Automated Discovery of Perceived Health-related Outcomes about E-cigarettes from Reddit

Alexandra DeLucia¹, MSE, Adam Poliak², PhD, Zechariah Zhu³, BS, Stephanie R. Pitts⁴, PhD, Mario Navarro⁴, PhD, Sharareh Shojaie⁴, MSE, John W. Ayers³, PhD, Mark Dredze¹, PhD ¹Johns Hopkins University; ²Byrn Mawr College; ³University of California San Diego; ⁴ Center for Tobacco Products, US Food and Drug Administration

Introduction

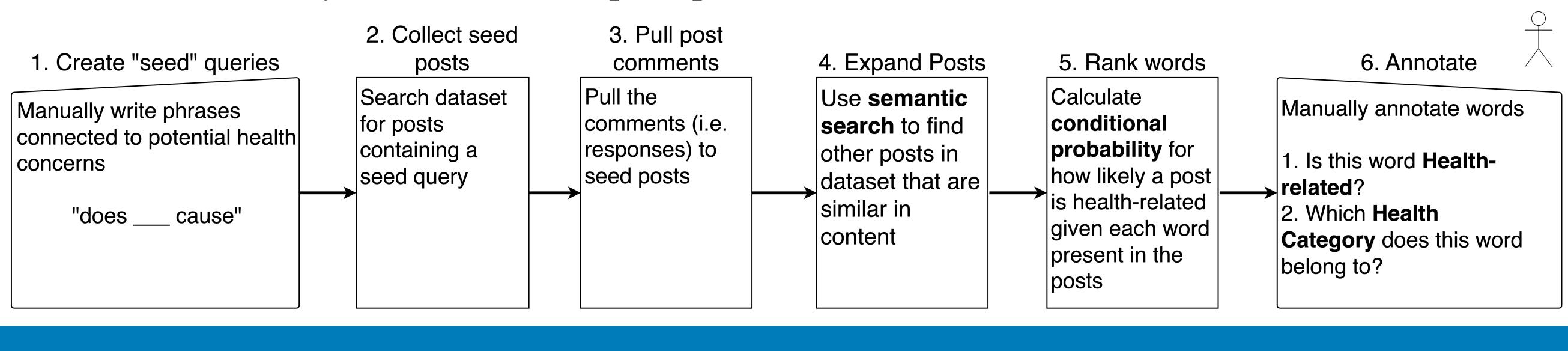
- concerns or lack of opportunity for respondents to include them
- ENDS users can share their concerns on Reddit, a publicly available social media platform where cigarettes
- health concerns associated with e-cigarettes

Methods

- Data: 10,403,433 posts from the subreddit " $r/electronic_cigarette$ " posted between 09/2008-04/2022
- Identification of health-related posts :

 - 2. Collected all replies to these posts
 - 3. Leveraged a *semantic search* model to find 10,905 posts that are similar in content (not necessarily in words) describing a health-concern
- Compared these 10,905 health-related posts with a random sample of posts from *r/electronic_cigarette* to discover words indicative of health-concerns
- A word is indicative of a health concern if there is a high *conditional probability* of a post being health-related given the presence of the word
- Words with at least 0.8 conditional probability were annotated with two questions: Is this word health-related? If so, what health category does this word belong to?

Overview of discovery of health-related perceptions



¹Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. MPNet: Masked and Permuted Pre-training for Language Understanding. In Advances in Neural Information Processing Systems, 2020. This work was funded by FDA via the Johns Hopkins University Center of Excellence in Regulatory Science and Innovation (CERSI Grant #UFD005942B, FOA: RFA-FD-18-011.

• Traditional surveys may fail to capture perceived health concerns from e-cigarettes, vapes, and other electronic nicotine delivery system (ENDS) users due to researchers being unaware of specific

people participate in discussions, including those focusing on perceived health concerns about e-

• Our goal is to leverage this abundant, publicly available data to automatically detect and categorize

I. Identified posts asking questions about health-related outcomes, e.g., "does vaping ecig cause."



Results & Discussion

Health-related words most associated with identified health-related posts

Term	Health Category	Total Occurrences	P(health-related word)
hydrated	Non-specific/Other	53	100%
copd	Respiratory	51	100%
suicidal	Psychiatry	41	100%
acne	Dermatology	31	100%
diketones	Respiratory	30	100%
bronchiolitis	Respiratory	29	100%
cilium	Respiratory	27	100%
obliterans	Respiratory	24	100%
inflammation	Inflammation	23	100%
popcorn	Respiratory	336	99%

Example posts containing words highly associated with health-related outcome perceptions

$\hat{\mathbf{C}}$	Posted by 💿 8 years ago	Ţ	$\tilde{\bigcirc}$	
2	Nose bleeds?			
\bigcirc	nsfw		✓ Posted by <a>10 years ago	\square
	I'd ask my doctor, but she has confessed to be ignorant of ecigs. I have avoided PG for the past few years due to a sore, scratchy throat and general malaise. I recently reconsidered, and joined that juice tasting program, and received two bottles. I've been pretty steadi		It's okay. You're not dying. It's all part of the switching process! Insfw	•
	vaping one with 25% PG from my Kayfun. No sore throat or malaise, but I've been getting noseble	eds.	(talk to your doctor before writing anything off) actually completely normal reactions to no longer smoking.	
	I take Flonase now (my car is diesel, I need the help), but I don't think that's related. I drink almost water, excepting my morning decaf and nightly lager, so I am extremely hydrated. Is this a norma		For me, the effects were:	
	thing, or should I switch back to max VG?		Coughing Shortness of breath Sore/itchy lungs Congestion Digestive Abnormalities Massive taste change Tingling/warmth in extremities Dizziness Insomnia Acne. Lots.	
	💭 14 Comments 🛱 Award 🧭 Share 💭 Save 🚥		More or less in that order.	
	Posted by 🔄 4 years ago	Ļ	There are probably more I'm not thinking of, but I'm sure y'all have some to offer.	
0 4	News sites running wild with new vaping study. They're saying that vap causes same type of cell damage as COPD. Anyone got any insight?		Dusically, you may reer nice sint. Stay hydrated. Talk to your doctor, but don't parte.	
	nsfw			
	☐ 2 Comments → Share ☐ Save ···			

- **367** unique terms emerged, of which 120 were concerns grouped into **14** categories
- Example: "cancer" mentioned 195 times "cough" 166 times
- Most cited categories of concerns: respiratory (3,957), addiction (1,147), allergy (643), oral health (589), mental health (191), cardiovascular (278), gastrointestinal (141), oncologic (227), inflammation, (64) neurological (207), dermatological (31), orthopedic (28), sleep (79)
- Conclusion: We provide a case study of how social media surveillance can help automate and routinize discovery of perceived health-related outcomes or safety concerns about e-cigarette use



