

# UC Davis

## Dermatology Online Journal

### Title

Correlation between cancer incidences and Google searches in the United States

### Permalink

<https://escholarship.org/uc/item/1m07s7b5>

### Journal

Dermatology Online Journal, 22(9)

### Authors

Wehner, Mackenzie R  
Nead, Kevin T  
Linos, Eleni

### Publication Date

2016

### DOI

10.5070/D3229032532

### Copyright Information

Copyright 2016 by the author(s). This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

## Abstract

### Correlation between cancer incidences and Google searches in the United States

Mackenzie R. Wehner MD, MPhil<sup>1</sup>, Kevin T. Nead MD, MPhil,<sup>2</sup> Eleni Linos MD, DrPH<sup>3</sup>

Dermatology Online Journal 22 (9)

<sup>1</sup>Department of Dermatology, University of Pennsylvania

<sup>2</sup> Department of Radiation Oncology, University of Pennsylvania

<sup>3</sup> Department of Dermatology, University of California, San Francisco

---

## Introduction

Despite being highly prevalent, keratinocyte carcinomas (basal cell and squamous cell carcinomas) lack nationwide registries. Internet search data has emerged as a new method to evaluate previously difficult to quantify public health outcomes and may be useful in keratinocyte carcinoma research.

## Objective

We aimed to evaluate whether Google search density correlated with known incidences of common cancers in the United States.

## Methods

We used the Center for Disease Control's National Program of Cancer Registries age-adjusted cancer incidences (2008-2012). We collected Google search data, normalized for total search volume, using Google trends ([google.com/trends](http://google.com/trends)). We collected data on the ten most incident cancers in the United States: lung, breast, colon, prostate, melanoma, endometrial, bladder, thyroid, Non-Hodgkin's lymphoma, kidney/renal pelvis. We utilized Pearson's correlation coefficient to evaluate the relationship between known cancer incidence and Google search density by state.

## Results

Four cancers (endometrial, bladder, thyroid, kidney/renal pelvis) had insufficient Google search quantity among individual states to be evaluated. Lung cancer ( $R^2=0.70$ ,  $p<0.001$ ), colon cancer ( $R^2=0.60$ ,  $p<0.001$ ), melanoma ( $R^2=0.42$ ,  $p=0.002$ ), and Non-Hodgkin's lymphoma ( $R^2=0.47$ ,  $p=0.006$ ) had statistically significant correlations between actual incidences and Google searches. Breast and prostate cancer incidences were not correlated ( $p>0.05$ ).

## Discussion

Four of the six highly incident cancers evaluated had statistically significant correlations between known incidence and Google search density. Internet search data may be a novel tool to estimate geographical incidence and prevalence of disease. This methodology may be particularly useful for keratinocyte carcinomas, which currently lack nationwide registries.