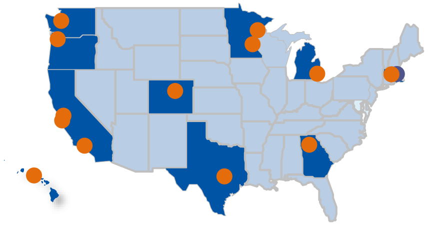
Precision Prediction for Suicide Prevention: Defining the Jobs Before Building the Tools

Greg Simon
Kaiser Permanente Washington Health Research Institute

Mental Health Research Network

KAISER PERMANENTE®



MHRN Mental Health Research Network

KP Washington
- Yates Coley
- Eric Johnson
- Belinda Operskalski
- Robert Penfold
- Julie Richards
- Susan Shortreed
- Chris Stewart
- Rod Walker
- Rob Wellman
- Rebecca Ziebell

Univ. of Washington
- Noah Simon

KP Southern California
- Karen Coleman
- Jean Lawrence
- Tae Yoon

KP Northwest
- Phil Crawford
- Frances Lynch
- Bobbi Jo Yarborough

KP Colorado
- Arne Beck
- Jennifer Boggs
- David Tabano

KP Hawaii
- Yihe Daida
- Beth Waitzfelder
- Carmen Wong

HealthPartners
- Rebecca Rossom
- Sheryl Kane

Henry Ford Health System
- Brian Ahmedani
- Bin Liu

Microsoft Research
- Rich Caruana
- Ran Gilad-Bachrach

Supported by NIMH Cooperative Agreement U19 MH092201 and by FDA Contract under BAA 18-00123

KAISER PERMANENTE®

Outline

- Why precision prediction?
- Four questions:
  - WHO is at risk?
  - WHEN are they at risk?
  - WHY are they at risk?
  - WHAT should we do?
  - (and WHAT could go wrong?)

Why precision prediction?

- Prediction is a precursor for prevention
  - Most relevant to selective or secondary prevention
  - Necessary, but not sufficient

Why precision prediction?

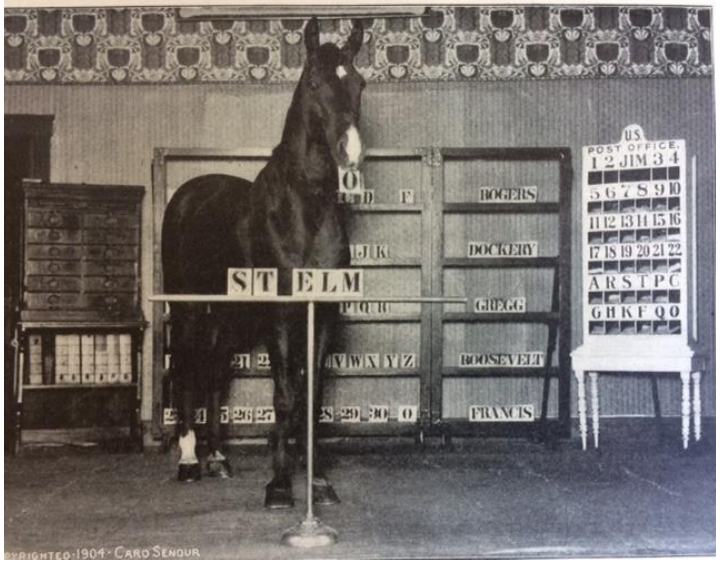
- Prediction is a precursor for prevention
  - Most relevant to selective or secondary prevention
  - Necessary, but not sufficient
- Precision prediction is a precursor for precision prevention
  - Who, When, Why, What

Prediction serves specific "customers"

- Clinical decision or choice point
- Information available
- Potential actions
- Consequences of errors

Tools searching for jobs

Clever Hans, the Calculating Horse



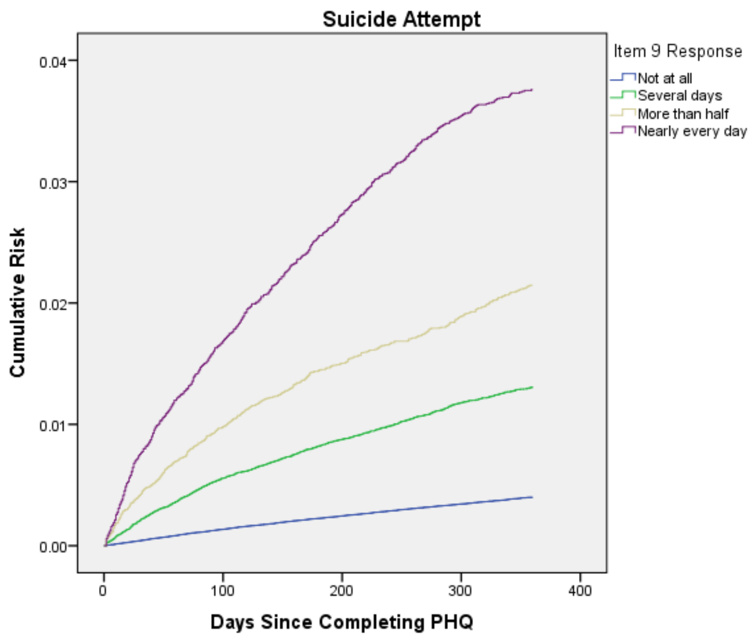
Mining Twitter for winter depression

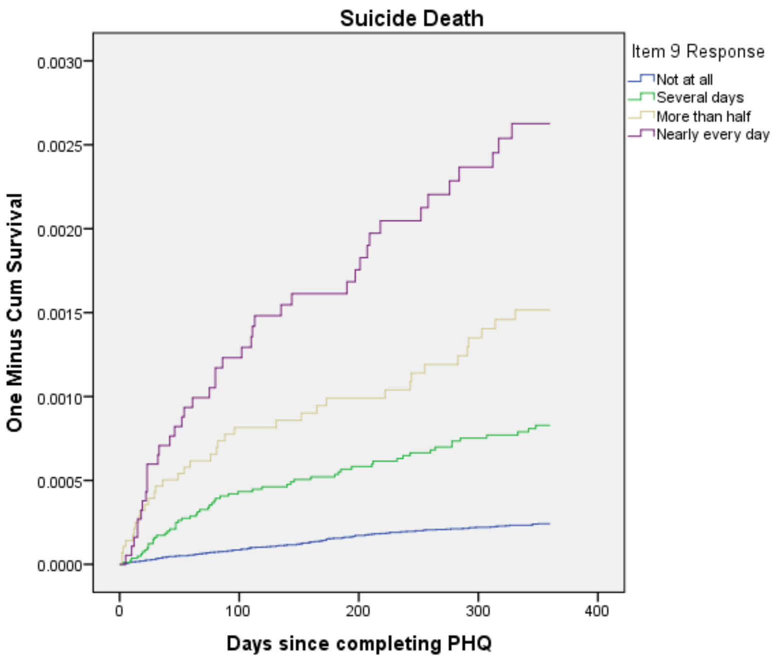
Across the cities, we observe that SMDI in certain cities is more associated with weather conditions than others. Jacksonville and Seattle are both ranked high in terms of depression rates, however the variation in SMDI trend for Seattle (σ^2=4.45) is much higher than that for Jacksonville (σ^2=0.85). In fact, the percent difference between Seattle and Jacksonville's SMDI during winter is 8% higher than that during summer. Note that Seattle's seasonal weather variations are more extreme than those for Jacksonville, per National Oceanic and Atmospheric Administration (NOAA). As also supported by clinical literature, we thus conjecture that Twitter users based in Seattle are more prone to depressive symptoms during winter than in Jacksonville, or other low weather variability cities.

Why precision prediction?

- Prediction is a precursor for prevention
  - Most relevant to selective or secondary prevention
  - Necessary, but not sufficient
- Precision prediction is a precursor for precision prevention
  - Who, When, Why, What
- Pragmatic precision prediction is a precursor for pragmatic precision prevention

WHO is at risk:
Suicidal behavior following completion of PHQ9





Mental Health Research Network

KAISER PERMANENTE®

WHO is at risk:
Shortcomings of risk stratification using PHQ9

Mental health specialty visits - Suicide attempt within 90 days

% of Visits	Item 9 Score	Actual Risk	% of Suicide Attempts
2.5%	3	2.3%	20%
3.5%	2	1.4%	19%
11%	1	0.7%	26%
83%	0	0.2%	35%

PPV: 2.3% in highest tier
Efficiency: Top 6% identifies 39% of events
Sensitivity: Depending on threshold 35% or 61% missed
AND – PHQ9 scores missing for significant minority of visits

WHO is at risk: Machine learning prediction of suicidal behavior

- 7 MHRN health systems with combined enrollment of 8 million
- 20 million visits by 4 million members aged 13 or older
  - Mental health specialty visits
  - General medical visits with mental health or substance use diagnosis
- Linked to nonfatal suicide attempt or suicide death within 90 days
- Approximately 150 potential predictors (and 200 possible interactions)
  - Demographic characteristics (age, sex, race/ethnicity, SES)
  - Current/recent/past mental health diagnoses
  - Current/recent/past mental health medications
  - Current/recent/past acute care utilization for mental health diagnosis
- Prediction models developed in 65% sample, validated in 35%

WHO is at risk:
Improved concentrationat the top AND fewer events "missed" at the bottom,

Thoughts of death or self-harm	% of Visits	% of Suicide Deaths
Nearly every day	2.5%	20%
More than half the days	3.5%	19%
Several days	11%	26%
Not at all	83%	35%

Excludes all those missing PHQ9!

Percentile of Visits	% of Suicide Deaths
>99.5th	12%
99th to 99.5th	11%
95th to 99th	25%
90th to 95th	16%
75th to 90th	16%
50th to 75th	13%
<50th	6%

WHO is at risk:
Do these predictions generalize?

- Across health systems?
  Similar integrated health system (KP Northern Cal) – YES
  Completely different health system (Southcentral Foundation) - YES

- Across time?
  ICD9 to ICD10? – YES
  2012 to 2015 to 2018? - YES

- Across age groups?
  Adolescents? – YES
  Seniors? - YES

- Across racial and ethnic groups?
  Prediction of suicide attempt? - YES
  Prediction of suicide death? – MAYBE (poorer performance in some)

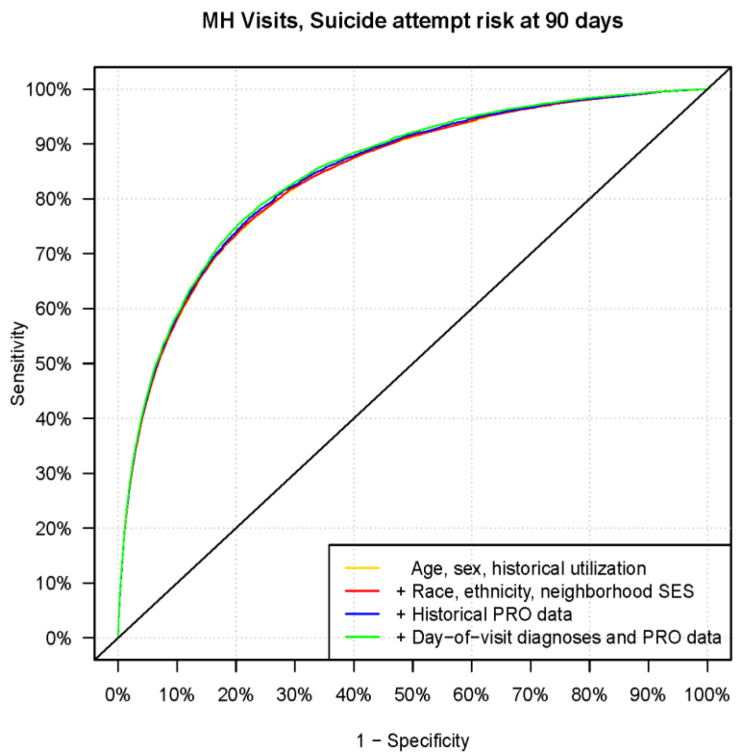
WHO is at risk: Can we do better?

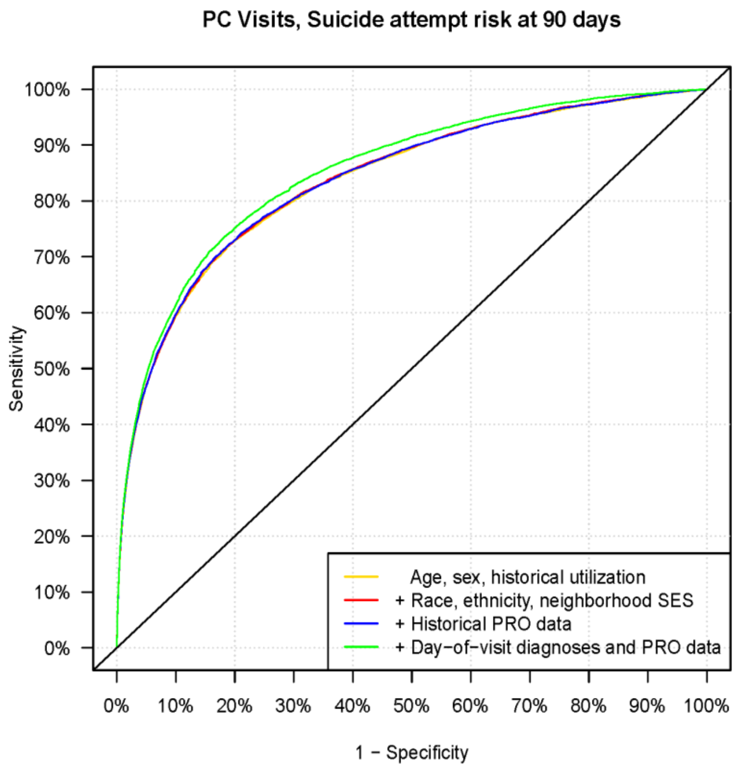
- Wider range of interactions - NO
- More sophisticated modeling methods (RF, ANN, Ensemble) – NO
- More detailed temporal encoding (48 patterns vs. 3) – NO

WHEN are people at risk? Strongest predictors are long-term

SUICIDE ATTEMPT IN 30 DAYS AFTER MENTAL HEALTH SPECIALTY VISIT (of 78 predictors selected)	SUICIDE DEATH IN 30 DAYS AFTER MENTAL HEALTH SPECIALTY VISIT (of 29 predictors selected)
Depression diagnosis in last 5 yrs.	Mental health ER visit in last 3 mos
Age 13-17 with Female	2^{nd} Gen. Antipsychotic Rx in last 5 years
Drug abuse diagnosis in last 5 yrs.	Hypnotic Drug Rx. in last year.
PHQ-9 Item 9 score =3 in last 90 days	Benzodiazepine Rx. in last 3 mos.
Drug use disorder Diag. in last 5 yrs	Mental health inpatient stay in last year
Suicide attempt diagnosis in last year	Mental health inpatient stay in last 3 mos
Mental health inpatient stay in last year	Antidepressant Rx. in last 3 mos.
Suicide attempt diagnosis in last 3 mos.	Charlson Comorbidity Score
Antidepressant Rx. in last yr.	Number of PHQ9 responses in last 90 days
Personality disorder diag. in last 5 yrs.	PHQ-9 item 9 score = 1 with Age
Benzodiazepine Rx. in last 3 mos.	PHQ-9 item 9 score = 3 with Age
PHQ-9 Item 9 score=2 in last year	Suicide attempt diagnosis in last 5 yrs.
Self-inflicted laceration in last 5 yrs	Depression diagnosis in last 5 yrs.
Antidepressant Rx. in last 3 mos.	PHQ-9 Item 9 score = 2 with Age
Eating Disorder diagnosis in last 5 yrs.	Schizophrenia diagnosis in last 5 yrs.

WHEN are people at risk?
Today's information adds little to prediction





WHEN are people at risk:
We'll need to look elsewhere

- Historical information (that we can easily extract from records):
  - Includes more stable indicators
  - Identifies more stable risk
- What information (and sources) might help identify shorter-term risk:
  - Financial stresses and housing instability (from credit agencies)
  - Relationship disruption (from social media)
  - Arrest/incarceration (from criminal justice system)

WHEN are people at risk?
Who is our customer and what do they need?

- “Imminent risk” most relevant to urgent care settings (Inpatient, emergency department, crisis services)
- Most outpatient interventions act over weeks or months
- We must consider harms of falsely identifying imminent risk

WHY are people at risk?
Prediction vs. Inference

- Inference: Is it true?

  Interpretation is the main point.

- Prediction: Does it serve?

  Interpretation is beside the point.

WHY are people at risk?
Strongest predictors are exactly what you expect

SUICIDE ATTEMPT IN 30 DAYS AFTER MENTAL HEALTH SPECIALTY VISIT (of 78 predictors selected)	SUICIDE DEATH IN 30 DAYS AFTER MENTAL HEALTH SPECIALTY VISIT (of 29 predictors selected)
Depression diagnosis in last 5 yrs.	Mental health ER visit in last 3 mos
Age 13-17 with Female	2nd Gen. Antipsychotic Rx in last 5 years
Drug abuse diagnosis in last 5 yrs.	Hypnotic Drug Rx. in last year.
PHQ-9 Item 9 score =3 in last 90 days	Benzodiazepine Rx. in last 3 mos.
Drug use disorder Diag. in last 5 yrs	Mental health inpatient stay in last year
Suicide attempt diagnosis in last year	Mental health inpatient stay in last 3 mos
Mental health inpatient stay in last year	Antidepressant Rx. in last 3 mos.
Suicide attempt diagnosis in last 3 mos.	Charlson Comorbidity Score
Antidepressant Rx. in last yr.	Number of PHQ9 responses in last 90 days
Personality disorder diag. in last 5 yrs.	PHQ-9 item 9 score = 1 with Age
Benzodiazepine Rx. in last 3 mos.	PHQ-9 item 9 score = 3 with Age
PHQ-9 Item 9 score=2 in last year	Suicide attempt diagnosis in last 5 yrs.
Self-inflicted laceration in last 5 yrs	Depression diagnosis in last 5 yrs.
Antidepressant Rx. in last 3 mos.	PHQ-9 Item 9 score = 2 with Age
Eating Disorder diagnosis in last 5 yrs.	Schizophrenia diagnosis in last 5 yrs.

But which of those are causal?
And what about the interactions?

WHAT should we do to reduce risk?
Specificity of empirically supported interventions

- Clozapine – people with psychotic disorders and recent self-harm
- Lithium – people with bipolar disorder
- Dialectical Behavior Therapy – women (primarily) with history of repeated self-harm (often with diagnosis of personality disorder)
- Cognitive Behavior Therapy – people with recent hospitalization for self-harm

WHAT should we do to reduce risk?
Population-based prevention vs. traditional clinical trials

WHAT could go wrong?

- False positive errors
  - Most concerning if interventions are intrusive or could cause harm
  - But must also consider alert fatigue or dilution of attention
- False negative errors
  - Sensitivity is improved, but still only mediocre
  - Must clearly communicate that other risk indicators still matter
- Reinforcing health disparities
  - The math is not biased, but the data often are
  - Must consider consequences of false positive and false negative errors
- Disrespecting autonomy
  - Often misunderstood as risk (e.g. data breaches, reidentification)
  - Autonomy interests are harder to measure or address

The dilemma:

FULL TEXT ARTICLE

Patient perspectives on acceptability of, and implementation preferences for, use of electronic health records and machine learning to identify suicide risk

Bobbi Jo H. Yarborough and Scott P. Stumbo

General Hospital Psychiatry, 2021-05-01, Volume 70, Pages 31-37,

Using external data is very/extremely important to help identify risk: 70%

Using external data to identify suicide risk is acceptable: 34%

Who, When, Why, What: Summary by Dr. Wenowdis

Summary

- WHO is at risk?

  We know this.
- WHEN are people at risk?

  We know that we don’t know much about this.
- WHY are people at risk?

  Why do we want to know this?
- WHAT should we do to reduce risk?

  We definitely don’t know this.

Given what we know, what jobs can we do?

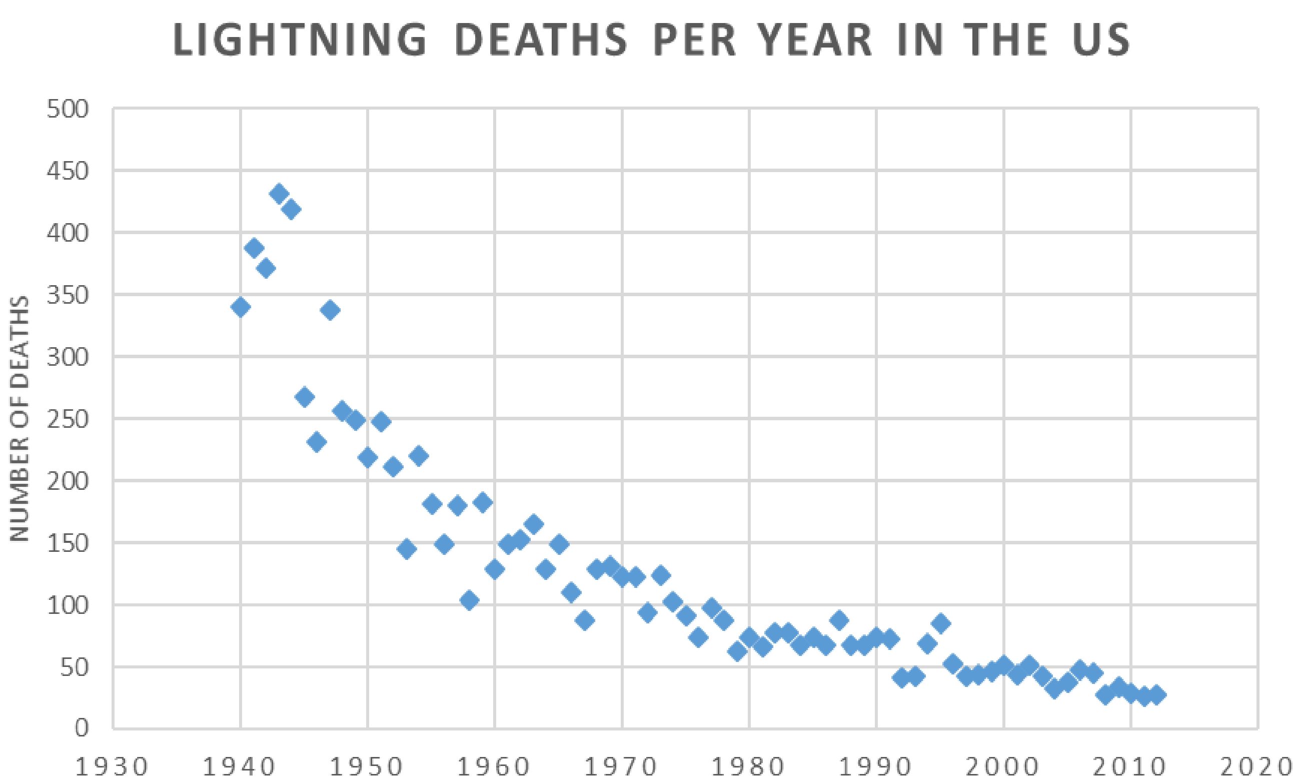
This

Definitely not this

Given what we know, we should:

- Err on the side of sensitivity
- Accept frequent "false positives"
- Think carefully about consequences of unnecessary intervention
- Keep human hands on the wheel

From prediction to prevention: What's actually possible?



Illinois State Climatologist Office, 2013
Data Source: NOAA NWS

KAISER PERMANENTE®