

Researchers mine Twitter for cardiovascular disease research

September 28 2016



For years, marketers and other commercial data-miners have been using Twitter's vast database of "tweets" to gauge consumer attitudes and track events. Now medical researchers are getting in on the trend. Researchers from the Perelman School of Medicine at the University of Pennsylvania completed a pilot analysis of archived tweets on cardiovascular disease.

In a study published today in *JAMA Cardiology*, researchers sifted through a sample of approximately ten billion [tweets](#) posted between 2009 and 2015, and found more than 500,000 English-language,

U.S.-originating tweets that related to cardiovascular disease.

"We demonstrated that Twitter can provide important information about heart disease, and represents a unique opportunity to listen to patients and understand more about what they talk about and care about related to [cardiovascular health](#)," said senior author Raina M. Merchant, MD, MSHP, an assistant professor of Emergency Medicine and director of Penn's Social Media and Health Innovation Lab.

Users in this sample who tweeted about cardiovascular themes were older and more likely to be female than the average Twitter user. The tweets mostly concerned risk factors, awareness and management of [cardiovascular disease](#) and related conditions such as diabetes and hypertension. Tweets included facts and statistics, tips, and links to new research related to heart health. Among examples: "Chronic Health Failure: Iron deficiency was found to be associated with 58% increased risk." "October is Sudden Cardiac Arrest Month. How can you protect yourself and your loved ones?" "Exercise 'just as good as drugs' for treating heart failure and stroke." "Working out for just 30 min a day, 5 days a week may help protect your body against diabetes." Twitter is a free online social messaging and "microblogging" service with more than 300 million active users worldwide. Twitter messages are 140 characters in length, and although private messages are possible, most "tweets" are public and go, at the rate of half a billion per day, into Twitter's ever-expanding archive which now includes roughly one trillion tweets. Twitter offers researchers several options for accessing these data, including high-cost access to the full database ("full firehose"), lower-cost access to a randomly sampled tenth of the database ("decahose") and free access to a 1/100th sample of the database ("Twitter spritzer").

Merchant's team used a combination of the decahose and spritzer options covering a period from July 2009 to February 2015. For finer-grained analysis they took a random subsample of 2,500 tweets and coded the

contents of each – "self-reported diagnosis," "news," "advertisement," "sentiment," "symptoms" – to assess the incidence of tweets in different categories. For example, 42 percent of the tweets in the 2,500-tweet sample contained references to cardiovascular risk factors.

The information gleaned from the exploratory study is basic, but has paved the way for deeper research by demonstrating that Twitter data can be mined to obtain clear and potentially useful information.

Merchant and her colleagues now are beginning a randomized clinical trial in which people with hypertension will join in a Twitter community with other participants and care providers, to see whether exposure to "heart health" messages by this medium lowers blood pressure.

"We are currently also working on using Twitter for epidemiologic purposes and mapping hypertension and diabetes across the US using Twitter data," Merchant said.

Provided by Perelman School of Medicine at the University of Pennsylvania

Citation: Researchers mine Twitter for cardiovascular disease research (2016, September 28) retrieved 1 May 2024 from

<https://medicalxpress.com/news/2016-09-twitter-cardiovascular-disease.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--