

## **COVID-19 predictability in Portugal using Google Trends**

**Gustavo Paiva Monteiro**

*G Paiva Monteiro<sup>1,2</sup>, S Barroso Travassos<sup>1</sup>, J Miguel Carvalho<sup>1</sup>,  
R Marques<sup>3</sup>, D Fernandes<sup>1</sup>, D Portela<sup>1,4</sup>*

<sup>1</sup>Public Health, ACeS Entre Douro e Vouga I - Feira/Arouca, Santa Maria da Feira, Portugal

<sup>2</sup>Faculdade de Economia da Universidade do Porto, Porto, Portugal

<sup>3</sup>Centro Hospitalar Entre Douro e Vouga, Santa Maria da Feira, Portugal

<sup>4</sup>Faculdade de Medicina da Universidade do Porto, Porto, Portugal

Contact: gpmonteiro@arsnorte.min-saude.pt

### **Background:**

Infodemiology is a valuable public health tool which analyses internet sources and real-time data to establish populational trends, thus assisting disease surveillance. Google trends (GT) is a popular infodemiology source widely used in previous reports to study the correlation between internet relative search volume (RSV) and surges in various pathologies. This study aims to explore the association between GT and Covid-19 symptoms and to access the possibility of predicting new surges from internet searches.

**Methods:**

Individual data was extracted from GT RSV on four main Covid-19 related symptoms in the Health category (fever, headache, cough and shortness of breath) in Portugal between 2020/03/02 and 2021/02/15, corresponding to each of the identified surges of daily new cases (DNC) in Portugal, retrieved from GitHub. Pearson's-correlation coefficient was used for assessment. Additionally, a 14 days time-lag correlation analysis between data for the same period of time was performed.

**Results:**

Statistically significant correlations were found between 'fever' web searches and the DNC in the first ( $p=0.02$ ) and third ( $p=0.02$ ) wave. No statistically significant correlations were found between any other variables. Through time lag analysis, we found a maximum Pearson association between web searches for 'cough' and the DNC during the first wave at 14 days ( $r=0.55$ ), as well as during the third wave, with a maximum association at 3 days time lag ( $r=0.55$ ).

**Conclusions:**

Monitoring behaviour and public interest in health related issues, such as crisis, is necessary and may help in the establishment of better and target oriented health policies. Despite previously stated potential, constraints such as the exclusion of social media platforms or internet users' representativeness, could partly explain our limited results for portuguese predictability of new COVID-19 surges. A better understanding of GT's algorithm may lead to more detailed and precise data.

**Key messages:**

- Monitoring behaviour and public interest in health related issues is necessary to establish better and more targeted health policies.
- Google trends seems to be an helpful infodemiology source, but doesn't allow for full representation of the population and needs better understanding.