Data in Brief 10 (2017) 457–464

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Infodemiological data concerning silicosis in the USA in the period 2004–2010 correlating with real-world statistical data

Nicola Luigi Bragazzi [a,*], Guglielmo Dini [b], Alessandra Toletone [c], Francesco Brigo [d,e], Paolo Durando [b,f]

[a] School of Public Health, Department of Health Sciences (DISSAL), University of Genoa, Via Antonio Pastore 1, Genoa 16132, Italy
[b] Department of Health Sciences, Postgraduate School in Occupational Medicine, University of Genoa and Occupational Medicine Unit, I.R.C.C.S. University Hospital San Martino - IST National Institute for Cancer Research, Genoa, Italy
[c] Department of Health Sciences, Postgraduate School in Occupational Medicine, University of Genoa, Genoa, Italy
[d] Department of Neurology, Franz Tappeiner Hospital, Merano, Italy
[e] Department of Neurological, Biomedical, and Movement Sciences, University of Verona, Italy
[f] Unità operativa Medicina del lavoro, IRCCS AOU San Martino-IST, Genoa, Italy

ARTICLE INFO

Article history:
Received 21 August 2016
Received in revised form
25 October 2016
Accepted 3 November 2016
Available online 13 November 2016

Keywords:
Infodemiology and infoveillance
Internet
Occupational medicine and hygiene
Web 2.0
Work-related diseases

ABSTRACT

This article reports data concerning silicosis-related web-activities using Google Trends (GT) capturing the Internet behavior in the USA for the period 2004–2010. GT-generated data were then compared with the most recent available epidemiological data of silicosis mortality obtained from the Centers for Disease Control and Prevention for the same study period. Statistically significant correlations with epidemiological data of silicosis ($r=0.805$, p-value <0.05) and other related web searches were found. The temporal trend well correlated with the epidemiological data, as well as the geospatial distribution of the web-activities with the geographic epidemiology of silicosis.

* Corresponding author.
E-mail address: robertobragazzi@gmail.com (N.L. Bragazzi).

http://dx.doi.org/10.1016/j.dib.2016.11.021

Specifications Table

Subject area	*Medicine*
More specific subject area	*Occupational medicine*
Type of data	*Figure, tables*
How data was acquired	*Outsourcing of Google Trends site and the Centers for Disease Control and Prevention (CDC) site*
Data format	*Raw, analyzed*
Experimental factors	*Google Trends search volumes were obtained through graphs and heat-maps*
Experimental features	*Validation of Google Trends-based data with "real-world" data taken from the CDC site was performed by means of correlational analysis*
Data source location	USA
Data accessibility	Data are within this article

Value of the data

- Google Trends (GT)-based data (*infodemiological* data) could be useful for scientific community, researchers and occupational physicians in that they show good correlation with "real world" data obtained from the Centers for Disease Control and Prevention site, thus proving to be reliable.
- These data could be further statistically processed, analyzed, refined and validated in such a way to complement traditional surveillance of silicosis, providing data quicker and in real time.
- These data could be used to understand occupational diseases-related web activities.
- To our knowledge, this is the first analysis of web search behavior related to an occupational disease, namely silicosis, carried out both in quantitative and qualitative terms.

1. Data

This article contains infodemiological data on silicosis searched in the USA in the study period 2004–2010, obtained from Google Trends (GT) (Fig. 1). These data well correlated with "real-world" data obtained from the Centers for Disease Control and Prevention (CDC) site for the same study period (Tables 1–3).

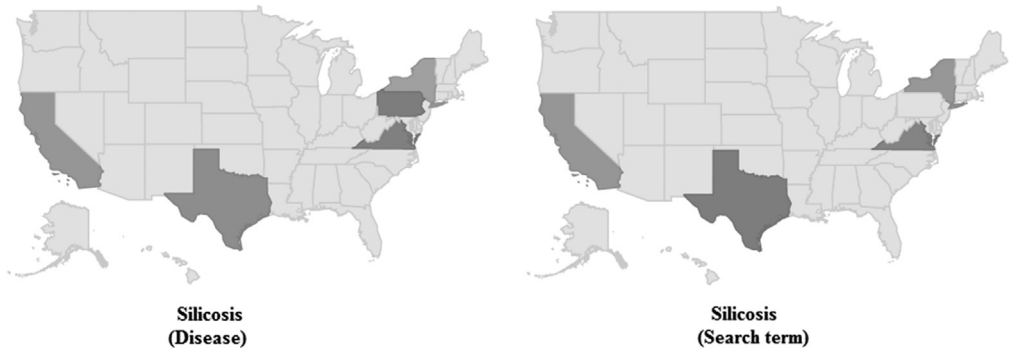

Fig. 1. Google Trends-generated heat-map showing the regional interest for silicosis in the USA. In particular, it can be noticed that silicosis-related web searches are concentrated in some counties (namely, California, Texas, New York, Pennsylvania, and Virginia).