

Google Flu Trends

7 January 2016

Google Flu Trends

<https://www.google.org/flutrends/ca/#CA>

google.org Flu Trends Language: English (United States)

[Google.org home](#)
[Dengue Trends](#)
Flu Trends
[Home](#)
Canada
National
[Download data](#)
[How does this work?](#)
[FAQ](#)

Explore flu trends - Canada

We've found that certain search terms are good indicators of flu activity. Google Flu Trends uses aggregated Google search data to estimate flu activity. [Learn more >](#)

National

2015-2016 Past years

Intense
High
Moderate
Low
Minimal

Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun

Estimates were made using a model that proved accurate when compared to historic official flu activity data. Data current through August 13, 2015.

Flu symptoms?

If you have flu-like symptoms, stay home and rest until they are gone. If you get flu-like symptoms and are pregnant, have underlying health problems, or if they get worse, contact your healthcare provider.

Flu symptoms include:

- Almost always: cough, fever
- Common: fatigue, muscle aches, sore throat, headache, decreased appetite, runny nose
- Sometimes: nausea, vomiting, diarrhea

More official info: www.fightflu.ca

Public Health Agency of Canada / Agence de santé publique du Canada

Animated Flu Trends in Google Earth

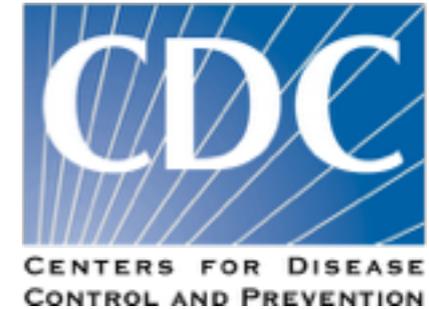
[Download and explore](#) Flu Trends data in Google Earth. Need Google Earth? [Download it here.](#)

Embed this chart

Use [this embed code](#) to show this chart on your website.



Flu Data



Proportion of Influenza-Like-Illnesses (PILI):
proportion of doctor visits where patient has an
“influenza-like illness”

ILINet in U.S., compiled and published by CDC.

Takes time to record, prepare, and publish
(week or two). Might like to have results faster.

Google Flu tries to predict what the CDC data will be.

Google Flu Trends Model (Classic Edition)



Query fraction (QF): proportion of Google searches that match a given **Query**

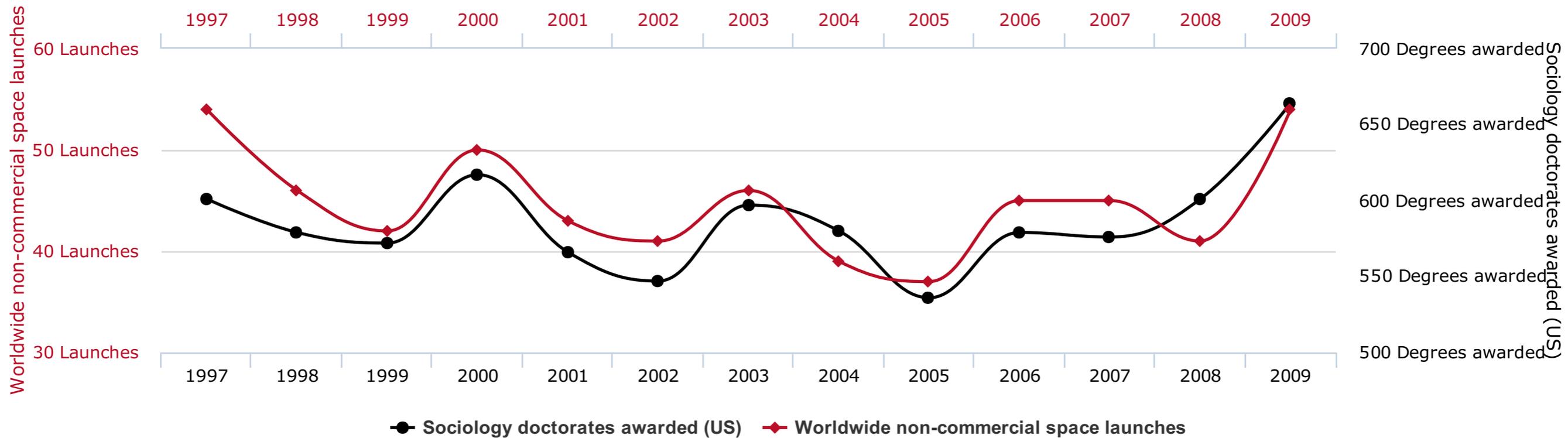
E.g., 16 per 100000 searches
are the word “fever”

50 million distinct queries,
unknown number of searches (billions)

A little EDA

Step 1: Find terms whose **QF *correlates with* PILI**

Worldwide non-commercial space launches
correlates with
Sociology doctorates awarded (US)



Step 2: Keep the best 45

- Most terms are secret; categories are:
Influenza Complication, Cold/Flu Remedy, General Influenza Symptoms, Term for Influenza, Specific Influenza Symptom, Symptoms of an Influenza Complication, Antibiotic Medication, General Influenza Remedies, Symptoms of a Related Disease, Antiviral Medication, Related Disease
- [symptoms of bronchitis],[pnumonia]*,[fever],
[early signs of the flu],[robitussin],[influenza a],
[amoxicillin],[strep throat]

Step 3: Build Model

- Let **S** be the sum of the QF of the 45 terms over a 1-week period.
- Let **Y** be the PILI for next week, provided by CDC
- Find **w** and **c** so that

$$\text{logit}(\mathbf{Y}) \approx \mathbf{w} \times \text{logit}(\mathbf{S}) + \mathbf{c}$$

as much as possible, on **Y** and **S** from historical data.

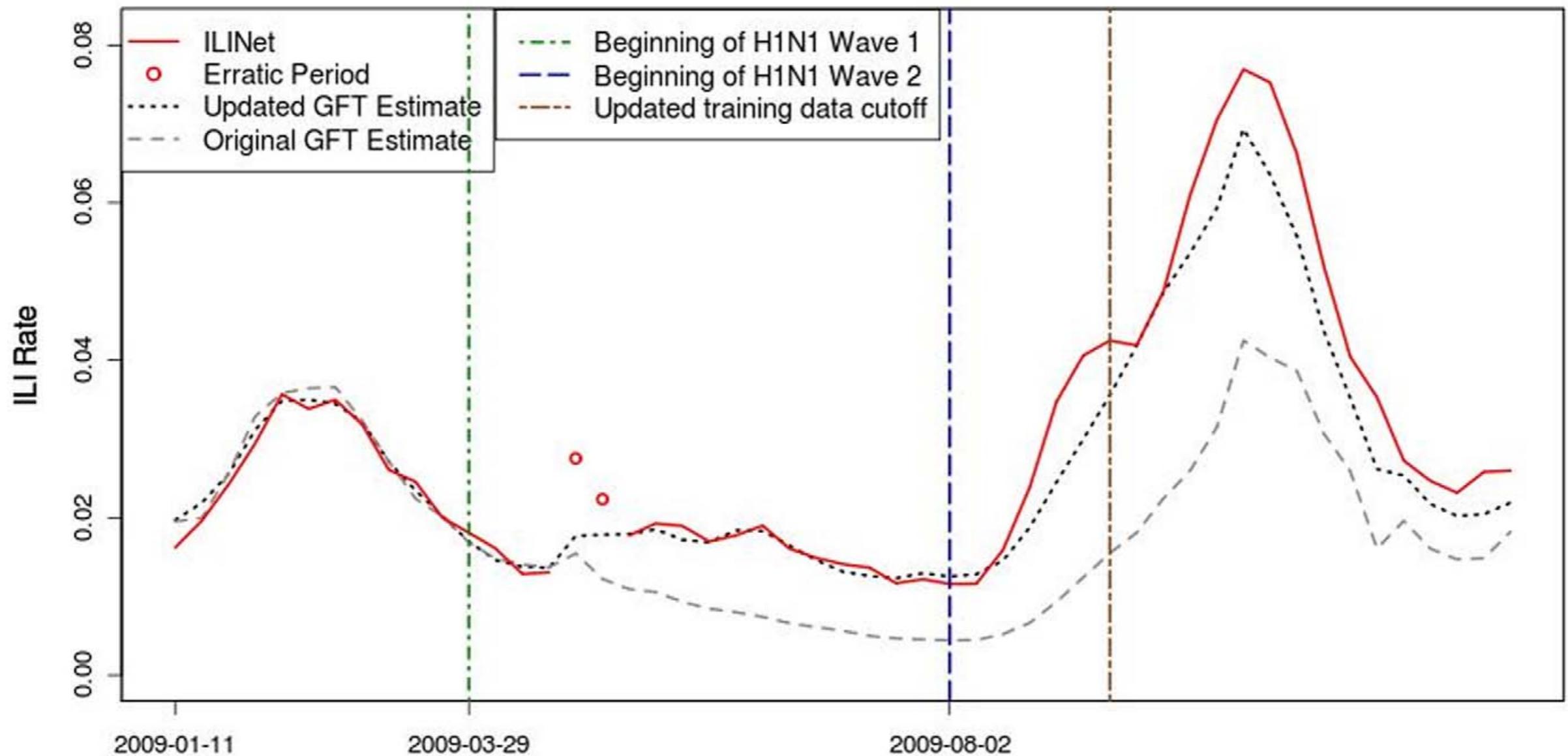
- **c** determines predicted “baseline ILI” if **S** = 0,
w determines how fast predicted **Y** grows with **S**
- To predict **Y** in the future, stick the learned **w** and **c**,
and the new **S**, into the model, get result.

Step 2: Keep the best 45

- Why 45?
 - Chosen to maximize accuracy on unseen data

Flu Trends Ups and Downs

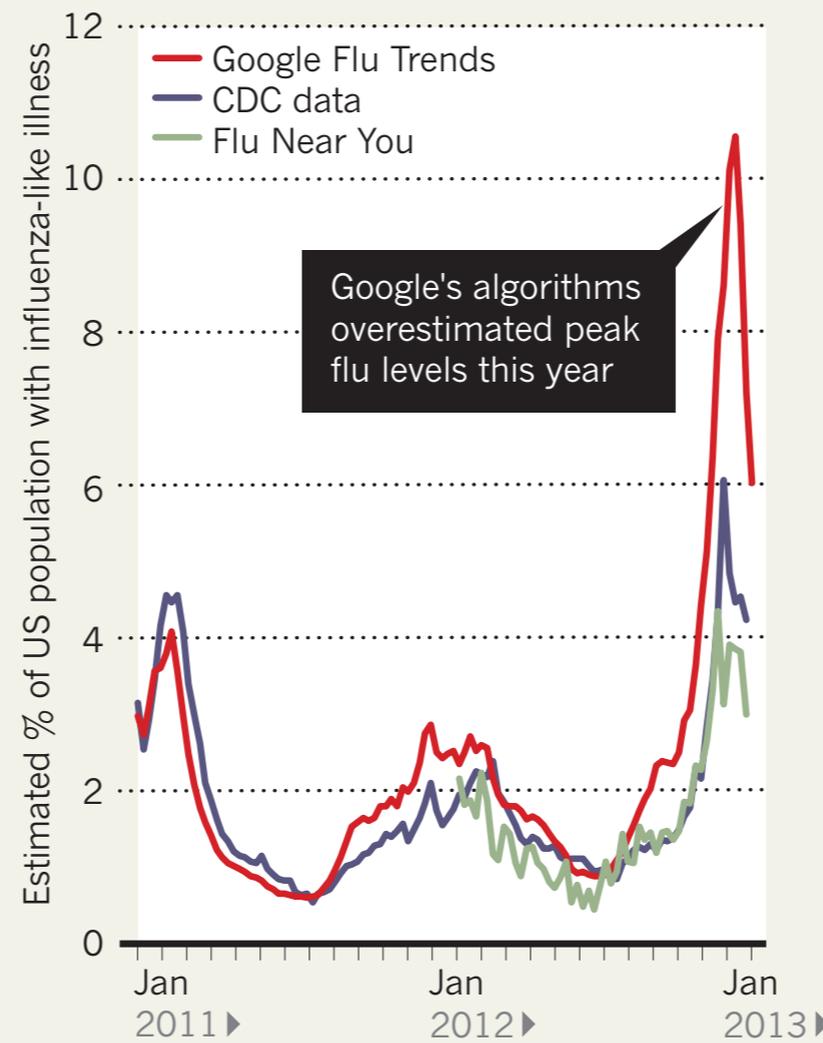
A ILINet Data and GFT Estimates: 2009



Flu Trends Ups and Downs

FEVER PEAKS

A comparison of three different methods of measuring the proportion of the US population with an influenza-like illness.



Step 2: Keep the best 45

- Why 45?
 - Chosen to maximize accuracy on unseen data
 - *Other search queries in the top 100, not included in our model, included topics like **“high school basketball”***

Why

“high school basketball”?

Table 1. Comparison of relative query category volume in original and updated United States GFT models.

Query Category	Sample Query	Original Model Relative Category Volume	Updated Model Relative Category Volume
Symptoms of an influenza complication	[symptoms of bronchitis]	6%	11%
Influenza complication	[pnumonia]*	42%	6%
Specific influenza symptom	[fever]	6%	39%
General influenza symptoms	[early signs of the flu]	2%	30%
Cold/flu remedy	[robitussin]	12%	4%
Term for influenza	[influenza a]	<1%	3%
Antibiotic medication	[amoxicillin]	12%	0%
Related disease	[strep throat]	16%	<1%

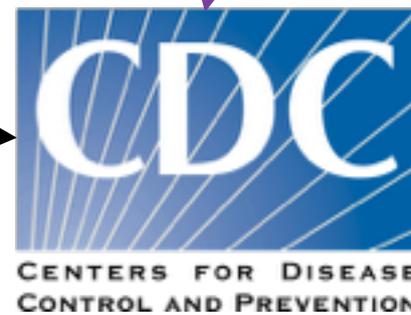
*Search users often misspell the word *pneumonia*.

doi:10.1371/journal.pone.0023610.t001

$$\text{logit}(\mathbf{Y}) \approx \mathbf{w} \times \text{logit}(\mathbf{S}) + \mathbf{c}$$

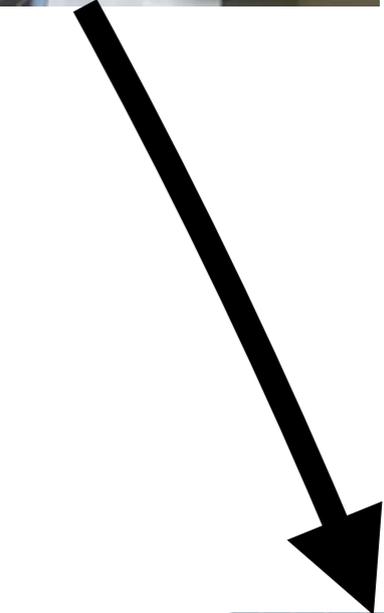
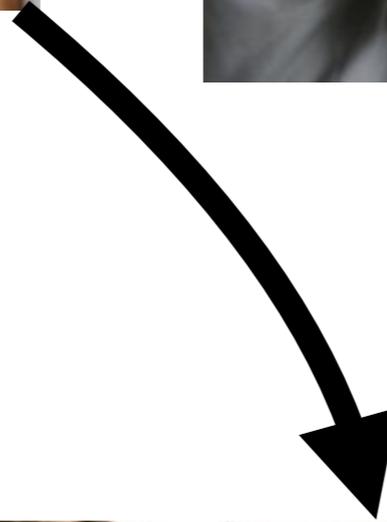
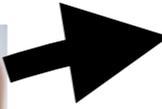
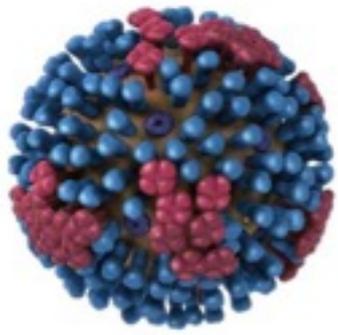


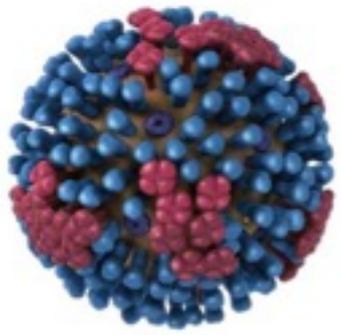
© Alamy



CENTERS FOR DISEASE
CONTROL AND PREVENTION

**What causes
flu-related searches?**





...

