

# How informative are web searches for risk communication during COVID-19 in Germany?

Kaja Kristensen

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Eva Lorenz

Bernhard Nocht Institute for Tropical Medicine, Department of Infectious Disease Epidemiology, Hamburg, Germany

Jürgen May

Bernhard Nocht Institute for Tropical Medicine, Department of Infectious Disease Epidemiology, Hamburg, Germany

Ricardo Strauss (✉ [ricardo.strauss@bntm.de](mailto:ricardo.strauss@bntm.de))

Bernhard Nocht Institute for Tropical Medicine, Department of Infectious Disease Epidemiology, Hamburg, Germany

---

## Research Article

**Keywords:** Google Trends, risk communication, covid-19, public concern

**Posted Date:** October 2nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-84751/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Scientific Reports on October 2nd, 2020. See the published version at <https://doi.org/10.1038/s41598-021-85873-4>.

# Abstract

**Background:** Risk communication during pandemics is an element of paramount importance. Understanding the level of public concern implicates expensive and time-consuming surveys. We hypothesize that the relative search volume from Google Trends could be used as an indicator of public concern towards prevention measures as well as of the adequacy of the official messages spread.

**Methods:** The search terms '*RKI*', '*corona*' and '*protective mask*' in German language were shortlisted. Cross-correlations between these terms and the reported cases from February 15<sup>th</sup> to April 27<sup>th</sup> were conducted for each German federal state. The findings were contrasted against a timeline of official communications concerning COVID-19.

**Results:** The highest correlations of the term '*RKI*' (Robert Koch Institute, national public health authority in Germany) with reported COVID-19 cases were found between lags of -2 and -12 days, meaning web searches were already performed two to twelve days before case numbers increased. A similar pattern was seen for the term '*corona*'. Cross-correlations indicated that most searches on '*protective mask*' were performed six to twelve days after the increase of cases.

**Conclusions:** The results for the term '*protective mask*' indicate some degree of confusion in the population, which is supported by the contradictory recommendations on the wearing of face masks over time. In addition, the relative search volumes could be a useful tool to provide timely information on location-based risk communication strategies.

## Introduction

COVID-19 (Corona Virus Disease 2019) is caused by the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV-2). The global spread of the virus led to COVID-19 being classified as a pandemic in March 2020, affecting the lives of billions of people. At the time of writing, more than 29 million confirmed cases and nearly 950.000 deaths have been reported worldwide<sup>1</sup>. In Germany, around 260.000 confirmed cases and 9.400 deaths have been reported in its first wave<sup>2</sup>.

The secondary attack rate of the infection and its enormous deadly toll highlighted the importance of public preventive measures to contain the spread of the virus and reduce the burden on health systems. In addition to government-imposed contact restrictions and other policies, there are a number of individual hygiene measures such as social distancing, regular hand washing and the wearing of masks in order to contain the spread and for self-protection<sup>3</sup>.

Risk communication strategies aim at increasing the capacity of the public to act as an effective response partner by encouraging prevention measures among the population<sup>4,5</sup>. In past research, several factors have been found to influence the persuasiveness of communications. These include authority credibility, transparency, perceived expertise and motivation to strive for a common interest<sup>6</sup>.

During a pandemic, the actual change in behaviour is of great importance as the implementation of measures on an individual level is imperative. In several health behaviour theories, such as the Health Belief Model or the Protection Motivation Theory, the perceived severity of an event is, among other factors, a key element in the process of developing and adapting health behaviour<sup>7,8</sup>. Therefore, the timing, and more importantly, the content of risk communication procedures need to be appropriately adapted to its target population.

Collecting data on the perceived severity and public concern regarding an event is both time-consuming and costly. Therefore, in order to retrieve faster and cheaper insights, it may be worthwhile to look at web searches since the Internet is a common source for finding health-related information<sup>9</sup>. Web searches are not only a valuable resource for individuals seeking health information, but also for the scientific community, as search queries contain geospatial and temporal data. Numerous studies have shown how analysing web search queries can assist in describing and predicting outbreaks especially in the context of sparse data<sup>10-12</sup>. Less frequently, web searches are analysed to describe the concern in the population in connection with an event or to get an idea of the general mood towards certain topics<sup>13</sup>.

The aim of this study is to explore the usefulness of clustered web search queries as an indicator of public concern for informing and adapting risk communication strategies in the context of the ongoing pandemic in Germany.

## **Materials And Methods**

### **2.1. Data sources**

#### **2.1.1. Officially reported cases**

Daily incidence data on officially reported COVID-19 cases in Germany from February 15th 2020 to April 27th 2020 were obtained from the Robert Koch Institute (RKI) which is the national public health authority in Germany. Data are publicly available ([https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea\\_0?geometry=-19.624%2C46.270%2C35.879%2C55.886](https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea_0?geometry=-19.624%2C46.270%2C35.879%2C55.886)). The daily number of new cases were retrieved for each of the 16 federal states in Germany separately as well as clustered national data.

#### **2.1.2. Web search queries**

Google Trends provides access to Internet search patterns by analysing web queries on the Google search browser and other affiliated Google sites. It provides a relative search volume (RSV), i.e. the query share of a particular term for a given location and time period, normalized by the highest query share of that term over the given time period. Hence, data is presented on a relative scale from 0 to 100, with 0 representing no related queries and 100 representing the maximum number of related queries.

In the period from February 15th to April 27th, various German search terms concerning information seeking and protective behaviour in the context of COVID-19 were explored. Search query volumes related to the disease (i.e. '*corona*', '*coronavirus*', '*covid-19*' and '*covid 19*'), information sources (i.e. '*corona podcast*', '*corona info*', '*corona hotline*', '*Robert Koch Institut*', '*RKI*', and '*Gesundheitsamt*'), as well as preventive measures including disinfection and handwashing (i.e. '*Händewaschen*', '*Hände waschen*', '*Händedesinfektion*', '*Händedesinfizieren*' and '*Hände desinfizieren*'), and facial masks (i.e. '*Mundschutz*', '*Maske*', '*Gesichtsmaske*', '*Schutzmaske*', '*Stoffmaske*') were retrieved. To shortlist the final set of terms, all terms within one category were entered simultaneously to compare their frequencies. The most frequently searched terms were then selected. Finally, daily data from February 15<sup>th</sup> to April 27<sup>th</sup> 2020 on web searches for '*RKI*', '*Mundschutz*' (protective mask) and '*corona*' were obtained separately from Google Trends (<https://trends.google.com/trends/>) on May 1<sup>st</sup> 2020. Data were retrieved for each federal state in Germany and nationally clustered for the whole country. Next, data was downloaded in csv-format.

### **2.1.3. Risk communication milestones**

Important announcements in the period from February 15<sup>th</sup> to April 27<sup>th</sup> were retrieved from timelines published by the German Ministry of Health (MoH) (<https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus.html>) and the World Health Organization (WHO) (<https://www.who.int/news-room/detail/29-06-2020-covidtimeline>). In addition, short reports during the same period from a major national public media network (<https://www.tagesschau.de/faktenfinder/hintergrund/corona-chronik-pandemie-101.html>) were screened.

## **2.2. Data analysis**

Statistical analyses were conducted using SPSS version 25 and R version 3.6.2. Line plots in combination with histograms, and cross-correlations were generated to assess the association between Google search trends and the officially reported COVID-19 cases.

To characterize time dependence between time-series from two different data sources, cross-correlations were used. The time dependence between two time-series is termed as lag and indicates the direction of the two time-series correlated. A lag of -1 suggests that a peak in Google trends precedes a peak in the officially reported COVID-19 cases by one day, and vice versa for positive lags.

Both time series were normalized to account for the different units and scales (RSV and absolute number of daily new cases). Thresholds for interpreting a correlation coefficient suggested by Hinkle et al. were applied<sup>14</sup>.

The risk communication milestones were reviewed based on their content and relevance for the present study by two independent reviewers. Disagreement was solved by discussion.

# **Results**

### **3.1. Time trends**

Histograms of officially reported cases on a national level in combination with line graphs of Google search queries for '*RKI*', '*corona*' and '*protective masks*' at state level are shown in Figures 1 to 3. Most relevant communications and announcements are also presented in these figures. Figure 1 and 2 show that the RSV of the search terms '*RKI*' and '*corona*' reach their peak (mid-March) before the peak of the epidemic (beginning of April) for all German federal states. The RSV of '*corona*' shows a second peak of searches in mid-April. Meanwhile the RSV of '*protective mask*' presents with three peaks with the highest peak at a point where the curve of confirmed cases is already flattening (end of April) (Figure 3).

### **3.2. Time-series cross-correlations**

The linear association between officially reported COVID-19 cases and Google Trend patterns for the search terms '*RKI*', '*corona*' and '*protective mask*' were assessed using cross-correlations, as shown in Table 1 to Table 3.

Visual assessment of Table 1 and Table 2, which refer to the search terms '*corona*' and '*RKI*' respectively, reveals that the correlations reach a high level in most federal states with the exception of Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland. These federal states are described separately from the rest in the following.

For the group of federal states showing high correlations, the search term '*corona*' reached high coefficients ( $r = 0,707$  to  $r = 0,804$ ) preceding the reported cases for a couple of days (range of lags: -5 to -12 days). In the same group of states, high correlations ( $r = 0,711$  to  $r = 0,856$ ) for '*RKI*' were reached when searches preceded the reported cases between two to ten days (range of lags: -2 to -10 days).

In the federal states Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland only moderate correlations ( $r = 0,453$  to  $r = 0,688$ ) were reached for the term '*corona*'. Correlations were highest when the searches preceded the reported cases (range of lags: -5 to 14 days). For these states, a similar trend was observed for the term '*RKI*' ( $r = 0,449$  and  $r = 0,682$ ; range of lags: -10 to 0 days).

A different picture results from the search term '*protective mask*'. In contrast to the other two search terms, reported cases preceded the web searches for '*protective mask*' by five to 14 days (range of lag +5 to +14 days) with correlation coefficients indicating low to moderate positive correlations in all federal states. ( $r = 0,400$  to  $r = 0,602$ )

### **3.3. Risk communication milestones**

Following the timeline of the data period (February 15<sup>th</sup> to April 27<sup>th</sup>) chronologically (Figure 4), an important announcement was made on March 2<sup>nd</sup> when the RKI classified the risk as "moderate". Just over a week later, on March 11<sup>th</sup>, COVID-19 was classified as a pandemic by WHO. On March 13<sup>th</sup>, most of the German states announced that they would temporarily close the schools. On March 17<sup>th</sup> the RKI classified the risk as "high" and all schools and day care centres were closed. Five days later, on March

22<sup>nd</sup>, the Federal Government and the federal states agreed on guidelines for restricting social contacts. Central elements were the prohibition of staying in public places with more than one other person not living in the same household, and the closure of restaurants, bars, and service providers with physical contact, such as hairdressers or massage studios.

On March 31<sup>st</sup>, the German government expressed its opposition to the obligation to wear masks as medical masks should be reserved for health care personnel. On April 2<sup>nd</sup>, the RKI extended its recommendation regarding wearing masks. While wearing masks was previously recommended for people with respiratory diseases, they now recommended it for the general population. A few days later (April 6<sup>th</sup>), compulsory wearing of masks was introduced in Jena, a city in the federal state of Thuringia. In the following weeks, several federal states introduced compulsory wearing of masks, while there was no nationwide regulation. Meanwhile, it was announced on April 15<sup>th</sup> that the contact restrictions, announced on March 22<sup>nd</sup>, would be extended until May 3<sup>rd</sup> and that the schools would reopen on May 4<sup>th</sup>. On April 17<sup>th</sup>, the Federal Minister of Health stated on national television that the first wave of infection had been well overcome and that the country should gradually return to normality while staying well-prepared and alert. On April 22<sup>nd</sup>, a nationwide mask obligation was announced which came into force on April 27<sup>th</sup>.

## Discussion

This study aimed to explore the usefulness of clustered web search queries as an indicator of the degree of public concern for informing and adapting risk communication strategies in the context of the ongoing COVID-19 pandemic in Germany.

The findings showed that different search terms related to the pandemic had different trends over the course of the first wave. While searches for the terms '*RKI*' and '*corona*' reached their peak before the number of cases peaked, the term '*protective mask*' reached its peak when the curve was already flattening.

The fact that the RSV of '*RKI*' followed the same trend as the searches for the term '*corona*' could indicate that the institute is trusted by the public. At times when the need for information on '*corona*' was at its highest, web searches for the national public health authority were also at their highest level. This hypothesis is supported by data from the GESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in Germany. The data shows, that the RKI enjoys the greatest public trust in dealing with the pandemic compared to other institutions<sup>15</sup>. The temporal relationship between the RSV of '*corona*' and the reported cases suggests that the term has a certain predictive value that could be used in the context of sparse data. A study by Effenberger et al. examined this temporal relationship in different countries and also found that the peaks of web searches for '*coronavirus*' preceded the peaks of the reported cases by 11.5 days on average<sup>16</sup>. However, the visual examination of data from this study also suggests that

the RSV is closely linked to the announcements of new policies or policy changes, since the second peak of the RSV coincided exactly with the announcement of the extension of the contact restriction policies.

The observation that an increase in web searches on '*RKI*' and '*'corona'*' preceded the increase of reported cases was also described in a similar study that monitored search terms related to COVID-19 such as '*corona*', '*handwashing*' and '*masks*'. In contrast to this study, however, web searches for '*face mask*' and '*surgical mask*' behaved in the same way and were performed before most cases were reported<sup>13</sup>.

Now, why did the search term '*protective mask*' show a different trend in Germany? The long absence of a nationwide uniform obligation to wear a mask, while such obligations have already been introduced at local level, could have led to confusion and an increased need for research on this topic on the Internet. In addition, there have been conflicting recommendations from different institutions as to whether masks should be worn by the general public or primarily by healthcare professionals. This public debate and uncertainty is reflected in the web searches for '*protective masks*'.

The previous example shows how RSV of search terms can help to inform risk communication strategies in order to assess the degree of public concern about certain issues. In a similar study that analysed trends in influenza-related search queries, it was also suggested that Google Trends identifies public perceptions of the influenza threat rather than representing the incidence of influenza<sup>17</sup>.

On the other hand, the RSV of search terms can help to evaluate the adequacy of the announcements and messages disseminated to the target public. We were able to show that at a time when web searches for '*corona*' and the public concern was at its highest level, relatively few people informed themselves about preventive measures such as the wearing of protective masks. In retrospect, this is important information for the risk communication strategy, as it would be desirable for the public to inform itself about preventive measures at the peak of the pandemic, or even better, earlier.

The results of this study are subject to some limitations that should be considered. First of all, the intentions behind the web searches analysed are unclear. Different motives such as seeking health information, updating oneself on current policies, or purchasing face masks might have led individuals to search for the terms and not necessarily represent certain behaviours towards any preventive measures or the disease itself. Furthermore, it is known from literature that disease-related web searches are subject to media-driven interest known as "celebrity phenomena"<sup>18</sup>, which basically refers to increased internet queries for certain terms that are temporarily popular in the media.

Another important limitation arises from the limited applicability of the analysis method in federal states with low absolute populations, as was the case in Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland. As the results showed, no high correlations were found in these states. It is therefore questionable to what extent the presented method can be applied in areas with a low population.

Some search terms that are highly relevant for disease prevention and personal protection, namely hand washing and disinfection, could not be included in the study due to the low search frequency at state level. There could be various reasons why these terms were not searched for as frequently. It is possible that there was no need in the population to inform themselves about these topics.

As a final limitation, it should be mentioned that RSVs do not provide information about the absolute number of web searches. Therefore, it is not possible to draw conclusions about the actual amount of searches performed, but only about their trends over time.

Nevertheless, in this study we could present a novel tool for the fast assessment of public concern in times of an ongoing pandemic and show how web searches can also inform about the adequacy of risk communication strategies.

## Conclusion

In the context of an ongoing pandemic using RSVs of key terms related to the disease can help to inform about the public concern. Results suggest that the data can also be used to assess the appropriateness of the messages disseminated. The example of the protective masks showed how initial public controversies and contradictions were reflected in the search behaviour of the population, which only informed itself about protective masks when the curve already flattened. Further studies are needed to assess what is the real impact of inaccurate risk communication in the actual development of the pandemic.

## Declarations

### Conflict of Interest

The authors declare no competing interests.

### Funding Declaration

None.

### Data availability

The Google Trends dataset generated for the current study is available and can be replicated at:  
<https://trends.google.com/trends/>

Daily incidence data on officially reported COVID-19 cases are publicly available: [https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea\\_0?geometry=-19.624%2C46.270%2C35.879%2C55.886](https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea_0?geometry=-19.624%2C46.270%2C35.879%2C55.886)

### Authors Contribution

RS developed the concept and design of the study and made substantial contributions to the analysis, interpretation of data and critically revised the manuscript. KK performed the analyses, interpreted the data, and wrote the first draft of the article and made substantial contributions to the conception of the work. EL made substantial contributions to the analysis, interpretation of data and critically revised the manuscript. JM revised the last version of the manuscript before submission. All authors have read and approved the manuscript in its current form.

### Ethics declarations

Ethical approval and consent to participate were not necessary as the study was based on openly available aggregated data.

## Abbreviations

COVID-19 - Corona Virus Disease 2019

MoH – Ministry of Health

RKI – Robert Koch Institut

RSV – Relative Search Volume

SARS-CoV-2 - Severe Acute Respiratory Syndrome Coronavirus 2

WHO – World Health Organization

## References

- 1 European Centre for Disease Prevention and Control (ECDC). *COVID-19 situation update worldwide, as of 3 August 2020*, <<https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>> (2020).
- 2 Robert Koch Institute. *COVID-19: Fallzahlen in Deutschland und weltweit - Stand: 3.8.2020*, <[https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Fallzahlen.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html)> (2020).
- 3 Robert Koch Institute. Leitfaden für den Öffentlichen Gesundheitsdienst zum Vorgehen bei Häufungen von COVID-19. (Berlin, 2020).
- 4 Blendon, R. J. *et al.* Public response to community mitigation measures for pandemic influenza. *Emerging infectious diseases* **14**, 778-786, doi:10.3201/eid1405.071437 (2008).
- 5 Holmes, B. J. Communicating about emerging infectious disease: The importance of research. *Health Risk Soc* **10**, 349-360, doi:10.1080/13698570802166431 (2008).

- 6 Aven, T. & Renn, O. in *Risk Management and Governance: Concepts, Guidelines and Applications* 159-179 (Springer Berlin Heidelberg, 2010).
- 7 Rosenstock, I. M. Why people use health services. *The Milbank Memorial Fund quarterly* **44**, Suppl:94-127 (1966).
- 8 Rogers, R. W. A Protection Motivation Theory of Fear Appeals and Attitude Change1. *J Psychol* **91**, 93-114, doi:10.1080/00223980.1975.9915803 (1975).
- 9 Baker, L., Wagner, T. H., Singer, S. & Bundorf, M. K. Use of the Internet and e-mail for health care information: results from a national survey. *Jama* **289**, 2400-2406, doi:10.1001/jama.289.18.2400 (2003).
- 10 Carneiro, H. A. & Mylonakis, E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **49**, 1557-1564, doi:10.1086/630200 (2009).
- 11 Milinovich, G. J. *et al.* Using internet search queries for infectious disease surveillance: screening diseases for suitability. *BMC infectious diseases* **14**, 690, doi:10.1186/s12879-014-0690-1 (2014).
- 12 Strauss, R. *et al.* Investigating the utility of Google trends for Zika and Chikungunya surveillance in Venezuela. *BMC public health* **20**, 947, doi:10.1186/s12889-020-09059-9 (2020).
- 13 Husnayain, A., Fuad, A. & Su, E. C.-Y. Applications of Google Search Trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. *Int J Infect Dis* **95**, 221-223, doi:10.1016/j.ijid.2020.03.021 (2020).
- 14 Hinkle, D., Wiersma, W. & Jurs, S. *Applied Statistics for the Behavioral Sciences*. (Houghton Mifflin, 2003).
- 15 Gesis Panel Team. (GESIS Datenarchiv, Köln. ZA5667 Datenfile Version 1.1.0, <https://doi.org/10.4232/1.13520>, 2020).
- 16 Effenberger, M. *et al.* Association of the COVID-19 pandemic with Internet Search Volumes: A Google TrendsTM Analysis. *Int J Infect Dis* **95**, 192-197, doi:<https://doi.org/10.1016/j.ijid.2020.04.033> (2020).
- 17 Dugas, A. F. *et al.* Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **54**, 463-469, doi:10.1093/cid/cir883 (2012).
- 18 Milinovich, G. J., Williams, G. M., Clements, A. C. A. & Hu, W. Internet-based surveillance systems for monitoring emerging infectious diseases. *Lancet Infect Dis* **14**, 160-168, doi:10.1016/S1473-3099(13)70244-5 (2014).

# Tables

Table 1. Cross-correlation coefficients displaying the linear association between the RSV of the term '*corona*' and officially reported COVID-19 cases

Lag	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Baden-Württemberg	0,685	0,736	0,771	0,779	0,761	0,742	0,731	0,73	0,728	0,718	0,689	0,644	0,597	0,544	0,497
Bavaria	0,734	0,772	0,791	0,782	0,763	0,75	0,736	0,72	0,703	0,678	0,631	0,58	0,534	0,489	0,446
Berlin	0,627	0,689	0,734	0,742	0,736	0,73	0,73	0,734	0,746	0,718	0,679	0,626	0,57	0,515	0,462
Brandenburg	0,59	0,613	0,645	0,62	0,599	0,61	0,577	0,565	0,58	0,588	0,571	0,538	0,513	0,469	0,422
Bremen	0,262	0,259	0,315	0,311	0,341	0,357	0,412	0,442	0,452	0,391	0,391	0,375	0,342	0,339	0,293
Hamburg	0,5	0,559	0,601	0,641	0,677	0,704	0,72	0,76	0,804	0,802	0,792	0,762	0,737	0,681	0,627
Hesse	0,597	0,644	0,675	0,701	0,707	0,661	0,665	0,686	0,695	0,699	0,678	0,643	0,593	0,549	0,486
Lower Saxony	0,649	0,699	0,728	0,737	0,726	0,72	0,729	0,733	0,747	0,738	0,711	0,674	0,627	0,581	0,534
Mecklenburg-Western Pomerania	0,416	0,457	0,515	0,532	0,515	0,501	0,486	0,53	0,606	0,688	0,678	0,66	0,589	0,531	0,504
Northrhine-Westphalia	0,584	0,623	0,652	0,671	0,677	0,671	0,688	0,711	0,734	0,75	0,747	0,713	0,673	0,64	0,584
Rhineland-Palatinate	0,609	0,66	0,703	0,735	0,75	0,749	0,754	0,77	0,773	0,784	0,774	0,751	0,722	0,671	0,617
Saarland	0,687	0,673	0,655	0,66	0,632	0,601	0,562	0,549	0,51	0,479	0,447	0,408	0,373	0,346	0,31
Saxony	0,621	0,676	0,714	0,725	0,709	0,706	0,731	0,749	0,768	0,756	0,715	0,674	0,629	0,597	0,551
Saxony-Anhalt	0,59	0,634	0,666	0,697	0,713	0,705	0,737	0,734	0,771	0,74	0,713	0,683	0,662	0,593	0,536
Schleswig-Holstein	0,618	0,668	0,719	0,735	0,724	0,714	0,712	0,698	0,698	0,683	0,647	0,606	0,574	0,531	0,478
Thuringia	0,565	0,608	0,656	0,675	0,663	0,67	0,701	0,724	0,735	0,708	0,661	0,624	0,573	0,557	0,501
Germany	0,672	0,719	0,748	0,757	0,748	0,739	0,735	0,741	0,745	0,735	0,702	0,661	0,615	0,569	0,52

.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

Table 2. Cross-correlation coefficients displaying the linear association between the RSV of the term '*RKI*' and officially reported COVID-19 cases

Lag	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Baden-Württemberg	0,651	0,691	0,743	0,784	0,824	0,830	0,829	0,801	0,785	0,777	0,793	0,776	0,747	0,700	0,653
Bavaria	0,706	0,744	0,775	0,812	0,832	0,827	0,820	0,812	0,782	0,753	0,740	0,720	0,676	0,642	0,602
Berlin	0,643	0,651	0,686	0,707	0,738	0,783	0,800	0,824	0,806	0,768	0,747	0,729	0,710	0,702	0,675
Brandenburg	0,633	0,659	0,626	0,658	0,676	0,632	0,631	0,668	0,643	0,629	0,617	0,587	0,538	0,544	0,528
Bremen	0,276	0,328	0,244	0,337	0,317	0,381	0,347	0,369	0,380	0,362	0,377	0,387	0,422	0,393	0,371
Hamburg	0,528	0,563	0,612	0,657	0,711	0,738	0,779	0,811	0,796	0,802	0,791	0,788	0,787	0,768	0,718
Hesse	0,600	0,641	0,662	0,694	0,732	0,751	0,778	0,780	0,780	0,758	0,754	0,751	0,743	0,718	0,688
Lower Saxony	0,619	0,670	0,697	0,732	0,764	0,789	0,804	0,817	0,809	0,803	0,805	0,789	0,765	0,739	0,717
Mecklenburg-Western Pomerania	0,249	0,289	0,336	0,340	0,398	0,447	0,609	0,563	0,535	0,576	0,580	0,542	0,558	0,673	0,647
Northrhine-Westphalia	0,547	0,595	0,628	0,671	0,706	0,753	0,774	0,799	0,789	0,798	0,797	0,813	0,818	0,806	0,781
Rhineland Palatinate	0,546	0,610	0,661	0,704	0,747	0,779	0,803	0,825	0,845	0,843	0,829	0,856	0,832	0,795	0,759
Saarland	0,661	0,679	0,682	0,665	0,682	0,638	0,599	0,550	0,535	0,496	0,435	0,433	0,412	0,374	0,313
Saxony	0,592	0,637	0,713	0,759	0,778	0,795	0,811	0,802	0,785	0,786	0,795	0,767	0,727	0,695	0,648
Saxony-Anhalt	0,488	0,550	0,567	0,607	0,710	0,742	0,755	0,726	0,752	0,789	0,744	0,740	0,795	0,719	0,668
Schleswig-Holstein	0,564	0,640	0,670	0,715	0,729	0,761	0,791	0,798	0,797	0,749	0,757	0,733	0,741	0,694	0,672
Thuringia	0,563	0,627	0,690	0,662	0,652	0,666	0,671	0,659	0,684	0,711	0,666	0,645	0,610	0,584	0,598
Germany	0,651	0,697	0,734	0,772	0,803	0,821	0,831	0,834	0,824	0,812	0,805	0,792	0,768	0,740	0,702

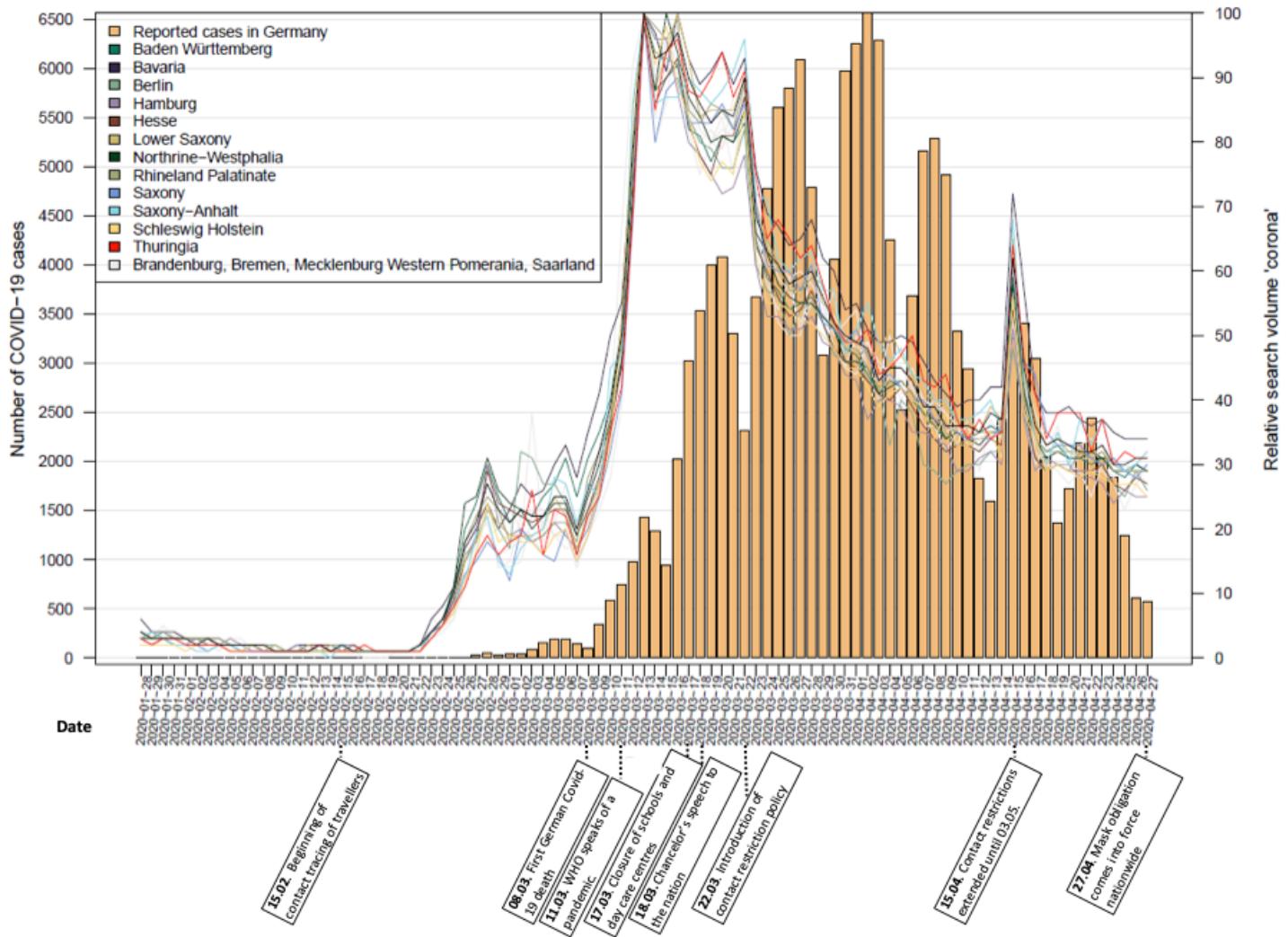
.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

Table 3. Cross-correlation coefficients displaying the linear association between the RSV of the term '*protective mask*' and officially reported COVID-19 cases

Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Baden-Württemberg	0,332	0,298	0,282	0,314	0,368	0,422	0,441	0,411	0,348	0,322	0,355	0,391	0,461	0,512	0,523
Bavaria	0,460	0,417	0,394	0,429	0,467	0,506	0,524	0,489	0,430	0,421	0,444	0,474	0,533	0,582	0,564
Berlin	0,475	0,453	0,441	0,491	0,519	0,577	0,548	0,505	0,458	0,394	0,411	0,413	0,459	0,464	0,421
Brandenburg	0,477	0,467	0,392	0,388	0,420	0,452	0,484	0,463	0,412	0,332	0,320	0,327	0,371	0,464	0,511
Bremen	0,602	0,510	0,390	0,406	0,409	0,394	0,395	0,371	0,291	0,243	0,207	0,238	0,257	0,238	0,203
Hamburg	0,258	0,227	0,216	0,233	0,308	0,338	0,374	0,356	0,332	0,354	0,358	0,422	0,471	0,477	0,510
Hesse	0,472	0,423	0,379	0,403	0,472	0,532	0,560	0,550	0,499	0,408	0,403	0,432	0,504	0,565	0,560
Lower Saxony	0,414	0,409	0,378	0,408	0,447	0,487	0,517	0,497	0,439	0,363	0,371	0,390	0,445	0,499	0,506
Mecklenburg-Western Pomerania	0,099	0,045	0,061	0,055	0,104	0,113	0,181	0,181	0,150	0,101	0,138	0,186	0,252	0,352	0,400
Northrhine-Westphalia	0,383	0,373	0,374	0,367	0,412	0,457	0,491	0,489	0,446	0,381	0,381	0,420	0,462	0,524	0,562
Rhineland-Palatinate	0,356	0,314	0,287	0,329	0,400	0,439	0,446	0,458	0,445	0,401	0,376	0,384	0,403	0,454	0,494
Saarland	0,319	0,273	0,203	0,123	0,140	0,180	0,220	0,247	0,219	0,209	0,161	0,201	0,328	0,491	0,539
Saxony	0,234	0,246	0,243	0,267	0,288	0,308	0,334	0,344	0,342	0,353	0,392	0,405	0,413	0,442	0,485
Saxony-Anhalt	0,310	0,285	0,216	0,271	0,417	0,409	0,379	0,366	0,294	0,176	0,183	0,257	0,311	0,363	0,390
Schleswig-Holstein	0,325	0,287	0,250	0,275	0,329	0,356	0,365	0,351	0,442	0,331	0,300	0,319	0,415	0,471	0,525
Thuringia	0,590	0,487	0,392	0,462	0,521	0,553	0,553	0,519	0,419	0,356	0,321	0,395	0,444	0,455	0,412
Germany	0,394	0,327	0,300	0,326	0,383	0,451	0,474	0,466	0,402	0,355	0,367	0,406	0,459	0,510	0,518

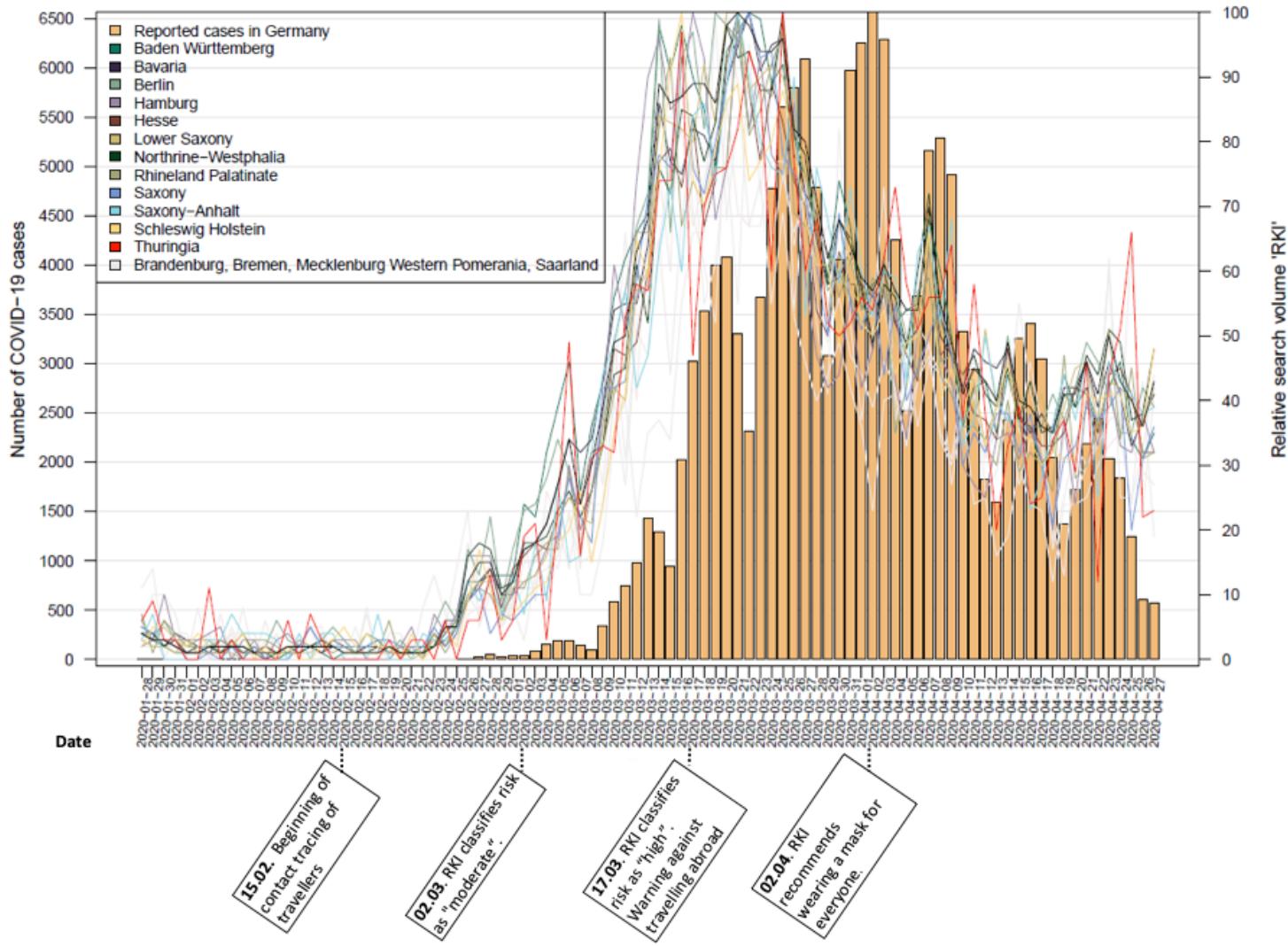
.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

## Figures



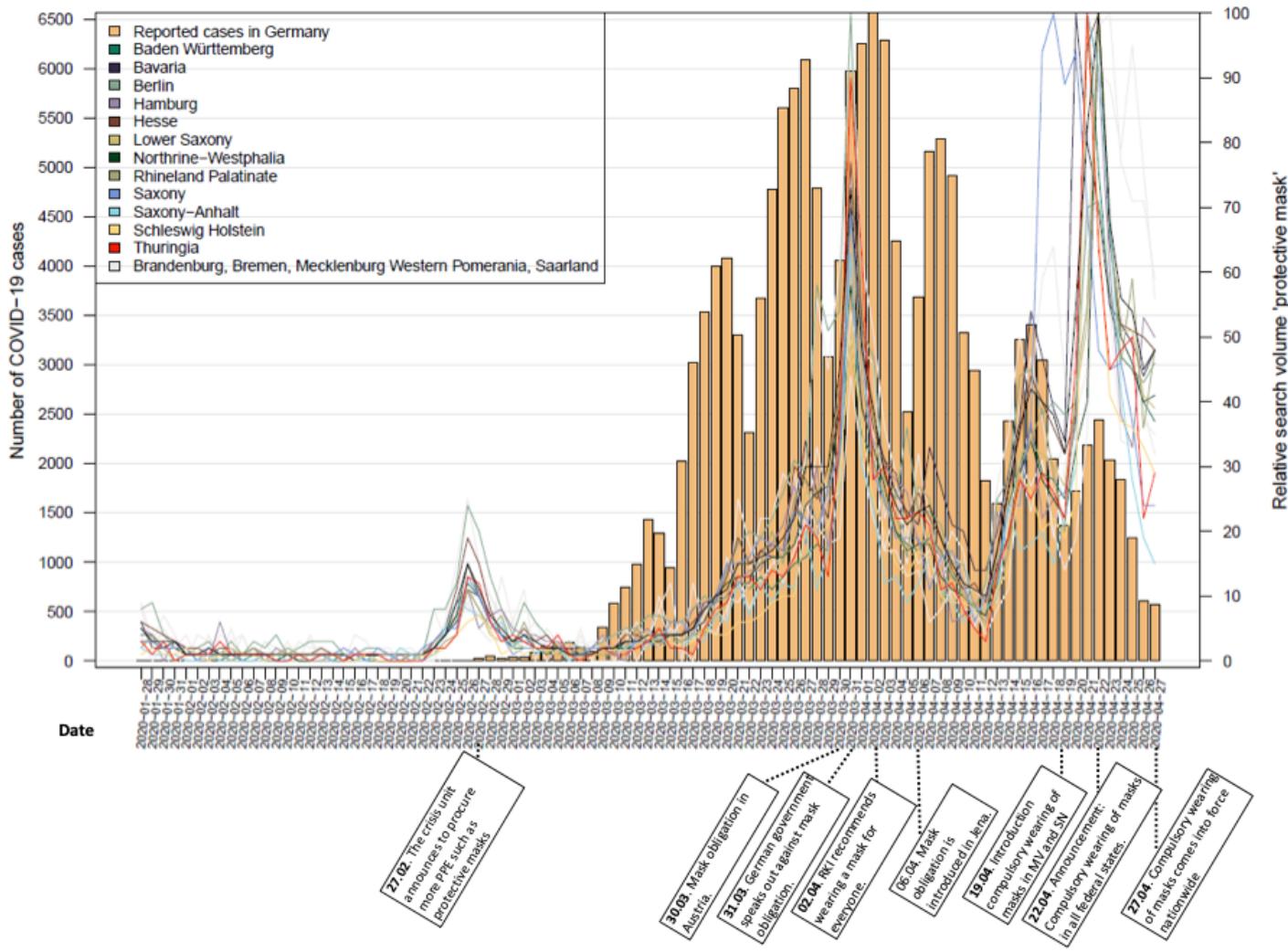
**Figure 1**

Nationally reported COVID-19 cases and RSV for the search term 'corona' at state level



**Figure 2**

Nationally reported COVID-19 cases and RSV for the search term 'RKI' at state level



**Figure 3**

Nationally reported COVID-19 cases and RSV for the search term 'protective mask' at state level



<b>February 15<sup>th</sup></b>	- Beginning of contact tracing of travellers
<b>February 27<sup>th</sup></b>	- The crisis unit announces to procure more PPE such as protective masks
<b>March 2<sup>nd</sup></b>	- RKI classifies risk as "moderate"
<b>March 3<sup>rd</sup></b>	- Prohibition of PPE exporting
<b>March 8<sup>th</sup></b>	- First German Covid-19 death
<b>March 10<sup>th</sup></b>	- Recommendation to cancel events with >1000 participants
<b>March 11<sup>th</sup></b>	- WHO speaks of a pandemic for the first time
<b>March 17<sup>th</sup></b>	- RKI classifies risk as "high" - Closure of schools and day care centres
<b>March 18<sup>th</sup></b>	- Chancellor's speaks to the nation on national TV
<b>March 22<sup>nd</sup></b>	- Introduction of contact restriction policy
<b>March 30<sup>th</sup></b>	- Mask obligation introduced in Austria
<b>March 31<sup>st</sup></b>	- German government speaks out against mask obligation
<b>April 2<sup>nd</sup></b>	- RKI recommends wearing a mask for everyone
<b>April 6<sup>th</sup></b>	- Mask obligation introduced in Jena
<b>April 15<sup>th</sup></b>	- Contact restriction policies extended until May 3 <sup>rd</sup>
<b>April 17<sup>th</sup></b>	- MoH on national TV: first wave of infection well overcome, Germany should gradually return to normality while staying alert
<b>April 19<sup>th</sup></b>	- Introduction compulsory wearing of masks in MV and SN
<b>April 22<sup>nd</sup></b>	- Announcement: Mask obligation in all federal states
<b>April 27<sup>th</sup></b>	- Mask obligation comes into force nationwide

**Figure 4**

Timeline of communication milestones from February 15th to April 27th 2020