

# Assessing Real-World Utility of Individualized Google Searches in the Early Detection of Suicidal Behaviors: Pilot Cohort Study

Abhishek Pratap, Ethan Kim, Ava Homiar, Patricia A Areánm, Honor Hsin, Tierney K Huppert, Karin E Hendricks, Patrick J Heagerty, Trevor Cohen, Courtney Bagge, Katherine Anne Comtois

### 1. The Problem

**700,000** people die from suicide worldwide each year.

On average, it takes between **11 to 25** nonfatal attempts before a death by suicide occurs.

Despite **50 years** of research, we know **who** is at risk but not **when** they are at the highest risk of suicide.

### 2. Our solution

Google processes over 8.5 billion searches per day, 7% of which are related to health. We may be able to use an individual's Google and YouTube data to predict when they are at the highest risk of self-harm - by analyzing the online web searches along with a timeline of their past self-harm behavior.

### 3. Analytical workflow

**Data-driven approach**: Using deep-learning based context-aware BERT models for Polarity detection (Finance, Education, Family) and Temporal topic modeling.

**Expert-curated approach**: Anxiety, Suicide methods, Alcohol, and other metrics. Psychometrics of word usage (LIWC).

### 4. Detailed analysis

#### Vertical real-world data integration

### 5. Results from pilot study

**Baseline online search behavior of a specific individual**

The red line shows online search behavior 7-60 days before a confirmed suicide attempt

*Areán PA\*, Pratap A\*, Hsin H, Huppert TK, Hendricks KE, Heagerty PJ, Cohen T, Bagge C, Comtois KA. Perceived Utility and Characterization of Personal Google Search Histories to Detect Data Patterns Proximal to a Suicide Attempt in Individuals Who Previously Attempted Suicide: Pilot Cohort Study. J Med Internet Res. 2021 May 6;23(5):e27918. doi: 10.2196/27918.*

### Search Data Featurization

- Semantic
- Behavioral

Aggregate statistics (mean, standard deviation, max and min) in the selected window

**Association Analysis**

- Baseline distributions
- Z-scores
- Placement value

Bootstrap distribution defining the typical search behavior for each feature

Aggregate statistics before the suicide attempt

--- Randomly selected day    ■ Window Period (7, 15, 30, 60 days) before the event of interest (randomly selected day or suicide attempt)

Sources: WHO, NIMH, Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research by Franklin JC et al