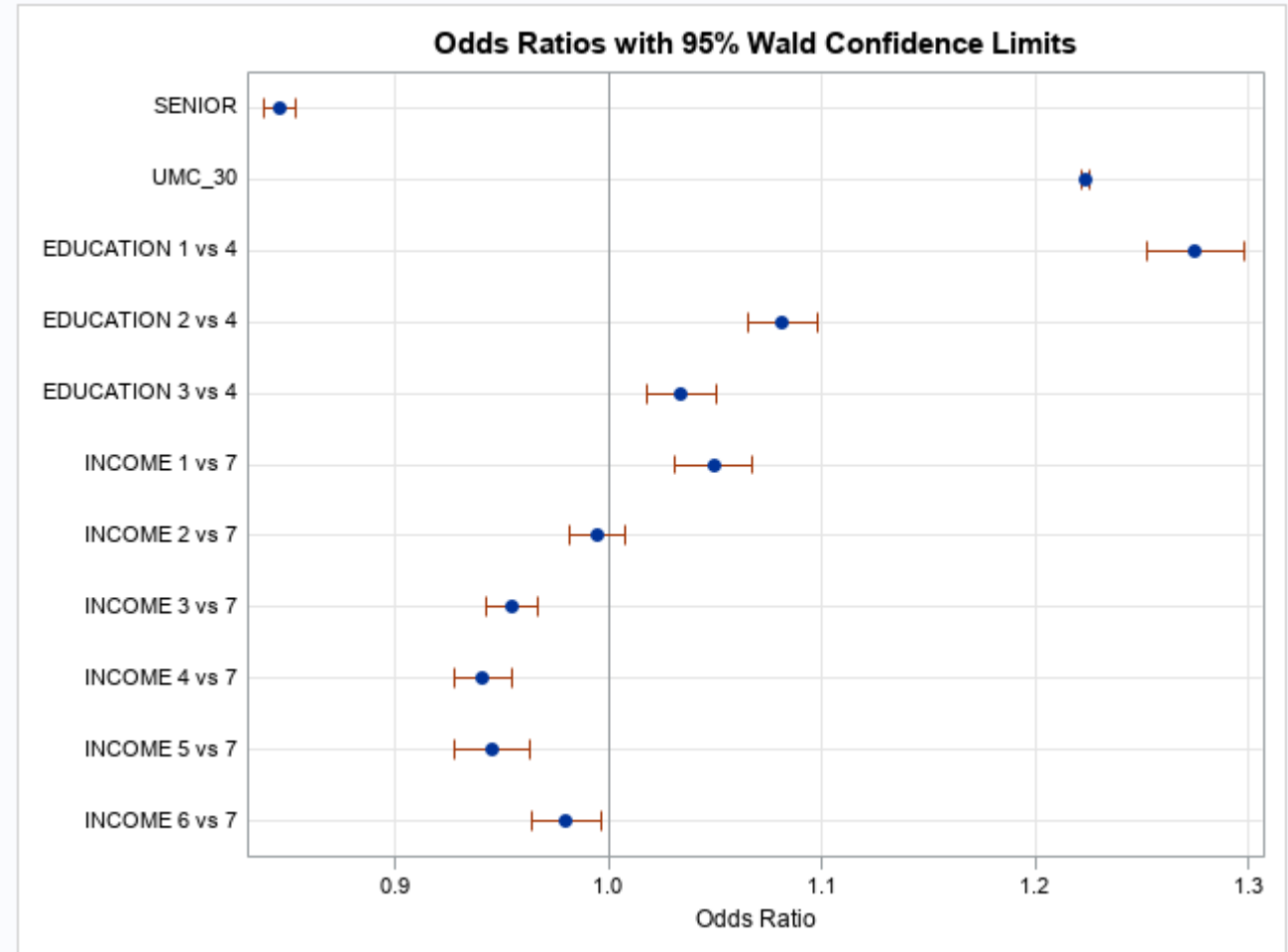
- Asian and Black populations are underrepresented in the linked dataset
- The dataset is not representative of the US population
- It comprises patients enrolled in doctors' offices that are served by a claims clearinghouse and EHR vendor

Risk of contracting Covid

Number of observations used: 9,446,214
Number contracted Covid in 2020: 258,814 (2.74%)

```
proc logistic data=DATA2020 descending;
class GENDER (ref='1') / param=ref;
class RACE (ref='1') / param=ref;
class INCOME (ref='7') / param=ref;
class EDUCATION (ref='4') / param=ref;
model COVID = SENIOR GENDER RACE UMC_30 EDUCATION INCOME GENDER*RACE;
oddsratio 'MALE VS FEMALE WHITE' GENDER / at(RACE='1' GENDER='2') CL=PL;
oddsratio 'MALE VS FEMALE AA' GENDER / at(RACE='2' GENDER='2') CL=PL;
oddsratio 'MALE VS FEMALE HISPANIC' GENDER / at(RACE='3' GENDER='2') CL=PL;
oddsratio 'M VS F NATIVE AMERICAN' GENDER / at(RACE='4' GENDER='2') CL=PL;
oddsratio 'MALE VS FEMALE ASIAN' GENDER / at(RACE='5' GENDER='2') CL=PL;
RUN;
```

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
SENIOR	0.846	0.838	0.853
UMC_30	1.224	1.222	1.226
EDUCATION 1 vs 4	1.275	1.252	1.298
EDUCATION 2 vs 4	1.081	1.065	1.098
EDUCATION 3 vs 4	1.034	1.018	1.050
INCOME 1 vs 7	1.049	1.031	1.068
INCOME 2 vs 7	0.995	0.982	1.008
INCOME 3 vs 7	0.955	0.943	0.967
INCOME 4 vs 7	0.941	0.927	0.955
INCOME 5 vs 7	0.945	0.928	0.963
INCOME 6 vs 7	0.980	0.964	0.997



*Education: Less than high school=1, High school=2, Bachelor degree=3, and Graduate degree=4

*Income: Under \$24,999=1, \$25,000 to \$49,999=2, \$50,000 to \$74,999=3, \$75,000 to \$99,999=4, \$100,000 to \$124,999=5, \$125,000 to \$149,999=6, and Over \$150,000=7

*Race: Unknown=0, White=1, African American=2, Hispanic=3, Native American and Pacific islander=4, and Asian=5

*Gender: Female=1, Male=2

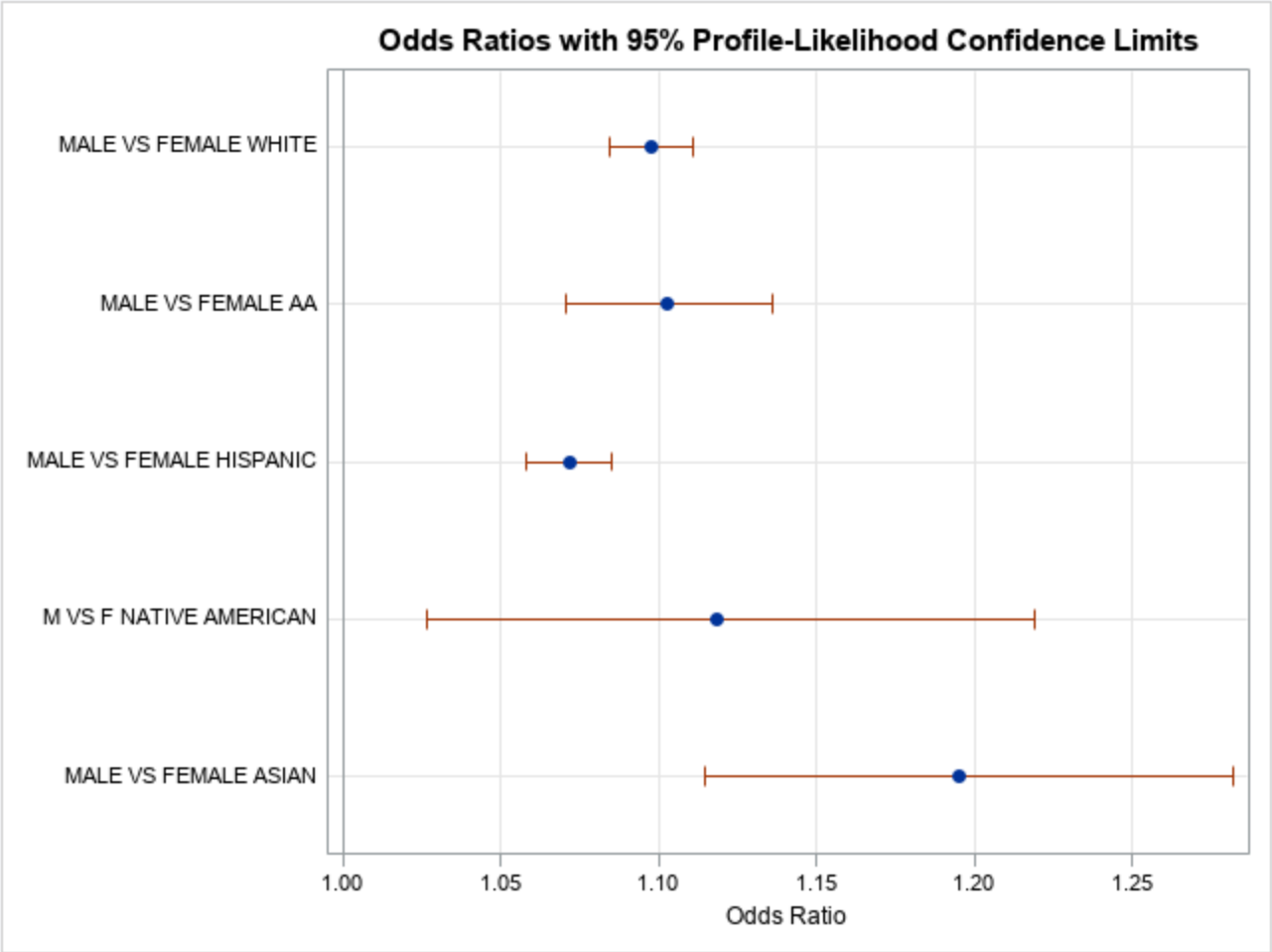
*Senior: Age 65 or above=1, Otherwise 0

*UMC_30: Count of underlying medical conditions out of total 30 chronic conditions that are risk factor for severe outcomes given Covid

Risk of contracting Covid, by gender

In all racial groups, males are likelier to contract Covid than females

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Label	Odds Ratio	Estimate	95% Confidence Limits	
MALE VS FEMALE WHITE	GENDER 2 vs 1 at RACE=1	1.097	1.084	1.111
MALE VS FEMALE AA	GENDER 2 vs 1 at RACE=2	1.103	1.070	1.136
MALE VS FEMALE HISPANIC	GENDER 2 vs 1 at RACE=3	1.072	1.058	1.085
M VS F NATIVE AMERICAN	GENDER 2 vs 1 at RACE=4	1.119	1.026	1.219
MALE VS FEMALE ASIAN	GENDER 2 vs 1 at RACE=5	1.195	1.115	1.282

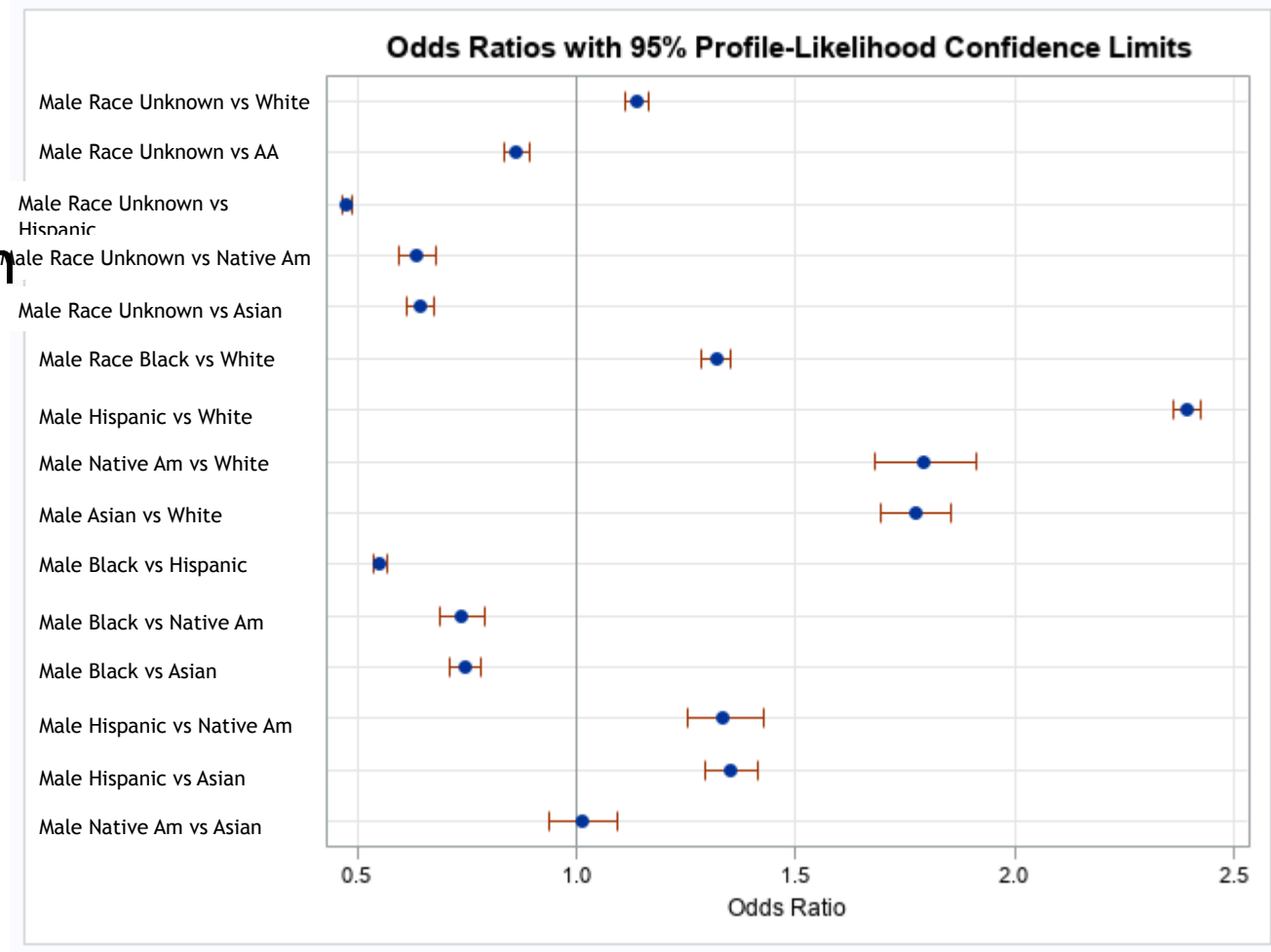


Risk of contracting Covid

Number of observations used: 9,446,214
Number contracted Covid in 2020: 258,814 (2.74%)

Risk of contracting Covid, by race

Males of minority racial groups
are likelier to contract Covid than
White males



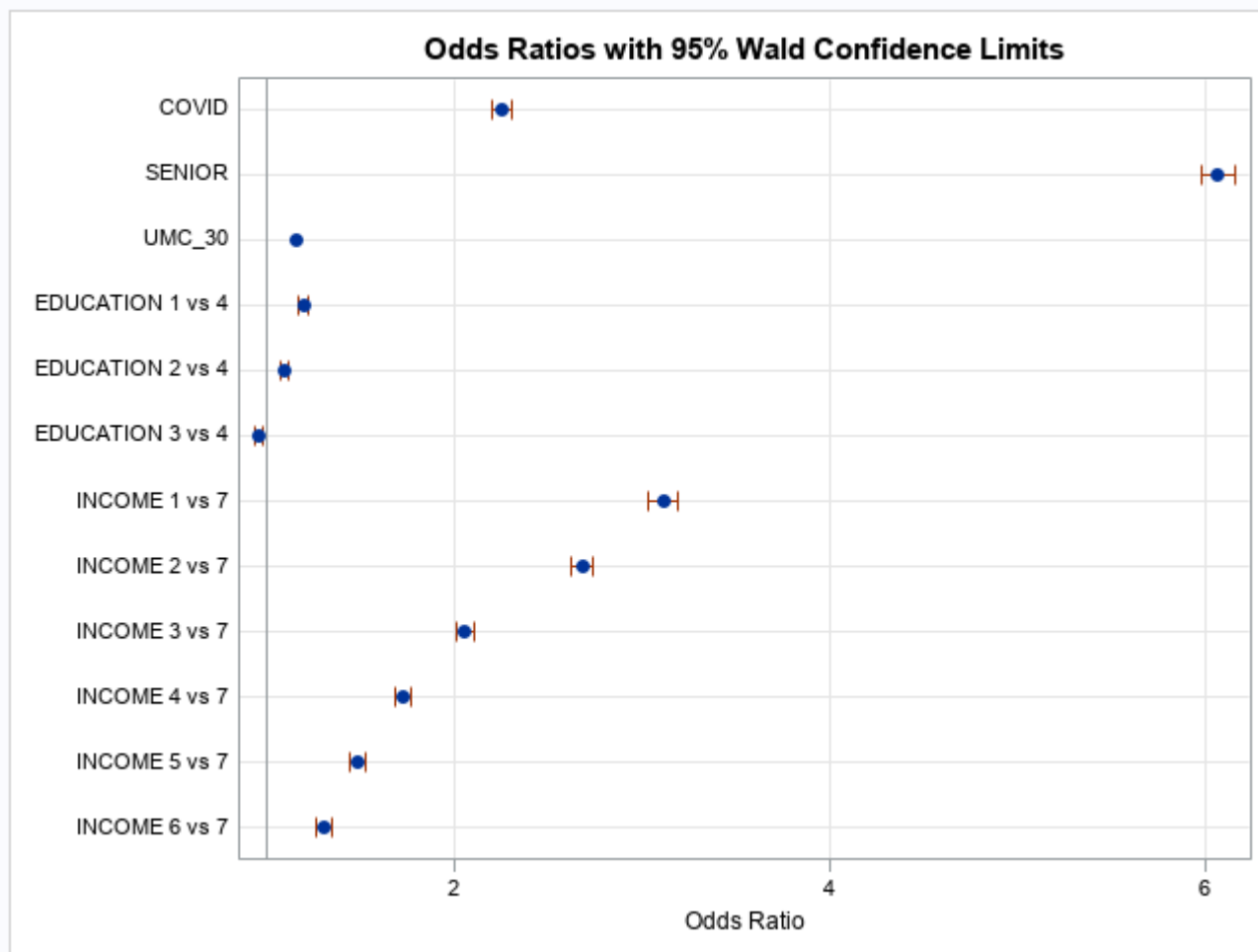
```
proc logistic data=DATA2020 descending;
class GENDER (ref='1') / param=ref;
class RACE (ref='1') / param=ref;
class INCOME (ref='7') / param=ref;
class EDUCATION (ref='4') / param=ref;
model DIED2020 = COVID SENIOR GENDER RACE UMC_30 EDUCATION INCOME GENDER*RACE;
oddsratio 'MALE vs FEMALE WHITE' GENDER / at(RACE='1' GENDER='2') CL=PL;
oddsratio 'MALE vs FEMALE AA' GENDER / at(RACE='2' GENDER='2') CL=PL;
oddsratio 'MALE vs FEMALE HISPANIC' GENDER / at(RACE='3' GENDER='2') CL=PL;
oddsratio 'M vs F NATIVE AMERICAN' GENDER / at(RACE='4' GENDER='2') CL=PL;
oddsratio 'MALE vs FEMALE ASIAN' GENDER / at(RACE='5' GENDER='2') CL=PL;
RUN;
```

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
COVID	2.253	2.204	2.303
SENIOR	6.064	5.974	6.156
UMC_30	1.155	1.153	1.158
EDUCATION 1 vs 4	1.196	1.167	1.226
EDUCATION 2 vs 4	1.097	1.075	1.119
EDUCATION 3 vs 4	0.957	0.936	0.977
INCOME 1 vs 7	3.111	3.034	3.190
INCOME 2 vs 7	2.680	2.622	2.740
INCOME 3 vs 7	2.054	2.008	2.101
INCOME 4 vs 7	1.724	1.681	1.768
INCOME 5 vs 7	1.485	1.443	1.529
INCOME 6 vs 7	1.306	1.267	1.346

Risk of death in 2020

Number of observations used: 9,446,214

Number died in 2020: 135,067 (1.43%)



*Education: Less than high school=1, High school=2, Bachelor degree=3, and Graduate degree=4

*Income: Under \$24,999=1, \$25,000 to \$49,999=2, \$50,000 to \$74,999=3, \$75,000 to \$99,999=4, \$100,000 to \$124,999=5, \$125,000 to \$149,999=6, and Over \$150,000=7

*Race: Unknown=0, White=1, African American=2, Hispanic=3, Native American and Pacific islander=4, and Asian=5

*Gender: Female=1, Male=2

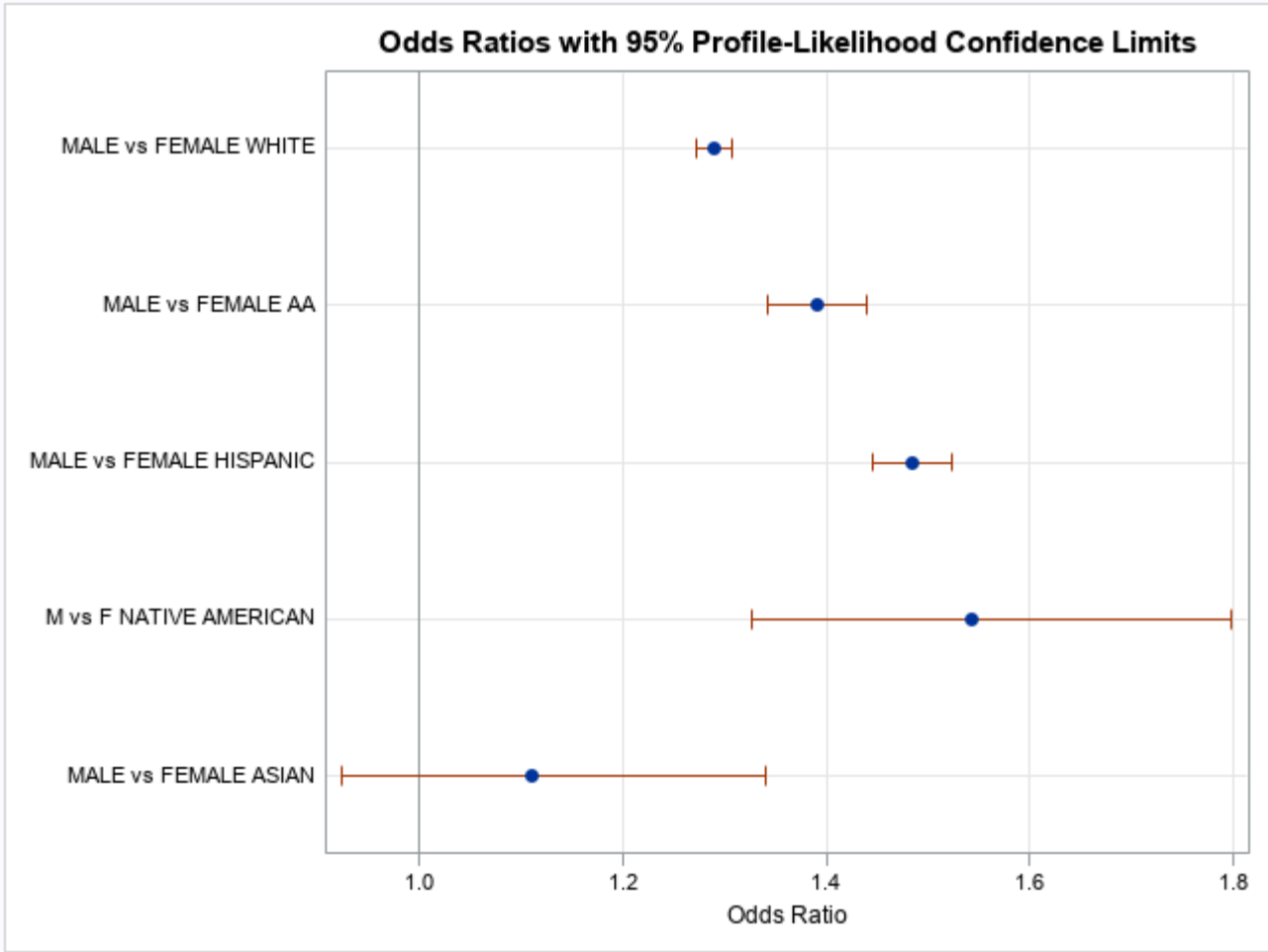
*Senior: Age 65 or above=1, Otherwise 0

*UMC_30: Count of underlying medical conditions out of total 30 chronic conditions that are risk factor for severe outcomes given Covid

Risk of death in 2020,
by gender

In all racial groups except
Asians, males were likelier to
die in 2020 than females

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Label	Odds Ratio	Estimate	95% Confidence Limits	
MALE vs FEMALE WHITE	GENDER 2 vs 1 at RACE=1	1.290	1.273	1.307
MALE vs FEMALE AA	GENDER 2 vs 1 at RACE=2	1.390	1.342	1.440
MALE vs FEMALE HISPANIC	GENDER 2 vs 1 at RACE=3	1.485	1.446	1.524
M vs F NATIVE AMERICAN	GENDER 2 vs 1 at RACE=4	1.544	1.327	1.797
MALE vs FEMALE ASIAN	GENDER 2 vs 1 at RACE=5	1.110	0.923	1.340

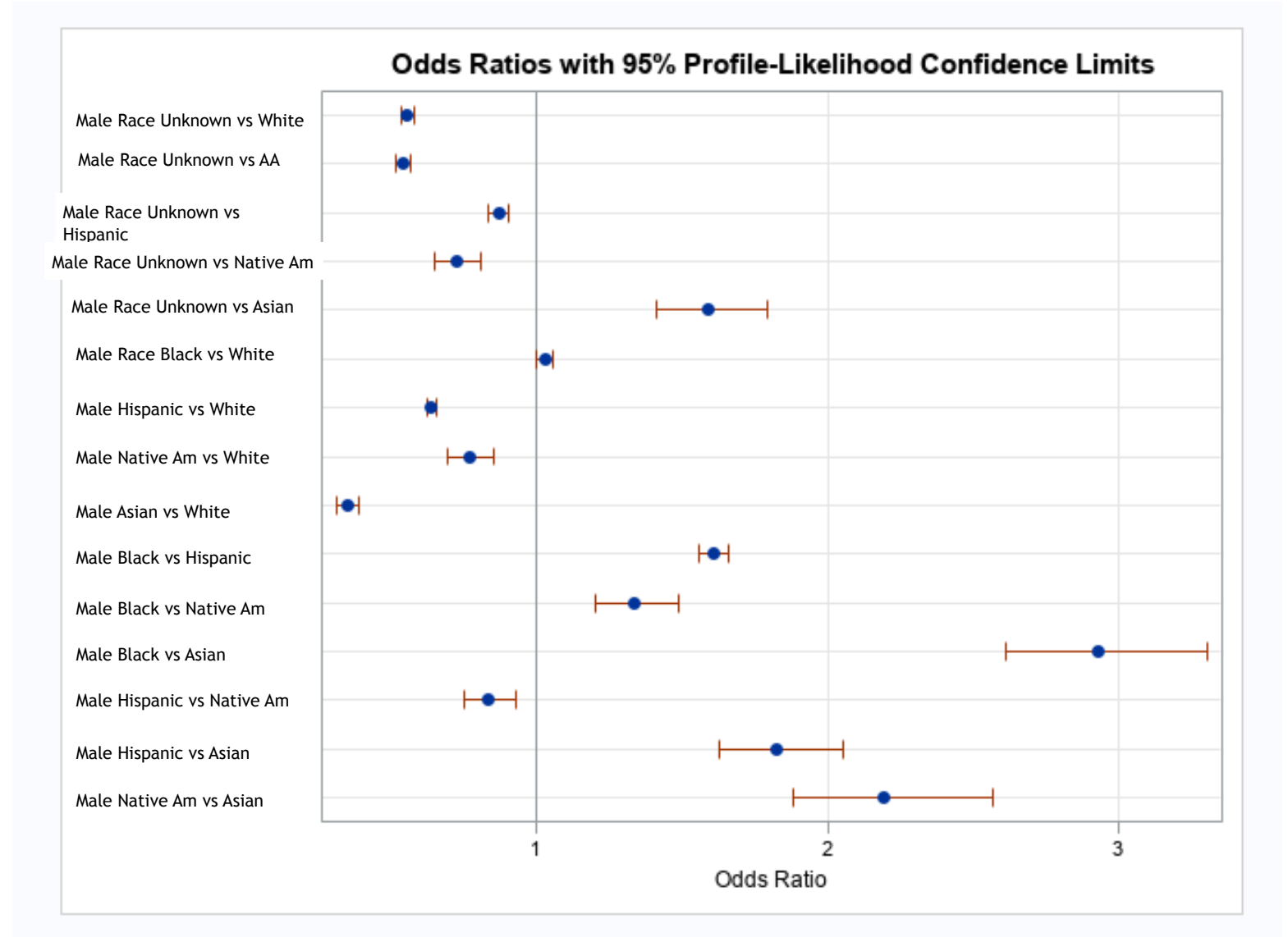


Risk of death in 2020

Number of observations used: 9,446,214
Number died in 2020: 135,067 (1.43%)

Risk of death in 2020,
by race

Black males were equally
likely to die in 2020 as their
White counterparts



Risk of death in 2020

Number of observations used: 9,446,214

Number died in 2020: 135,067 (1.43%)

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	174751.206	35	<.0001
Score	194871.175	35	<.0001
Wald	137106.854	35	<.0001

Joint Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
COVID	1	557.3104	<.0001
SENIOR	1	4251.0944	<.0001
GENDER	1	219.9255	<.0001
RACE	5	357.0190	<.0001
UMC_30	1	120.2393	<.0001
EDUCATION	3	10.5229	0.0146
INCOME	6	701.4263	<.0001
COVID*SENIOR	1	39.0688	<.0001
COVID*GENDER	1	2.5071	0.1133
COVID*RACE	5	26.5357	<.0001
UMC_30*COVID	1	788.9770	<.0001
COVID*EDUCATION	3	14.9084	0.0019
COVID*INCOME	6	19.2187	0.0038

*Education: Less than high school=1, High school=2, Bachelor degree=3, and Graduate degree=4

*Income: Under \$24,999=1, \$25,000 to \$49,999=2, \$50,000 to \$74,999=3, \$75,000 to \$99,999=4, \$100,000 to \$124,999=5, \$125,000 to \$149,999=6, and Over \$150,000=7

*Race: Unknown=0, White=1, African American=2, Hispanic=3, Native American and Pacific islander=4, and Asian=5

*Gender: Female=1, Male=2

*Senior: Age 65 or above=1, Otherwise 0

*UMC_30: Count of underlying medical conditions out of total 30 chronic conditions that are risk factor for severe outcomes given Covid

Parameter			DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Odds Ratio
COVID	0		1	-1.1147	0.0589	357.8496	<.0001	0.33
RACE	5		1	-0.9922	0.1524	42.4	<.0001	0.37
RACE	0		1	-0.4404	0.0516	72.7719	<.0001	0.64
RACE	3		1	-0.4397	0.0259	287.1882	<.0001	0.64
RACE	4		1	-0.3197	0.1255	6.4933	0.0108	0.73
SENIOR*COVID	0		1	-0.1981	0.0315	39.5809	<.0001	0.82
COVID*RACE	0	0	1	-0.1253	0.0535	5.4897	0.0191	0.88
RACE	2		1	-0.1101	0.0387	8.1043	0.0044	0.90
COVID*RACE	0	3	1	-0.0887	0.0273	10.5536	0.0012	0.92
SENIOR			1	1.9866	0.0305	4251.9874	<.0001	7.29
INCOME	1		1	1.053	0.0482	477.6008	<.0001	2.87
INCOME	2		1	0.9231	0.0428	466.1771	<.0001	2.52
INCOME	3		1	0.7005	0.0446	246.1939	<.0001	2.01
INCOME	4		1	0.571	0.049	135.7347	<.0001	1.77
INCOME	5		1	0.473	0.0562	70.7628	<.0001	1.60
GENDER	2		1	0.3137	0.0211	220.0745	<.0001	1.37
INCOME	6		1	0.3002	0.0588	26.0331	<.0001	1.35
EDUCATION	1		1	0.1249	0.0474	6.9299	0.0085	1.13
UMC_30*COVID	0		1	0.1112	0.00396	788.5009	<.0001	1.12
COVID*RACE	0	2	1	0.1005	0.04	6.3195	0.0119	1.11
UMC_30			1	0.0418	0.00381	120.3382	<.0001	1.04
EDUCATION	2		1	0.0796	0.0414	3.7042	0.0543	1.08
COVID*INCOME	0	1	1	0.0895	0.05	3.2093	0.0732	1.09
COVID*EDUCATION	0	3	1	-0.0745	0.0456	2.6712	0.1022	0.93
COVID*GENDER	0	2	1	-0.0349	0.0219	2.5285	0.1118	0.97
COVID*INCOME	0	2	1	0.0695	0.0443	2.4647	0.1164	1.07
COVID*INCOME	0	5	1	-0.0815	0.0583	1.954	0.1622	0.92
COVID*EDUCATION	0	1	1	0.0623	0.0492	1.602	0.2056	1.06
EDUCATION	3		1	0.0271	0.0442	0.3753	0.5401	1.03
COVID*INCOME	0	6	1	-0.035	0.061	0.3291	0.5662	0.97
COVID*INCOME	0	3	1	0.0241	0.0462	0.2726	0.6016	1.02
COVID*INCOME	0	4	1	-0.0252	0.0508	0.2467	0.6194	0.98
COVID*EDUCATION	0	2	1	0.0156	0.0427	0.1327	0.7157	1.02
COVID*RACE	0	4	1	-0.0372	0.1319	0.0795	0.7779	0.96
COVID*RACE	0	5	1	-0.0107	0.1599	0.0045	0.9466	0.99

Age and Covid

Prior to Covid

Odds of death prior to Covid Year 2019

N=19,483,137 patients active in the claims dataset linked with social data in 2019
Died in 2019: 140,193 (0.72%)

We estimated the odds of death in 2019 using logistic regression and age, race, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions as determinants.

I find, the odds of death were

- 92% higher for population in the lowest quartile, 58% higher for those in the second lowest quartile, 30% higher for those in the third quartile of home value when compared to those in the highest quartile
- 30% higher for men than women
- 9% higher for each additional underlying chronic medical condition
- 10% higher for Hispanics than Caucasians
- 9% higher for those who do not have access to primary care vs. those who have
- 7% higher for those who do not have health insurance vs. those who have
- 8% higher for each additional year in age
- 8% higher for grade schoolers vs those with graduate education
- 6% higher for heavy drinkers vs those who are not
- 5% higher for risk takers in health vs those who are not
- 22% lower for those who do not consume junk diet vs those who do
- 17% lower for non-daily smokers vs those who are
- 16% lower for non-overweight vs those who are
- 8% lower for non-stressed vs those who are stressed
- 6% lower for African Americans vs Caucasians

These findings are statistically significant at 99.99% confidence level (p-value =< 0.0001).

2019

Seniors

N=6,304,287 patients active in the claims dataset linked with social data in 2019
Died in 2019: 110,701 (1.76%)

The odds of death were

- 105% higher for population in the lowest quartile, 69% higher for those in the second lowest quartile, 33% higher for those in the third quartile of home value when compared to those in the highest quartile
- 28% higher for men than women
- 9% higher for each additional underlying chronic medical condition
- 14% higher for Hispanics than Caucasians
- 10% higher for those who do not have access to primary care vs. those who have
- 9% higher for each additional year in age
- 6% higher for grade schoolers vs those with graduate education
- 10% higher for heavy drinkers vs those who are not
- 5% higher for risk takers in health vs those who are not
- 23% lower for those who do not consume junk diet vs those who do
- 18% lower for non-overweight vs those who are
- 16% lower for non-daily smokers vs those who are
- 8% lower for African Americans vs Caucasians

These findings are statistically significant at 99.99% confidence level (p-value =< 0.0001).

2019

Age below 65

N=13,529,742 patients active in the claims dataset linked with social data in 2019
Died in 2019: 40,434 (0.30%)

The odds of death were

- 68% higher for population in the lowest quartile, 40% higher for those in the second lowest quartile, 22% higher for those in the third quartile of home value when compared to those in the highest quartile
- 58% higher for men than women
- 17% higher for each additional underlying chronic medical condition
- 17% higher for Hispanics than Caucasians
- 3% higher for those who do not have access to primary care vs. those who have
- 7% higher for each additional year in age
- 9% higher for grade schoolers vs those with graduate education
- 6% higher for heavy drinkers vs those who are not
- 3% higher for risk takers in health vs those who are not
- 17% lower for those who do not consume junk diet vs those who do
- 11% lower for non-overweight vs those who are
- 14% lower for non-daily smokers vs those who are
- 4% lower for African Americans vs Caucasians

These findings are statistically significant at 99.99% confidence level (p-value =< 0.0001).

After Covid

Year 2020

N=19,853,689 patients active in the claims dataset linked with social data in 2020

Died in 2020: 170,795 (0.86%)

Controlling for age, race, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions,

- The odds of death are **68% lower for those who did not have Covid vs those who had**
- This finding is statistically significant at 99.99% confidence level (p-value ≤ 0.0001)

2020

Seniors

N=6,334,552 patients active in the claims dataset linked with social data in 2020

Died in 2020: 140,966 (**2.22%**)

Controlling for age, race, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions,

- The odds of death are **68% lower for those who did not have Covid vs those who had**
- This finding is statistically significant at 99.99% confidence level (p-value =< 0.0001)

Age below 65

N=13,519,137 patients active in the claims dataset linked with social data in 2020

Died in 2020: 29,829 (**0.22%**)

Controlling for age, race, income, employment, job, politics, education, religion, home ownership, home value, professional work status, health insurance, health behaviors, health service utilization, and underlying medical conditions,

- The odds of death are 64% lower for those who did not have Covid vs those who had
- This finding is statistically significant at 99.99% confidence level (p-value =< 0.0001)