

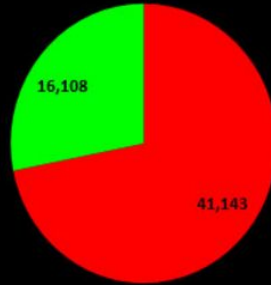
Predicting the Risk of Suicide

Analyzing the Text of Clinical Notes

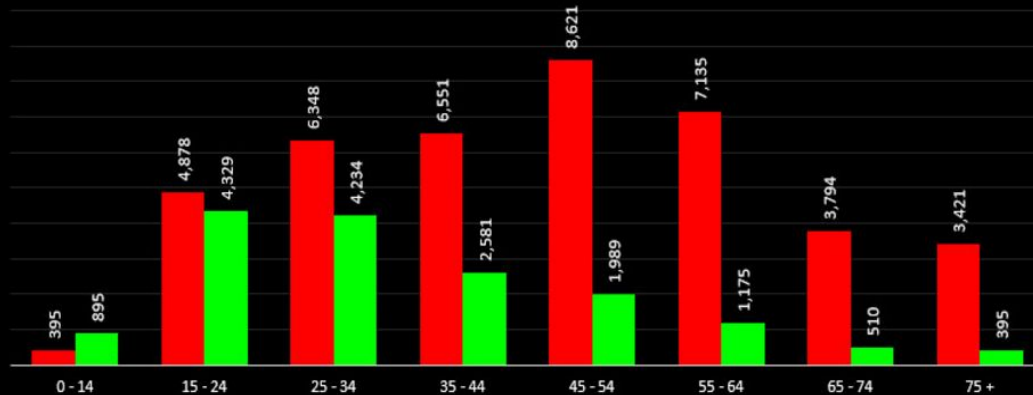
USA HOMICIDE VS SUICIDE

As America contemplates what appears to be an unprecedented wave of violence related deaths in other parts of the world here at home we continue to kill ourselves far more often than we kill each other. Below are total US deaths for selected ages for Homicide and Suicide. Official CDC final deaths for 2014 are used. Visit USA Health Rankings to see state rankings.

USA Total Deaths: Homicide - Suicide

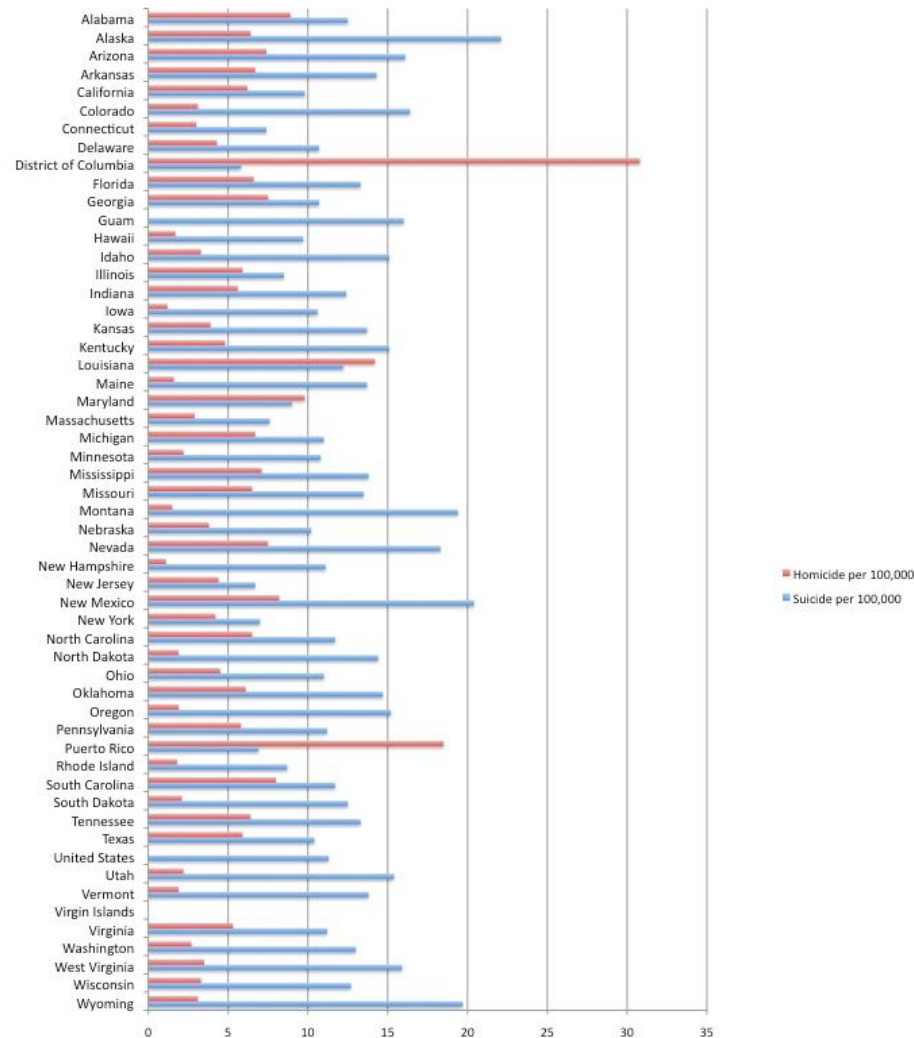


■ SUICIDE ■ HOMICIDE



Homicide vs Suicide

Suicide & Homicide by State



10 Leading Causes of Death by Age Group, United States – 2014

Rank	Age Groups										Total
	<1	1-4	5-9	10-14	15-24	25-34	35-44	45-54	55-64	65+	
1	Congenital Anomalies 4,746	Unintentional Injury 1,216	Unintentional Injury 730	Unintentional Injury 750	Unintentional Injury 11,836	Unintentional Injury 17,357	Unintentional Injury 16,048	Malignant Neoplasms 44,834	Malignant Neoplasms 115,282	Heart Disease 489,722	Heart Disease 614,348
2	Short Gestation 4,173	Congenital Anomalies 399	Malignant Neoplasms 436	Suicide 425	Suicide 5,079	Suicide 6,569	Malignant Neoplasms 11,267	Heart Disease 34,791	Heart Disease 74,473	Malignant Neoplasms 413,885	Malignant Neoplasms 591,699
3	Maternal Pregnancy Comp. 1,574	Homicide 364	Congenital Anomalies 192	Malignant Neoplasms 416	Homicide 4,144	Homicide 4,159	Heart Disease 10,368	Unintentional Injury 20,610	Unintentional Injury 18,030	Chronic Low. Respiratory Disease 124,693	Chronic Low. Respiratory Disease 147,101
4	SIDS 1,545	Malignant Neoplasms 321	Homicide 123	Congenital Anomalies 156	Malignant Neoplasms 1,569	Malignant Neoplasms 3,624	Suicide 6,706	Suicide 8,767	Chronic Low. Respiratory Disease 16,492	Cerebro-vascular 113,308	Unintentional Injury 136,053
5	Unintentional Injury 1,161	Heart Disease 149	Heart Disease 69	Homicide 156	Heart Disease 953	Heart Disease 3,341	Homicide 2,588	Liver Disease 8,627	Diabetes Mellitus 13,342	Alzheimer's Disease 92,604	Cerebro-vascular 133,103
6	Placenta Cord. Membranes 965	Influenza & Pneumonia 109	Chronic Low. Respiratory Disease 68	Heart Disease 122	Congenital Anomalies 377	Liver Disease 725	Liver Disease 2,582	Diabetes Mellitus 6,062	Liver Disease 12,792	Diabetes Mellitus 54,161	Alzheimer's Disease 93,541
7	Bacterial Sepsis 544	Chronic Low Respiratory Disease 53	Influenza & Pneumonia 57	Chronic Low Respiratory Disease 71	Influenza & Pneumonia 199	Diabetes Mellitus 709	Diabetes Mellitus 1,999	Cerebro-vascular 5,349	Cerebro-vascular 11,727	Unintentional Injury 48,295	Diabetes Mellitus 76,488
8	Respiratory Distress 460	Septicemia 53	Cerebro-vascular 45	Cerebro-vascular 43	Diabetes Mellitus 181	HIV 583	Cerebro-vascular 1,745	Chronic Low. Respiratory Disease 4,402	Suicide 7,527	Influenza & Pneumonia 44,836	Influenza & Pneumonia 55,227
9	Circulatory System Disease 444	Benign Neoplasms 38	Benign Neoplasms 36	Influenza & Pneumonia 41	Chronic Low Respiratory Disease 178	Cerebro-vascular 579	HIV 1,174	Influenza & Pneumonia 2,731	Septicemia 5,709	Nephritis 39,957	Nephritis 48,146
10	Neonatal Hemorrhage 441	Perinatal Period 38	Septicemia 33	Benign Neoplasms 38	Cerebro-vascular 177	Influenza & Pneumonia 549	Influenza & Pneumonia 1,125	Septicemia 2,514	Influenza & Pneumonia 5,390	Septicemia 29,124	Suicide 42,773

Data Source: National Vital Statistics System, National Center for Health Statistics, CDC.

Produced by: National Center for Injury Prevention and Control, CDC using WISQARS™.



Centers for Disease
Control and Prevention
National Center for Injury
Prevention and Control

The Study

Cohort 1

VA patients who did not use mental health services

Cohort 2

Used VHA National Suicide Registry to obtain a random sample of 100 VHA enrollees who committed suicide in 2009

Cohort 3

VHA patients who were hospitalized in inpatient psychiatric units at least once in 2009 but did not complete suicide

Primary Data

Notes by Nurses, Doctors, and
Other Healthcare Professionals

- Hospitalizations, procedures, surgeries, and other medical services
- Semi-automatic, script-generated tables (e.g. lists of medications)
- Psychological state, depression, alcoholism

Primary Data by Cohort

Cohort 1

1,913 Notes

27 Notes per Patient

Cohort 2

4,243 Notes

61 Notes per Patient

Cohort 3

5,388 Notes

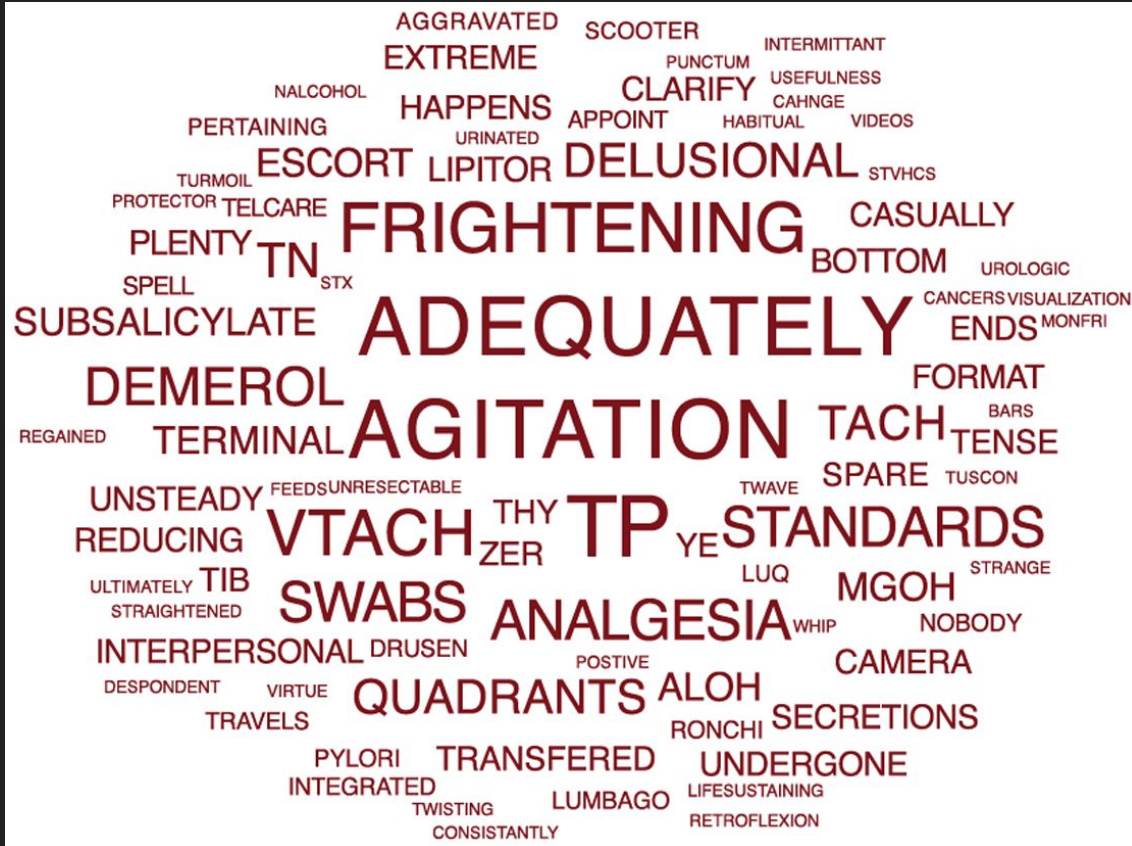
77 Notes per Patient

The Model

Supervised Training with
Genetic Programming

- Recognizes patterns associated with a known outcome
- Bag-of-words/phrases with frequency count in notes
- Words associated with suicide
- Mutual Information (variable co-dependence) for feature selection
- Ensemble of 100 models with 5-fold Cross Validation
- 3-bin classification, e.g. Cohort 1 vs Cohorts 2 & 3

Predictive Words for Cohort 2



More Resources

- [Depressed dogs, cats with OCD — what animal madness means for us humans](#)
- [How Will 'Cut And Paste' Technology Rewrite Our DNA?](#)
- [How CRISPR lets us edit our DNA](#)
- [CRISPR/Cas9 and Targeted Genome Editing: A New Era in Molecular Biology](#)

Thank you