

# Health Misinformation in Search and Social Media

By  
Amira Ghenai

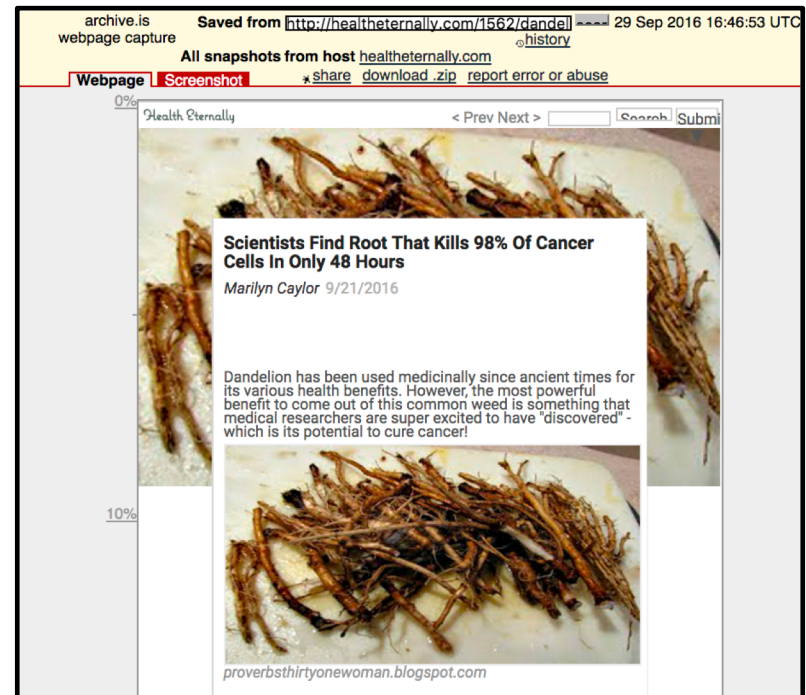
A thesis  
presented to the University of Waterloo in fulfillment of the  
thesis requirement for the degree of Doctor of Philosophy  
in  
Computer Science



UNIVERSITY OF  
**WATERLOO**

# Imagine

- Your friend on social media posted an article about a cancer treatment
- The post reached 1.4 m shares
- You are curious to know more about this..
- You turn to your search engine and look up **“dandelion weed cancer”**





## Integrative Medicine

Developing Your Personal Care Plan

Therapies, Classes & Workshops +

About Herbs, Botanicals & Other Products -

Overview

Search About Herbs

Herbs, Botanicals & Other Products: FAQs

About Herbs App

Email Us

About Mind-Body Therapies

Our Research & Clinical Trials

Multimedia +

# Dandelion



## Common Names

- Blowball
- Puffball
- Lion's tooth
- Pu gong ying
- Swine snout
- Wild endive

Jump to:

[For Patients & Caregivers](#)

[For Healthcare Professionals](#)

## For Patients & Caregivers

[How It Works](#)

[Purported Uses](#)

Evidence-based medicine

- To treat cancer

Laboratory studies have shown dandelion to have anticancer properties, but clinical studies have not shown this effect in humans.

- To treat diabetes

Laboratory studies have shown dandelion to lower blood sugar, however no clinical studies have evaluated its effects in diabetic humans.



About 436,000 results (0.80 seconds)

### Stage-4 Cancer Survivor - From New Cancer Treatments

[www.peggysuesurvived.com/](http://www.peggysuesurvived.com/) +1 760-614-2856

I'm a 10-Year Survivor Because I Had New Cancer Treatments At A Mexican Clinic.  
Cost of Treatments · Get Clinic Recommendation · 6 Treatments I Used · Types of Cancer Treated

### Dandelion Root Cancer Treatment - Cancer Tutor

<https://www.cancertutor.com/dandelionroot/>

Dandelion Root has been used in China for certain kinds of cancers for centuries. .... I hoe them when needed and keep the weeds and grass out of them.

### The Cancer Fighting Medicine That's Growing in Your Yard

<https://thetruthaboutcancer.com/cancer-treatments/>

Discover the cancer fighting medicine that's likely hiding in plain sight on your front ... Science is revealing that the dandelion (which many consider a weed) is ...

### Cancer-fighting tea set for clinical trials in Ontario - Calgary - CBC News

[www.cbc.ca/news/canada/calgary/calgary-anti-cancer-tea-1.3370691](http://www.cbc.ca/news/canada/calgary/calgary-anti-cancer-tea-1.3370691)

Dec 18, 2015 - 'Dandelion root extract has very potent anti-cancer activity,' biochemist ... to harness the potential cancer-fighting properties of a common weed.

### 30 patients to test dandelion's cancer-killing potential - Windsor - CBC ...

[www.cbc.ca/news/30-patients-to-test-dandelion-s-cancer-killing-potential-1.295981](http://www.cbc.ca/news/30-patients-to-test-dandelion-s-cancer-killing-potential-1.295981)

Feb 17, 2015 - ... are being recruited to take part in a study looking at the anti-cancer abilities of dandelion root extract, taken from the common yellow yard weed.

### Dandelion Root Kills 98% of Cancer Cells in 48 Hours? - Snopes.com

[www.snopes.com/dandelion-kills-cancer/](http://www.snopes.com/dandelion-kills-cancer/)

Claim: Dandelion root can kill 98 percent of cancer cells in 48 hours.

Claimed by: Internet

Fact check by Snopes.com: MOSTLY FALSE

[Feedback](#)

### Holistic Cancer Treatment - Dandelion Tea Kills Cancer Cells

<https://www.tylertolman.com/health-articles/holistic-cancer-treatment/>

I am a huge proponent when it comes to holistic cancer treatment. This common weed has shown to be super effective against cancer cells in just 48 hours and ...

### Dandelion root far more effective in fighting cancer cells than ...

[www.naturalnews.com/053880\\_dandelion\\_root\\_cancer\\_cell\\_suicide\\_chemotherapy.h...](http://www.naturalnews.com/053880_dandelion_root_cancer_cell_suicide_chemotherapy.h...)

Mar 3, 2016 - (NaturalNews) A so-called "weed" growing right in your front and back yards could hold the key to being the most effective cancer-fighting ...

### Dandelion Root Kills Cancer Cells In Lab Tests - Reset.me

[reset.me/story/dandelion-root-kills-cancer-cells-in-lab-tests/](http://reset.me/story/dandelion-root-kills-cancer-cells-in-lab-tests/)

Jul 2, 2015 - A common weed may hold the key to treating, and possibly curing, some types of blood cancers. Dandelion root may become an inexpensive ...

### Dandelion tea obliterates cancer cells - Natural Health 365

[www.naturalhealth365.com/benefits-of-dandelion-tea-cancer-cells-1344.html](http://www.naturalhealth365.com/benefits-of-dandelion-tea-cancer-cells-1344.html)


Mar 10, 2015 - In fact, science now confirms that dandelion tea kills cancer cells. ... (NaturalHealth365) Those bright yellow weeds blooming all over suburban ...


Google dandelion weed cancer


All Images News Videos Shopping More Settings Tools


About 436,000 results (0.80 seconds)


**Stage-4 Cancer Survivor - From New Cancer Treatments**  
[www.peggysuesurvived.com/](http://www.peggysuesurvived.com/) +1 760-614-2856  
 I'm a 10-Year Survivor Because I Had New Cancer Treatments At A Mexican Clinic.  
 Cost of Treatments · Get Clinic Recommendation · 6 Treatments I Used · Types of Cancer Treated


 **Dandelion Root Cancer Treatment - Cancer Tutor**  
<https://www.cancertutor.com/dandelionroot/>  
 Dandelion Root has been used in China for certain kinds of cancers for centuries. .... I hoe them when needed and keep the weeds and grass out of them.


 **The Cancer Fighting Medicine That's Growing in Your Yard**  
<https://thetruthaboutcancer.com/cancer-treatments/>  
 Discover the cancer fighting medicine that's likely hiding in plain sight on your front ... Science is revealing that the dandelion (which many consider a weed) is ...


 **Cancer-fighting tea set for clinical trials in Ontario - Calgary - CBC News**  
[www.cbc.ca/news/canada/calgary/calgary-anti-cancer-tea-1.3370691](http://www.cbc.ca/news/canada/calgary/calgary-anti-cancer-tea-1.3370691)  
 Dec 18, 2015 - 'Dandelion root extract has very potent anti-cancer activity,' biochemist ... to harness the potential cancer-fighting properties of a common weed.


 **30 patients to test dandelion's cancer-killing potential - Windsor - CBC ...**  
[www.cbc.ca/news/30-patients-to-test-dandelion-s-cancer-killing-potential-1.295981](http://www.cbc.ca/news/30-patients-to-test-dandelion-s-cancer-killing-potential-1.295981)  
 Feb 17, 2015 - ... are being recruited to take part in a study looking at the anti-cancer abilities of dandelion root extract, taken from the common yellow yard weed.

 **Dandelion Root Kills 98% of Cancer Cells in 48 Hours? - Snopes.com**  
[www.snopes.com/dandelion-kills-cancer/](http://www.snopes.com/dandelion-kills-cancer/)  
 Claim: Dandelion root can kill 98 percent of cancer cells in 48 hours.  
 Claimed by: Internet  
 Fact check by Snopes.com: MOSTLY FALSE  
 Feedback

 **Holistic Cancer Treatment - Dandelion Tea Kills Cancer Cells**  
<https://www.tylertolman.com/health-articles/holistic-cancer-treatment/>  
 I am a huge proponent when it comes to holistic cancer treatment. This common weed has shown to be super effective against cancer cells in just 48 hours and ...

 **Dandelion root far more effective in fighting cancer cells than ...**  
[www.naturalnews.com/053880\\_dandelion\\_root\\_cancer\\_cell\\_suicide\\_chemotherapy.h...](http://www.naturalnews.com/053880_dandelion_root_cancer_cell_suicide_chemotherapy.h...)  
 May 3, 2016 - (NaturalNews) A so-called "weed" growing right in your front and back yards could hold the key to being the most effective cancer-fighting ...

 **Dandelion Root Kills Cancer Cells In Lab Tests - Reset.me**  
[reset.me/story/dandelion-root-kills-cancer-cells-in-lab-tests/](http://reset.me/story/dandelion-root-kills-cancer-cells-in-lab-tests/)  
 Jul 2, 2015 - A common weed may hold the key to treating, and possibly curing, some types of blood cancers. Dandelion root may become an inexpensive ...

 **Dandelion tea obliterates cancer cells - Natural Health 365**  
[www.naturalhealth365.com/benefits-of-dandelion-tea-cancer-cells-1344.html](http://www.naturalhealth365.com/benefits-of-dandelion-tea-cancer-cells-1344.html)  
 Mar 10, 2015 - In fact, science now confirms that dandelion tea kills cancer cells. ... (NaturalHealth365) Those bright yellow weeds blooming all over suburban ...

'I'm living proof it works'

CBC: "researchers hoped to test dandelion root's potential.."

'Snopes' fact checking!

Like

Share

Social Media



cancer cure

Search

Save Search

Export

Page 1 of 449

How to run an Advanced Search

Sort by Total Engagements

Facebook Engagements, Twitter Shares, Pinterest Shares, Reddit Engagements, Number of Links, Evergreen Score, Total Engagement

Hemp Could Free Us From Oil, Prevent Deforestation, **Cure Cancer**, and It's Environmentally Friendly

By Truthcommand - Mar 31, 2018  
truthcommand.com

- Save, View Backlinks, View Sharers, Share



A YouTuber who claimed being vegan cured her cancer has died from cancer

By Harry Shukman - Feb 14, 2018  
babe.net

- Save, View Backlinks, View Sharers, Share



**Cancer vaccine that cured 97% of tumours in mice will be tested on humans**

By Alexandra Thompson Health... - Mar 28, 2018  
dailymail.co.uk

- Save, View Backlinks, View Sharers, Share



A New Israeli **Cancer Vaccine May Cure 90% of All Cancer Types in Just One Shot**

By Jews News - Nov 25, 2017  
jewsnews.co.il

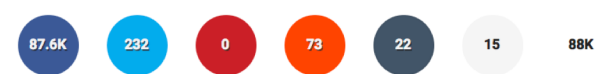
- Save, View Backlinks, View Sharers, Share



Woman '**cured**' her terminal cancer with cannabis oil

By Alexandra Thompson Health... - Mar 12, 2018  
dailymail.co.uk

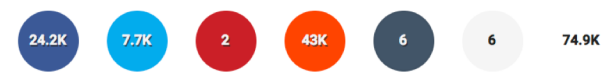
- Save, View Backlinks, View Sharers, Share



Ex-Clinton staffer: If Obama **cured cancer**, Trump would try to bring it back

By Avery Anapol - Dec 5, 2017  
thehill.com

- Save, View Backlinks, View Sharers, Share



**Vegan YouTuber Who Said Veganism Cures Cancer Has Died Of Cancer**

By Neelam Tailor - Feb 20, 2018  
unilad.co.uk

- Save, View Backlinks, View Sharers, Share



**DANDELION CURES CANCER, HEPATITIS, LIVER, KIDNEYS, STOMACH ... HERE'S HOW TO USE IT!**

By Admin - Mar 9, 2018  
dailynativenews.site

How Article

- Save, View Backlinks, View Sharers, Share



How to run an Advanced Search

Sort by Total Engagements

Hemp Could Free Us From Oil, Prevent Deforestation, Cure Cancer, and It's Environmentally

By Truthcommand - Mar 31, 2018 truthcommand.com

They are all unproven treatments

A YouTuber who claimed being vegan cured her cancer has died from cancer

By Harry Shukman - Feb 14, 2018 babe.net

Save View Backlinks View Sharers Share

Cancer vaccine that cured 97% of tumours in mice will be tested on humans

By Alexandra Thompson Health... - Mar 28, 2018 dailymail.co.uk

They manipulate real facts

A New Israeli Cancer Vaccine May Cure 90% of All Cancer Types in Just One Shot

By Jews News - Nov 25, 2017 jewsnews.co.il

View Sharers Share

Woman 'cured' her terminal cancer with cannabis oil

By Alexandra Thompson Health... - Mar 12, 2018 dailymail.co.uk

Save View Backlinks View Sharers Share

Cancer patients!

Ex-Clinton staffer: If Obama cured cancer, Trump would try to bring it back

By Avery Anapol - Dec 5, 2017 thehill.com

Vegan YouTuber Who Said Veganism Cures Cancer Has Died Of Cancer

By Neelam Tailor - Feb 20, 2018 unilad.co.uk

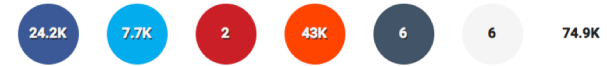
Save View Backlinks View Sharers Share

DANDELION CURES CANCER, HEPATITIS, LIVER, KIDNEYS, STOMACH ... HERE'S HOW TO USE IT!

By Admin - Mar 9, 2018 dailynativenews.site

Save View Backlinks View Sharers Share

Facebook Engagements Twitter Shares Pinterest Shares Reddit Engagements Number of Links Evergreen Score Total Engagement





# Problem Definition

Looking at two major online platforms (online search/social media), how does online health misinformation effect people's health-related decisions?

# Proposed Solution

## **In online search**

- Understand how search results influence decisions
  - Controlled laboratory studies
- > What factors contribute to people's final health-decisions?
- > How can we help people make correctly informed decisions?

## **In social media**

- Detect and track misinformation in social media
  - Content analysis, ML, observational studies
- > Can we automatically detect medical rumors?
- > Who propagates questionable medical advise?

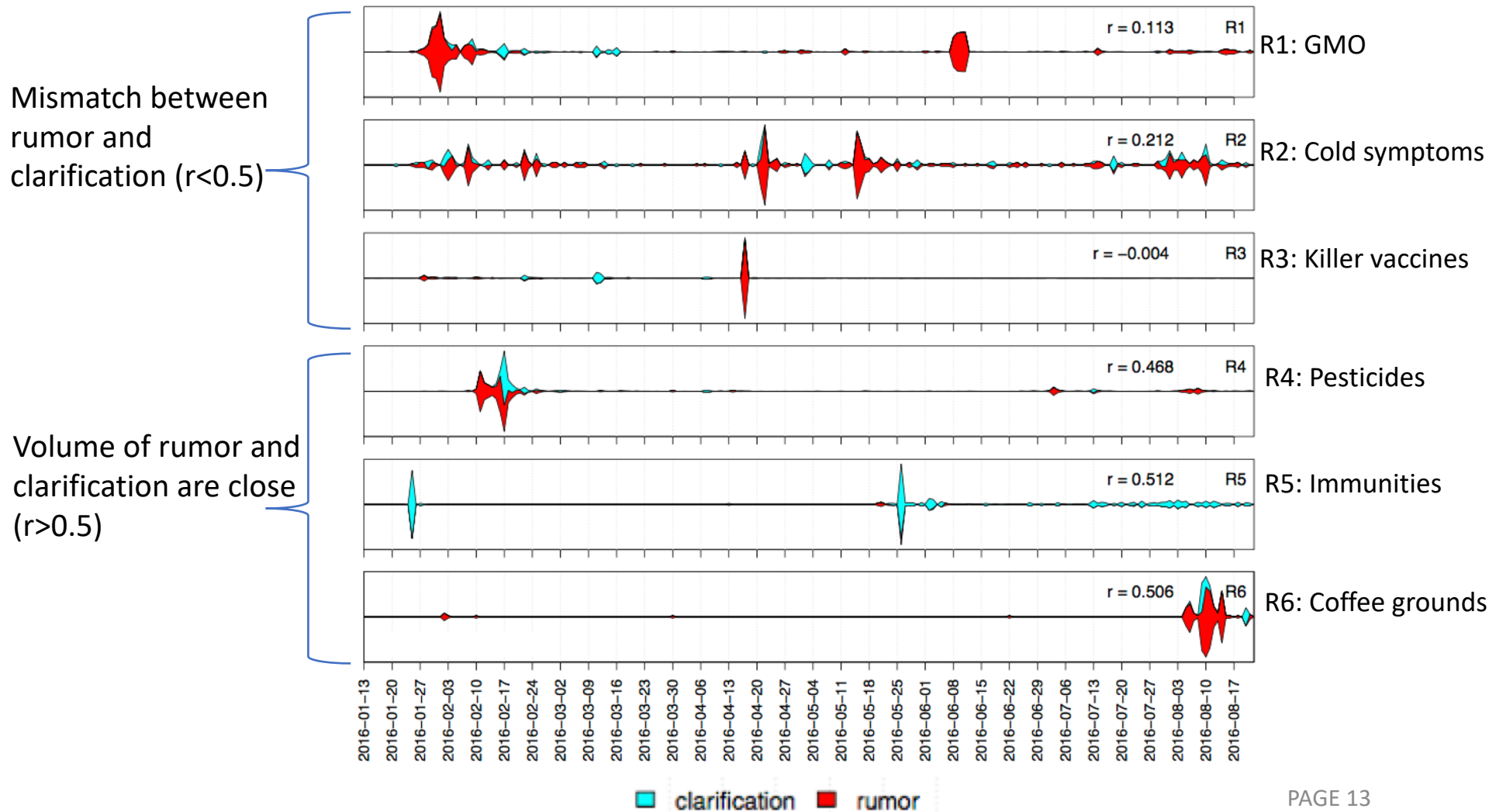
# List of Publications

1. **Amira Ghenai**, Yelena Mejova, 2017, January. Catching Zika Fever: Application of Crowdsourcing and Machine Learning for Tracking Health Misinformation on Twitter. The Fifth IEEE International Conference on Healthcare Informatics - **ICHI 2017**
2. **Amira Ghenai**, Yelena Mejova, 2018, November. Fake Cures: User-centric Modeling of Health Misinformation in Social Media. The 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing – **CSCW'18**
3. Frances Pogacar, **Amira Ghenai**, Mark D. Smucker, Charles L. A. Clarke, 2017, October. The Positive and Negative Influence of Search Results on People's Decisions about the Efficacy of Medical Treatments. The 3rd ACM International Conference on the Theory of Information Retrieval – **ICTIR'17**
4. **Amira Ghenai**, Mark D. Smucker and Charles L. A. Clarke. A Think-Aloud Study to understand Factors Affecting Online Health Search. [under review ACM **CHIIR'20**]

# Tracking Health Misinformation on Twitter (Chap. 3)

- Collected 13 million tweets regarding the Zika outbreak
- Selected 6 Zika rumors from WHO & Snopes
- Hand-craft queries to extract corresponding tweets
- Use crowdsourcing to identify rumor, clarification and other tweets
- Generated 48 different features (Twitter, linguistic, sentiment, medical and readability)
- Train classification model to identify rumor tweets

# Results



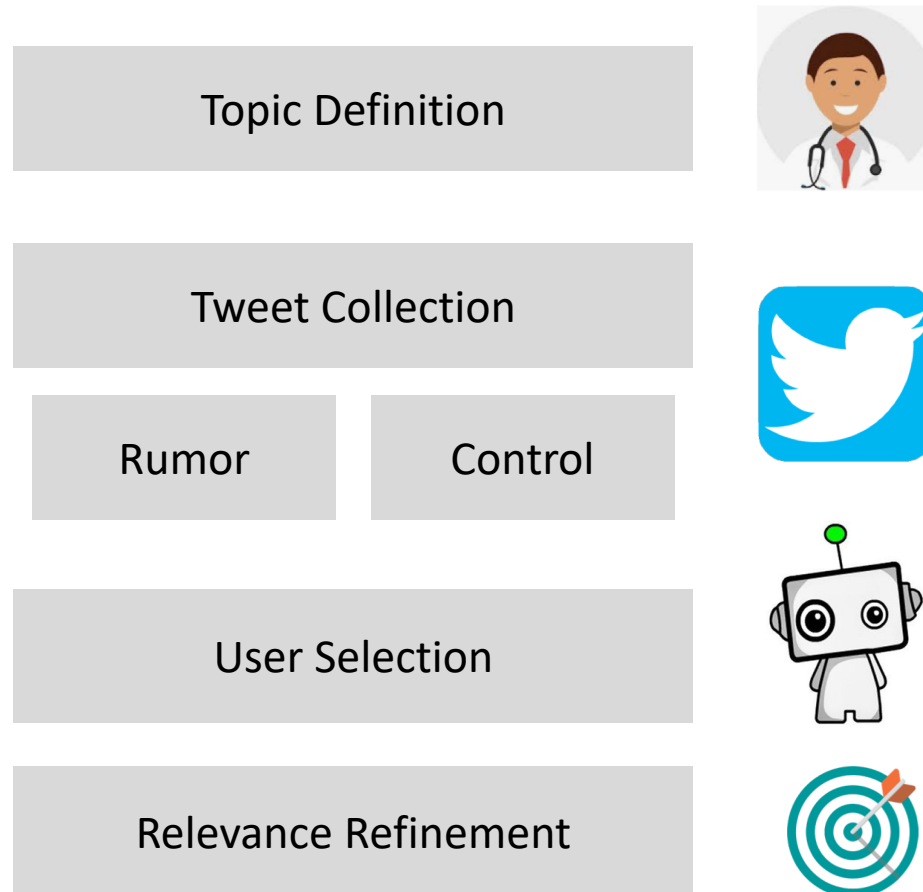
# Results

- Best features to predict if a tweet is a rumor
  - Medical features
  - Tweet text syntax
  - Sentiment features
  - Twitter features
- Classification model with high accuracy 0.92, precision 0.97, recall 0.95, F-measure 0.96 (90/20 training testing split)
- Training on 5 topics and testing on the 6<sup>th</sup>
  - New topic without labelled data when building the classifier
  - Low accuracy for new topics
  - Importance of labelled data about the topic being classified

We can automatically detect rumor tweets...what about possible future health rumors?

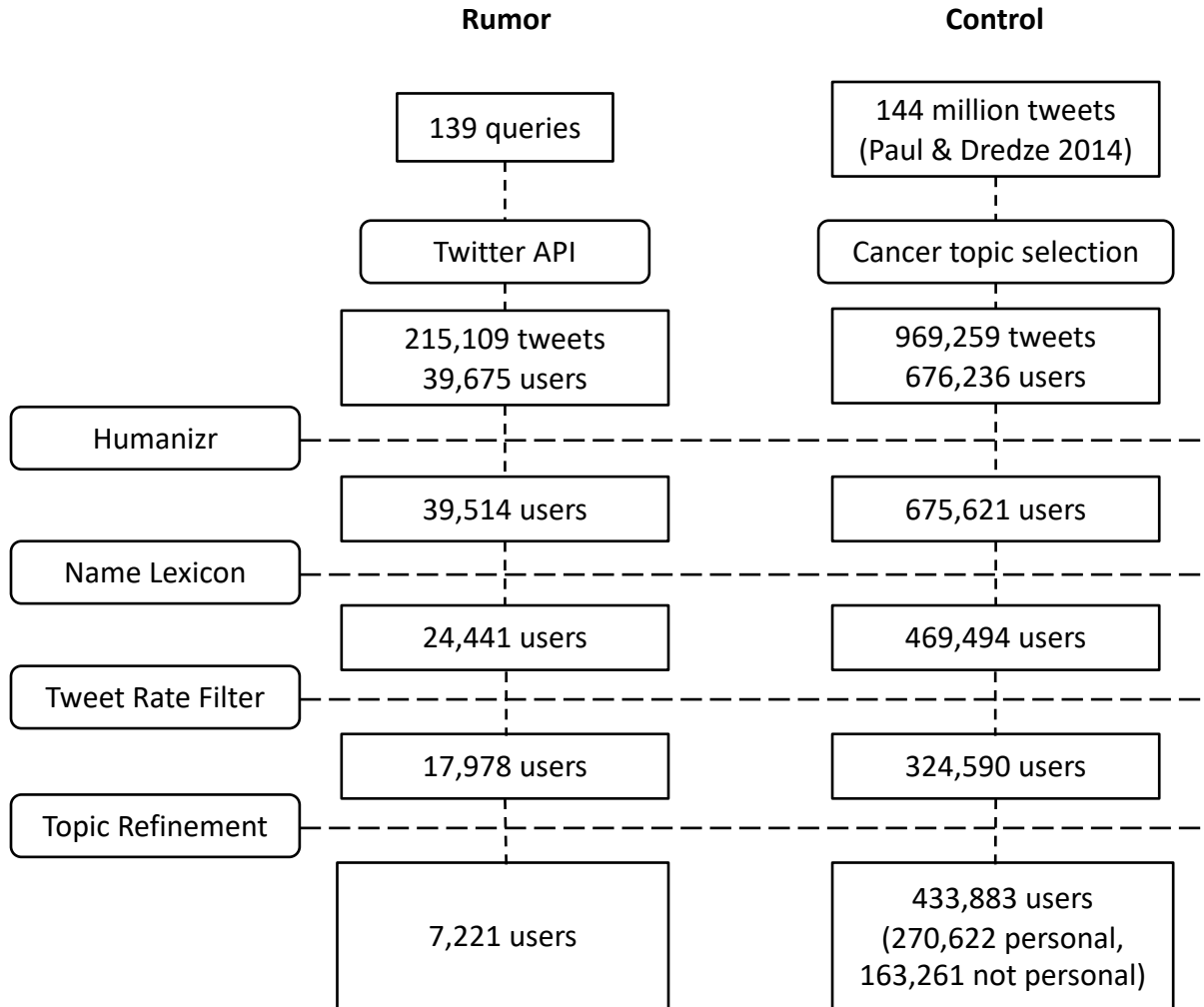
Looking at **who** propagates rumors might help predict potential health rumors!

# Health Misinformation User Modeling in Twitter (Chap. 4)





# User Selection



# Can we predict the “rumor spreading” behavior?

- Look at all the tweets before a users posts a tweet about the rumor
  - Rumor users: tweets before the first rumor post
  - Control users: (no date for first rumor!) sample users' dates from a normal distribution having mean and variance of first rumor in Rumor dataset
- At least 100 tweets of 4,212 rumor users, sample control users

# Can we predict the “rumor spreading” behavior?

- Use following feature types:
  - User features
  - Tweet features
  - Entropy: the intervals between posts to measure the predictability of retweeting patterns
  - LIWC (Linguistic Inquiry and Word Count): psycholinguistic measures shown to express user mindset
- Train logistic regression classifier to identify users that might be talking about rumors in the future using their historical timeline

Figure 2: Logistic regression with LASSO regularization model, predicting whether a user posts about a rumor, with forward feature selection.

McFadden R2 = **0.90**

Significance levels:

$p < 0.0001$  \*\*\*,  $p < 0.001$  \*\*,  $p < 0.01$  \*,  $p < 0.05$  .

variable	coefficient	std. error	p-value
(Intercept)	-6.160	1.405	***
Avg syllables per word	17.120	0.660	***
Is verified	-40.310	42310	
Percentage uppercase / lowercase	-0.201	0.018	***
Word count	1.491	0.131	***
SMOG readability score	-0.753	0.123	***
Percentage uppercase	0.191	0.019	***
Character count	-0.163	0.024	***
Number of cancer tweets	0.001	1.9E-04	***
LIWC48: ingest	1.839	0.722	*
Negative word count	-1.460	0.262	***
URL count	3.364	0.505	***
Is retweet	4.947	0.790	***
word2vec count	-0.634	0.165	***
LIWC55: focuspast	-1.636	0.567	**
LIWC37: tentat	2.531	0.859	**
Number of sentences	-0.610	0.205	**
LIWC32: male	-1.820	1.000	
Interval entropy	0.508	0.105	***
Account age	-0.001	2.7E-04	***
LIWC23: posemo	-0.490	0.384	
LIWC61: time	-1.431	0.378	***
LIWC13: adverb	1.758	0.536	**
LIWC20: number	2.936	1.317	*
Statuses count	7.1E-05	2.6E-05	**
LIWC42: hear	-4.742	1.799	**
Has 1st person pronoun	-1.504	0.662	*
LIWC62: work	1.591	0.665	*
LIWC40: percept	1.217	0.754	

We looked at cancer cures in social media.

What about using **online search** to answer health-related questions?

# Measuring search results effect on people's online health-search(Chap.5)

- Total of 60 participants were told to pretend to be searching for the answer to a question about the effectiveness of a treatment for a health issue
- Participants had to classify the medical treatments as
  - *Helpful*: Treatment has direct positive effect
  - *Unhelpful*: Treatment is ineffective or has a direct negative effect
  - *Inconclusive*: Unsure about the effectiveness
- They either received a search engine result page, or the control condition, with no SERP

# Medical treatments

- The medical treatments and associated medical conditions were all formulated as “Does X help Y?”
  - Each medical question was classified as *helpful* or *unhelpful*, as determined by the Cochrane Review by White and Hassan.
  - Each participant answers 10 questions (5 helpful and 5 unhelpful)
- Examples:
- Unhelpful: “Do insoles help back pain?”
  - Helpful: “Does caffeine help asthma?”

# Experimental Conditions

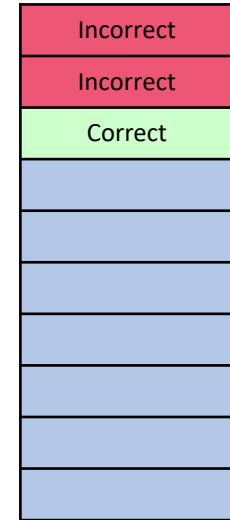
## Search Result Bias

- 8:2 ratio of results
- 8 correct, 2 incorrect
- 2 correct, 8 incorrect

➤ 10 × 10 Graeco-Latin square to fully balance the experimental conditions with the treatments

## Topmost Correct Rank

- Always had a correct result at rank 1 or rank 3





# User performance

## Accuracy

- Fraction of correct decisions
- A correct response agrees with the *authoritative answer*

➤ Generalized linear (logistic) mixed effect model for stat. sig

## Harm

- Fraction of harmful decisions
- A harmful decision is opposite of the *authoritative answer*
- *Inconclusive* is not considered a harmful decision

# Results - Accuracy

Bias	Topmost Correct Rank	Correct decisions	Average Accuracy
Incorrect	3	0.23 ± 0.04	0.23 ± 0.04
Incorrect	1	0.23 ± 0.04	
Control	No search results	0.43 ± 0.05	0.43 ± 0.05
Correct	3	0.59 ± 0.05	0.65 ± 0.05
Correct	1	0.70 ± 0.04	

Independent Variable	Dependent Variable	Pr(>Chisq)
Search Result Bias	Correct Decision	<< 0.001
Topmost Correct Rank	Correct Decision	0.16

# Results - Harm

Bias	Topmost Correct Rank	Harmful decisions	Average Harm
Incorrect	3	0.41 ± 0.05	0.38 ± 0.05
Incorrect	1	0.35 ± 0.04	
Control	No search results	0.20 ± 0.04	0.20 ± 0.04
Correct	3	0.13 ± 0.03	0.10 ± 0.03
Correct	1	0.06 ± 0.02	

Independent Variable	Dependent Variable	Pr(>Chisq)
Search Result Bias	Harmful Decision	<< 0.001
Topmost Correct Rank	Harmful Decision	0.06

People are influenced with the search result.

What **factors** contributed to their final decisions?

How can we help them make correct decisions?

# Factors affecting Online health-related search (Chap. 6)

- Total of 16 participants were asked to think aloud while they used search results to determine the efficacy of health treatments
- Procedure:
  - Concurrent think-aloud with eye tracking and video recording
  - Retrospective: Video recording reviewed by participants *post hoc* with further information elicited
  - Final questionnaire
- Think-aloud data transcribed and coded

# Factors affecting Online health-related search (Chap. 6)

- Previous study conditions (search bias/rank)
- 8 treatments out of the 10 treatments from the previous study
- Participants' performance (accuracy/harm)
- Coding scheme:
  - Think-aloud transcribed
  - Performed twice within different time periods
  - Mixed methods research approach to generated codes (top-down and bottom-up)

# Results – Search results bias

Results Bias	Correct decisions	Harmful decisions
Correct	0.67 ± 0.08	0.06 ± 0.03
Incorrect	0.32 ± 0.06	0.28 ± 0.06

Independent Variable	Dependent Variable	Pr(>Chisq)
Search Result Bias	Correct Decision	<< 0.001
Topmost Correct Rank	Correct Decision	0.8

# Results – Coding

No	Name	Participants	References
C1	Majority	14	36
C2	Authoritativeness	13	153
C2	Stats & studies	12	20
C6	Advertisements	7	16
C7	Date	7	15
C8	References	7	12
C9	Negative information	6	15
C10	Information representation	5	18
C12	Prior_belief	5	8
C14	Readability	4	8
C13	Relevance	4	7
C15	Past_experience	3	3
C16	Text_length	3	3
C17	Images	2	6
C18	Rank	2	4
C19	Social_factor	1	2



# Results – Coding

No	Name	Participants	References
C1	Majority	14	36
C2	Authoritativeness	13	153
C2	Stats & studies	12	20
C6	Advertisements	7	16
C7	Date	7	15
C8	References	7	12
C9	Negative information	6	15
C10	Information representation	5	18
C12	Prior_belief	5	8
C14	Readability	4	8
C13	Relevance	4	7
C15	Past_experience	3	3
C16	Text_length	3	3
C17	Images	2	6
C18	Rank	2	4
C19	Social_factor	1	2

The majority of the search results stating that the treatment helps or that the treatment does not help or looking for a consensus of different search results.

If participants are exposed to results geared towards a specific direction, they end up being influenced by what the majority of the search results state.

# Results – Coding

No	Name	Participants	References
C1	Majority	14	36
C2	Authoritativeness	13	153
C2	Stats & studies	12	20
C6	Advertisements	7	16
C7	Date	7	15
C8	References	7	12
C9	Negative information	6	15
C10	Information representation	5	18
C12	Prior_belief	5	8
C14	Readability	4	8
C13	Relevance	4	7
C15	Past_experience	3	3
C16	Text_length	3	3
C17	Images	2	6
C18	Rank	2	4
C19	Social_factor	1	2

The trustworthiness and reliability of the source of information.

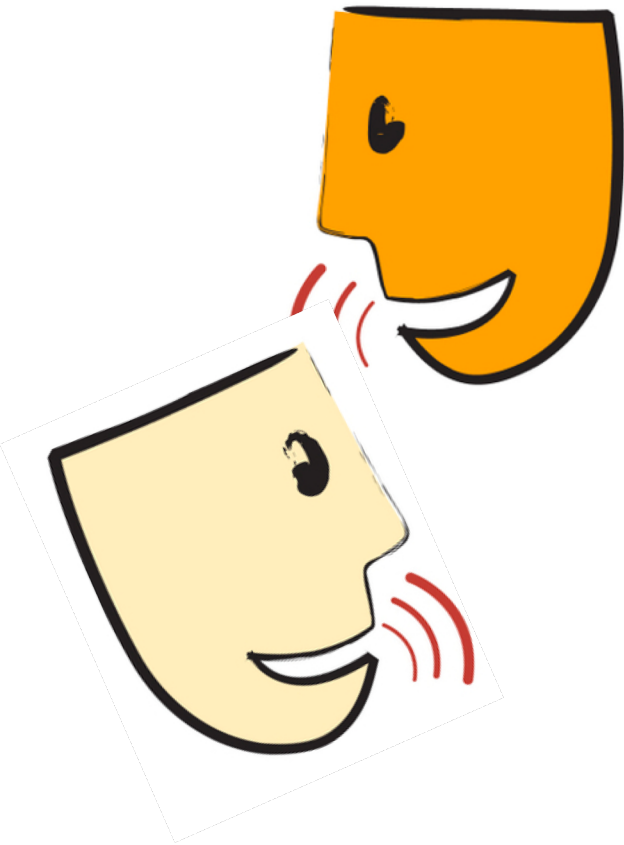
Participants pay attention to authoritativeness. (We did not control for authoritativeness)

# Results – Coding

No	Name	Participants	References
C1	Majority	14	36
C2	Authoritativeness	13	153
C2	Stats & studies	12	20
C6	Advertisements	7	16
C7	Date	7	15
C8	References	7	12
C9	Negative information	6	15
C10	Information representation	5	18
C12	Prior_belief	5	8
C14	Readability	4	8
C13	Relevance	4	7
C15	Past_experience	3	3
C16	Text_length	3	3
C17	Images	2	6
C18	Rank	2	4
C19	Social_factor	1	2

The quality of the search results page such as the presence of ads, research studies or reference/citations.

Participants pay attention to quality.



- Majority

*"I'm going to say helps because a lot of people, like it was just, the vast number were in agreement."*

- Authoritativeness

*"WebMD. It's a more trustworthy source, I think."*

- Quality

*"So this looks like a research study, so I think it's pretty reliable."*

# Retrospective think-aloud & post-task questionnaire

- Retrospective think-aloud to get insights on new strategies not discovered in the previous step
- Post task questionnaire aligns with the think-aloud collected data

# Conclusion

- Mixed-method approaches to address the health misinformation in online search and social media
- Online search:
  - Traditional search needs to incorporate a notion of negative gain to incorrect information
- Social media:
  - Detection - automatically detecting Twitter users who may post questionable information
  - Intervention- attempting to change those individuals' views
  - Prevention - quickly identifying and limiting the spread of misinformation

# Future work

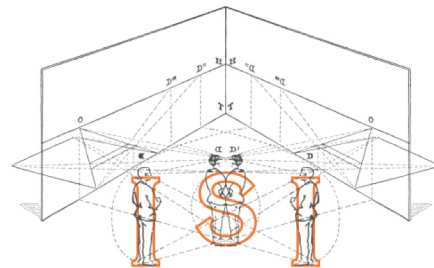
- Effect of authoritativeness in online health search
- Rank effect in online search
- User studies on different populations
- False advertisement campaigns in social media (Facebook) about cancer cures

Prof. Charlie L.A Clarke

Prof. Mark D. Smucker

Dr. Yelena Mejova

Frances A. Pogacar



ISI Foundation



معهد قطر لبحوث الحوسبة  
Qatar Computing Research Institute

عضو في مؤسسة قطر  
Member of Qatar Foundation





# ZIKA RUMOR LIST

Rumor Description	Example	#tweets
R1) Zika virus is linked to genetically modified mosquitoes <b>(WHO)</b>	<i>BIOWEAPON! #Zika Virus Is Being Spread by #GMO #Mosquitoes Funded by Gates!</i>	73,832
R2) Zika virus symptoms are similar to seasonal flu <b>(WHO)</b>	<i>The affects of Zika are same symptoms as the Common Cold. #StopSpreading- GMOMosquitos</i>	469
R3) Vaccines cause micro- cephal in babies <b>(WHO)</b>	<i>Government document confirms tdap vaccine causes microcephaly.. <a href="https://t.co/4ZVLbaabbG">https://t.co/4ZVLbaabbG</a></i>	4,329
R4) Pyriproxyfen insecticide causes microcephaly <b>(WHO)</b>	<i>"Argentine and Brazilian doctors sus- pect mosquito insecticide as cause of microcephaly"</i>	10,389
R5) Americans are immune to Zika virus <b>(Snopes)</b>	<i>Yup and Americans R immune to Zika, so why fund a response to it?</i>	351
R6) Coffee as mosquito- repellent to protect against Zika <b>(Snopes)</b>	<i>Bring on the Cuban coffee. Say Good- bye to Zika mosquitoes. Dee Lundy- Charles Fredric Sweeney Joshua Oates Laure... <a href="http://fb.me/tArL595b">http://fb.me/tArL595b</a></i>	202

# Geolocate Zika tweets:

1. Get GPS location (latitude and longitude) values
  - Very small portion has this information (less than 1%)
  - Convert GPS location to country name (World Borders API)
2. No GPS location, get country name from mentioned place in tweet
3. No place value, get country name from user location
  - field is very messy and not well formatted
  - Use Yahoo Placemaker API to get information about user mentioned place such as type (city, country, street..), GPS coordinates
4. Convert GPS coordinates of user location to country name (World Borders API)
5. No user location, country name is the country associated with the user in previously posted tweets

# Is This Tweet A Rumor?

## Instructions ▾

We are interested in identifying rumors in tweets saying **Americans are immune to Zika virus**, which is not true. So, given a tweet, annotate it whether

- **Rumor** -- stating or implying that Americans are immune to Zika virus
- **Clarification** -- stating that Americans do not have immunization to Zika virus
- **Other** -- neither of above

Instructions

For example:

- *I'm immune to Zika because I'm a Goddamn American*
  - is a **Rumor** (outright states Americans are immune to Zika)
- *Zika Strain From American Outbreak Spreads To Africa Where Some May Already Be Immune*
  - is a **Rumor** (implies that Americans are immune to Zika)
- *Yup and Americans R immune to Zika, so why fund a response to it?*
  - is a **Rumor** (outright states that Americans are immune to Zika)
- *Americans are not immune to the Zika virus. <http://www.snopes.com/americans-immune-zika-virus/> ...*
  - is a **Clarification** (outright states that Americans are not immune to Zika)
- *None of the illegals entering the United States are immunized --- exposing Americans to all types of diseases and viruses like ZIKA. Sigh.*
  - is a **Other** (about illegals entering US, but not about Americans being immune to Zika)

Examples

For more information on rumors see Snopes website <http://www.snopes.com/americans-immune-zika-virus/>

American Immunity to the Zika Virus: A dangerous and false rumor circulating on Twitter claims that Americans... <https://t.co/uubOo6pAYy>  
see [original tweet](#)

Is it a rumor about Americans being immune to Zika? (required)

- Rumor
- Clarification
- Other

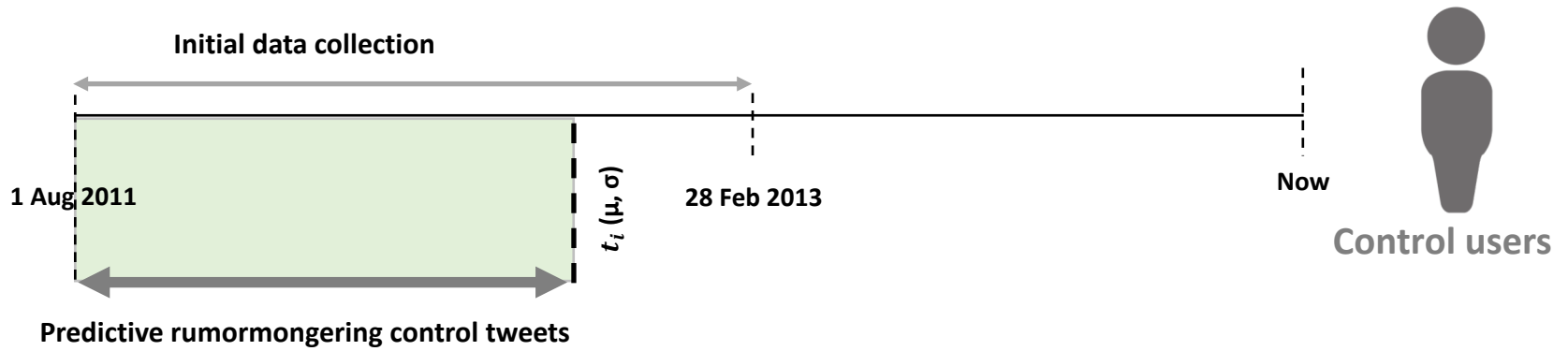
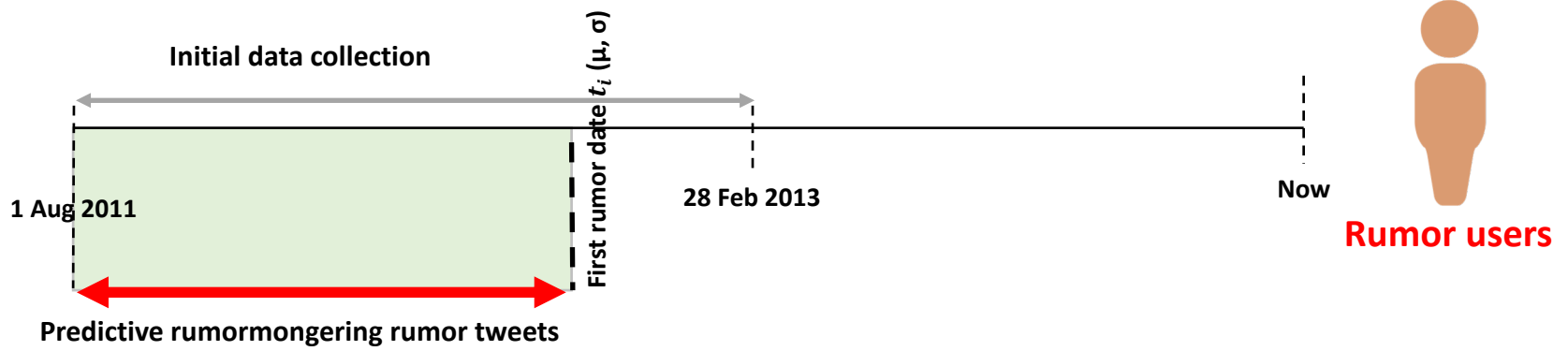
Labeling task

# ZIKA CLASSIFICATION - MEDICAL LEXICON

Social media

1. Download “Infectious disease” pages [~ 22 thousand words] => corpus  $M$
2. Get top ~22 thousand words from all Wikipedia pages => corpus  $W$
3. Compute the probability of every work in corresponding corpus:  
$$Mp_w = \frac{\text{count}_w}{\sum_w M} \quad Wp_w = \frac{\text{count}_w}{\sum_w W}$$
4. Compute difference in probabilities:  
$$p_w = Mp_w - Wp_w$$
5. Get words with highest  $p_w$

Word(w)	$Mp_w$	$Wp_w$	$p_w$	Rank
<b>syphilis</b>	0.01	-	0.01	4
<b>bronchitis</b>	0.002	-	0.002	81
<b>tetanus</b>	0.001	-	0.001	236
diarrhea	0.006	0.121	-0.121	13682
epidemiology	0.009	0.147	-0.138	15284
treatment	0.019	4.652	-4.633	33869
life	0.003	34.61	-34.608	35074



Control History				Rumor History				Rumor Misinformation			
love	1.95%	night	0.66%	good	1.01%	video	0.54%	cancer	1.43%	cells	0.50%
good	1.55%	life	0.63%	health	1.00%	food	0.54%	juice	0.81%	out	0.48%
day	1.34%	happy	0.60%	day	0.96%	back	0.50%	RT	0.77%	healthy	0.45%
time	1.22%	ill	0.59%	love	0.85%	free	0.46%	breast	0.73%	diabetes	0.44%
people	1.00%	hope	0.58%	time	0.78%	work	0.45%	risk	0.61%	prostate	0.44%
lol	0.99%	feel	0.55%	great	0.73%	diet	0.44%	help	0.58%	antioxidant	0.42%
today	0.96%	haha	0.51%	people	0.71%	healthy	0.40%	health	0.55%	pain	0.40%
back	0.94%	follow	0.51%	today	0.68%	post	0.38%	helps	0.54%	chronic	0.37%
great	0.73%	home	0.49%	news	0.62%	weight	0.38%	cure	0.54%	patients	0.37%
work	0.70%	man	0.47%	life	0.57%	blog	0.36%	treatment	0.53%	study	0.36%

Figure 3: Word frequency tables summarizing the top 20 most popular terms, excluding stopwords, in all historical tweets by control users (left), all historical tweets of rumor users (center), and only rumor tweets (right).

# SERP Page:

Question number 2 out of 10:

## Task Question: Does cinnamon help diabetes?

**Treatment:** Cinnamon is a sweet spice made from the bark of an Asian tree and used in cooking and baking.

**Health issue:** Diabetes is a serious disease in which the body cannot properly control the amount of sugar in your blood because it does not have enough insulin.

Showing results 1 to 10

### [Is it true that cinnamon can lower blood sugar in people who have diabetes?](#)

Whether cinnamon can lower blood sugar is a topic of debate – but recent research suggests that cinnamon may be helpful as a supplement to regular diabetes treatment in people with type 2 diabetes. A 2012 review of several recent studies concluded that the use of cinnamon had a potentially beneficial effect on glycemic control. One study published in 2009 found that a 500 mg capsule of cinnamon taken twice a day for 90 days improved hemoglobin A1C levels – a reflection of average blood sugar level for the past two to three months – in people with poorly controlled type 2 diabetes (hemoglobin A1C levels greater than 7 percent).

<http://www.mayoclinic.org/diseases-conditions/diabetes/expert-answers/diabetes/faq-20058472>

### [Cinnamon Improves Glucose and Lipids of People With Type 2 Diabetes](#)

**OBJECTIVE** – The objective of this study was to determine whether cinnamon improves blood glucose, triglycerides, total cholesterol, HDL cholesterol, and LDL cholesterol levels in people with type 2 diabetes. **RESEARCH DESIGN AND METHODS** – A total of 60 people with type 2 diabetes, 30 men and 30 women aged 52.2 ± 6.32 years, were divided randomly into six groups. Groups 1, 2, and 3 consumed 1, 3, or 6 g of cinnamon daily, respectively, and groups 4, 5, and 6 were given placebo capsules corresponding to the number of capsules consumed for the three levels of cinnamon.

<http://care.diabetesjournals.org/content/26/12/3215>

### [Is honey and cinnamon a hoax?](#)

Honey is almost entirely composed of different sugars, with water and trace minerals. Excess sugar (and we really get enough carbohydrates of different sorts in our diet anyway) gets stored as fat. So this diet cannot work unless you reduce your food intake to compensate for the excess sugar you are consuming. Certain types of cinnamon may well have useful properties that might influence insulin levels, but probably not at the recommended concentrations.

[http://www.answers.com/Q/Is\\_honey\\_and\\_cinnamon\\_a\\_hoax](http://www.answers.com/Q/Is_honey_and_cinnamon_a_hoax)

### [Article: Cinnamon Helps Diabetics](#)

Cinnamon a spice found in many kitchens around the world especially at Christmas time. Although it is used by Jewish people at Rosh Hashona (the Jewish New Year that falls in September each year). It has been found that this flavoursome spice has its uses for people who are Type 2 Diabetics and helps to control their Blood Sugar. What Is Diabetes? Diabetes is a disorder, where the body is unable to control the Blood Sugar. Insulin a hormone secreted by the Pancreas is essential in helping to control the Blood Sugar. Some unfortunate people are resistant to Insulin so they have to very strictly adhere to a reduced fat intake in their diet.

<http://www.naturalhealthweb.com/articles/Moffat3.html>

Submit Answer

Document title, snippet, url

**Instructions:** Use the search engine results to help you determine the effectiveness of the specified treatment for the corresponding health issue. Once you believe you have determined the effectiveness of the treatment, go to the bottom of the page and click the "Submit Answer" button to submit your answer on the next page.

Category	Explanation
Helps	A medical treatment <b>helps</b> if the treatment is effective and has a direct positive influence on the specified illness.
Inconclusive	The effectiveness of a medical treatment is <b>inconclusive</b> if medical professionals are still unsure if the treatment will have a positive, negative, or no influence on the illness.
Does not help	A medical treatment <b>does not help</b> if the treatment is ineffective and either has no effect or has a direct negative influence on the specified illness.

Instructions & classifications



# Confusion Matrices

- Control Condition

Decision	Total Responses	
Unhelpful	33%	26%
Helpful	33%	37%
Inconclusive	33%	37%

- With SERP

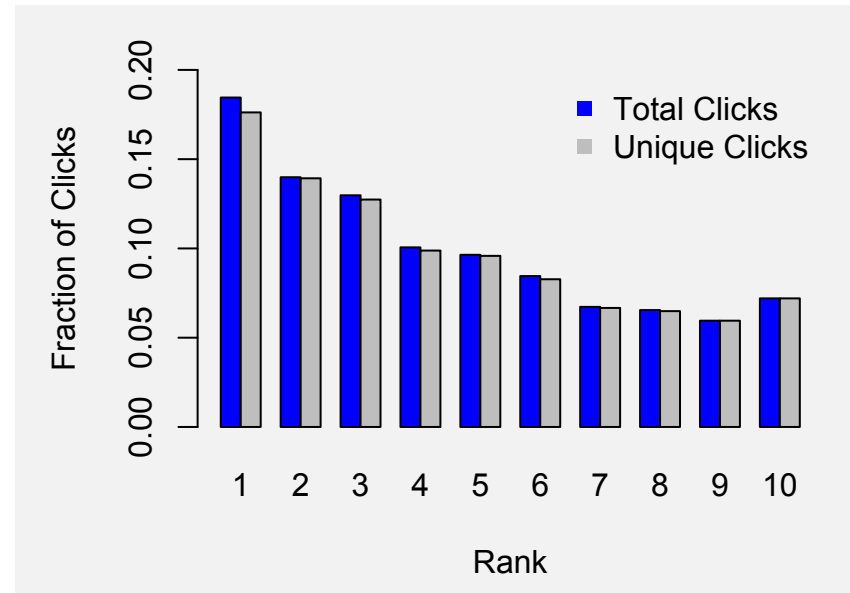
Decision	Total Responses	
Unhelpful	x%	
Helpful	x%	27%
Inconclusive	y%	41%
		32%

- Under the control we should expect an even percent of responses in each category.
- Under the biased conditions, we should expect an even amount between helpful and unhelpful.

➤ There is an overall bias to saying that a treatment is helpful.

# Results - Clicks

- We recorded the overall and unique clicks in each participant's session.
- Participants that interacted more with the search results were more likely to make a correct decision.



Dependent Variable	Mean Number of Clicks
Correct Decisions	3.73 ± 0.20
Incorrect Decisions	3.32 ± 0.2
Harmed Decisions	3.02 ± 0.30
Unharmed Decisions	3.65 ± 0.3

# Post Task Questionnaire

No	Question	Yes	No	Maybe
1.	Do you believe that exposure (i.e. most results say the treatment helps/does not help) is important in determining the effectiveness of the medical treatment? And why?	13	2	1
2.	Do you believe that rank (i.e. highly ranked results say the treatment helps/does not help) is important in determining the effectiveness of the medical treatment? And why?	9	6	1
3.	Do you believe that quality is important in determining the effectiveness of the medical treatment? And please elaborate on what quality means to you?	15	0	1
4.	Do you believe that the web page layout is important in determining the effectiveness of the medical treatment? And why?	12	2	2
5.	Do you believe that social factors (i.e. experience of other people you know such as friends, family etc.) is important in determining the effectiveness of the medical treatment? And why?	9	5	2
6.	Did you notice any manipulation of the search results? If yes, then can you guess what was it?	9	7	0
7.	How do you describe your experience with the think-aloud process?	-		