Article

Not peer-reviewed version

Digital Infoveillance for Measles Surveillance in Italy: Analysis of Google Trends and Wikipedia Data (2013–2025)

Omar Enzo Santangelo * , Anna Sole Pizzamiglio , Carlotta Vella , Danilo Greco , Alberto Firenze , Sandro Provenzano

Posted Date: 19 March 2026

doi: 10.20944/preprints202603.1557.v1

Keywords: infodemiology; infoveillance; epidemiological surveillance; measles; google trends; Wikipedia; infectious diseases; Italy

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 March 2026 doi:10.20944/preprints202603.1557.v1

Article

Digital Infoveillance for Measles Surveillance in Italy: Analysis of Google Trends and Wikipedia Data (2013–2025)

Omar Enzo Santangelo [1,*], Anna Sole Pizzamiglio [1], Carlotta Vella [1], Danilo Greco [2], Alberto Firenze [3] and Sandro Provenzano [2]

[1] Regional Health Care and Social Agency of Lodi, ASST Lodi, 26900 Lodi, Italy
[2] Local Health Unit of Trapani, ASP Trapani, 91100 Trapani, Italy
[3] Local Health Unit of Palermo, ASP Palermo, 90129 Palermo, Italy
* Correspondence: omarenzosantangelo@hotmail.it

Abstract

The increasing digitalization of information behaviors has created new opportunities in public health, promoting the development of infodemiology and infoveillance as complements to traditional epidemiological surveillance. This study aimed to assess the correlation and temporal association between official measles surveillance data published by the Italian National Institute of Health (ISS), Google Trends search volumes, and Wikipedia page views related to measles in Italy. A cross-sectional analysis was conducted using measles cases reported in Italy from January 2013 to November 2025. Monthly data from Google Trends and Wikipedia were aggregated and compared with the monthly incidence reported in ISS bulletins. Cross-correlation analyses and linear regressions were performed to evaluate associations. The results showed a statistically significant and strong correlation (r=0.81) between Google Trends search volume for "Morbillo" and reported measles cases. Linear regression analyses demonstrated significant associations between digital indicators and official surveillance data, suggesting that online user activity closely reflects measles epidemiological trends. These findings indicate that Google Trends and Wikipedia may serve as useful and timely tools to complement traditional measles surveillance. Although caution is needed due to methodological limitations and potential information bias, integrating digital data could enhance early outbreak detection and support more rapid public health responses.

Keywords: infodemiology; infoveillance; epidemiological surveillance; measles; google trends; Wikipedia; infectious diseases; Italy

1. Introduction

In recent years, public healthcare has undergone a profound transformation driven by the progressive digitalization of the population's information behaviour, with increasing use of online platforms for health-related information searches. Within this context, infodemiology and infoveillance have emerged as important tools for digital public health surveillance, two innovative approaches that apply the principles of epidemiology to the study of the production, dissemination, and use of health information in digital environments. Infodemiology, introduced by Gunther Eysenbach in the early 2000s, is defined as the science that analyses the distribution and determinants of information in digital contexts with the aim of supporting policy making and decision-making processes in public health [1,2]. It is based on the assumption that online information flows significantly influence individual and collective behaviour, with potential direct impacts on population health outcomes. Within this theoretical framework, infoveillance represents the most operational dimension of infodemiology, configuring itself as a continuous and real-time information surveillance system, based on the analysis of online searches, social media and digital contents [2]. Numerous studies have demonstrated how these data sources can anticipate the progress of health events compared to traditional surveillance systems, offering timely indicators relating to infectious diseases, health behaviours and risk perception in the general population [3–5]. The COVID-19 pandemic has further highlighted the crucial role of these tools in monitoring and combating the "infodemic", i.e., the overabundance of information, often inaccurate or misleading, which can compromise the effectiveness of institutional healthcare responses [6]. At the same time, the evolution of digital platforms and the availability of large volumes of user-generated data from search engines like Google and social networks have expanded the application possibilities of digital epidemiological surveillance. Tools such as Google Trends allow the analysis of relative search volume (RSV) of specific terms or topics, providing indirect proxies for the population's attention, interest, and concerns [7,8]. Several studies have highlighted a statistically significant correlation between online search trends and data from official surveillance systems for numerous infectious diseases [9–11], especially in contexts characterised by delays or incompleteness of traditional reporting systems. From this perspective, measles represents a particularly relevant case study. Despite the availability of a safe and effective vaccine, measles remains one of the most contagious infectious diseases and a significant cause of childhood morbidity and mortality globally, especially when vaccination coverage is insufficient [10,12]. In Europe and Italy, where the disease is subject to mandatory notification, epidemic outbreaks persist, particularly associated with poor adherence to vaccinations [12–14]. Measles continues to represent a substantial public health threat in Italy, particularly in regions with suboptimal vaccination coverage. Notably, recent outbreaks, including those observed in 2017–2018, highlight the critical need for robust and timely surveillance systems to mitigate the spread of the disease [15]. Traditional measles surveillance, based on clinical and laboratory reports, is often burdensome and characterized by a time delay in data availability; for this reason, the integration with digital data from the web and social media appears to be a promising strategy to improve early detection and monitoring of the disease progression [9–11,16,17]. Overall, infodemiology and infoveillance emerge as key tools for digital public health, capable of complementing traditional epidemiological surveillance systems, improving institutional communication, and strengthening the capacity to respond to health emergencies. However, their use requires careful consideration of methodological limitations, algorithmic biases, data representativeness, and ethical implications related to privacy protection and the responsible use of information [18]. In this scenario, harmonizing analysis methods and multidisciplinary integration are fundamental steps to fully exploit the potential of these innovative public health tools. This study was conducted with the primary objective of evaluating the correlation and temporal association between data from Google Trends (GT), Wikipedia searches, and conventional surveillance data relating to measles infection, reported in the bulletins of the Italian National Institute of Health (ISS). Specifically, the study aims to explore the applicability of digital surveillance tools as a support to traditional systems for monitoring communicable diseases, with specific reference to measles

2. Materials and Methods

This study employs a time-series epidemiological approach, analyzing monthly data from 2013 to 2025 to examine trends in measles cases alongside digital indicators. Measles cases were selected between January 2013 and November 2025. The ISS publishes a monthly bulletin containing the number of measles cases reported as possible, probable, and confirmed in the previous months [15]. Internet search data were obtained from GT [11]. Data were extracted on December 31, 2025 using the Italian terms in the "Health" category: "Morbillo" as topic (Measles in English) and "Sintomi Morbillo" as search term (Measles Symptoms in English), in the time period between January 1, 2013 (since data became available) and November 30, 2025; data were aggregated on a monthly basis. The file was downloaded in ".CSV" format. Google Trends provides a relative search volume (RSV), calculated as the percentage of queries related to a given term for a specific location and time period, normalized on a scale from 0 to 100. From Wikipedia Trends [19] it is possible to obtain information

 doi:10.20944/preprints202603.1557.v1

on the number of views of a specific page by users. The data were extracted as daily data and then aggregated on a monthly basis. The following pages were therefore considered: “Morbillo” (Measles in English), “Vaccino del Morbillo” (Measles vaccine in English), “Vaccinazione MPR” (MMR Vaccination in English), “Esantema” (Exanthema in English) and “Macchie di Koplik” (Koplik spots in English), for the period between 1 July 2015 (since data became available) and 30 November 2025. Data from Wikipedia and GT have been temporally aligned with the monthly incidence of cases reported in the ISS epidemiological bulletins. Then a cross-correlation analysis was done. In order to allow direct comparison with the Wikipedia data and with the corresponding subset of the ISS data, and therefore make the time intervals superimposable, a second dataset was created containing the GT data relating to the keywords “Morbillo” and “Sintomi Morbillo” for the period 1 July 2015 – 30 November 2025, conventionally called “GT15”. The cross-correlation results were obtained as product-moment correlations between the time series. The main advantage of this approach lies in the ability to consider the temporal dependence between the analysed variables. Statistical analyses were performed using the Pearson correlation coefficient (r). Correlation coefficients were interpreted based on commonly accepted thresholds (strong if >0.7, moderate if between 0.3 and 0.7, and weak if <0.3 [20]. Linear regressions were also performed, considering as dependent variables the Wikipedia searches for the page “Morbillo”, the Google Trends RSV values for the term “Morbillo” and the ISS cases; as independent variables, the GT search terms, the Wikipedia pages analyzed and, where applicable, the ISS cases were used. The results are expressed as regression coefficients with 95% confidence intervals (95% CI). Potential autocorrelation was ascertained through the calculation of the Durbin-Watson (DW) statistics. The DW test is a statistic test used to detect the presence of autocorrelation in the residuals (prediction errors) from a regression analysis [21]. The DW test statistic or d always lies between 0 and 4. If the d is substantially less than 2, there is evidence of positive serial correlation, while values greater than 2 suggest no autocorrelation. Linear and exponential models were also estimated for the graphical representation of the associations, calculating the coefficient of determination (R^2), the data are reported in supplementary materials, Figure S1. The statistical significance level was set at 0.05. Analyses were conducted using STATA statistical software, version 14.

3. Results

The findings demonstrate a strong association between ISS-reported measles cases and online search activity, particularly as captured through Google Trends (Figure 1a–c, Table 1). Broad search terms, such as “morbillo” and “sintomi morbillo”, exhibited the highest correlations with ISS cases (r = 0.81–0.89, p<0.001; Table 1), indicating that public online interest closely mirrors epidemiological trends. In contrast, Wikipedia pages addressing more specific symptoms, such as “Esantema Wiki” and “Macchie di Koplik Wiki,” showed moderate correlations with ISS cases (r = 0.44–0.74, p<0.001; Table 1). This likely reflects lower search volume and more targeted interest in these topics, which are primarily consulted during epidemic peaks or periods of heightened media coverage. Linear regression analyses further support the predictive potential of online search data for measles incidence. For instance, Google Trends queries for “Morbillo GT15” and “Sintomi Morbillo GT15” were strongly associated with ISS cases (coef. 9.34 and coef = 10.03, respectively; p<0.001; Table 2), while Wikipedia page views also showed significant, albeit smaller, effects (coef. 0.012–0.22; p<0.001; Table 2, Figure S1). Durbin–Watson statistics suggest some residual autocorrelation, indicating that models accounting for temporal structure may enhance predictive accuracy. Overall, these results indicate that digital search behavior can serve as a complementary indicator of measles activity, while highlighting the limitations of interpreting searches for more specific and less frequently consulted topics.

 doi:10.20944/preprints202603.1557.v1

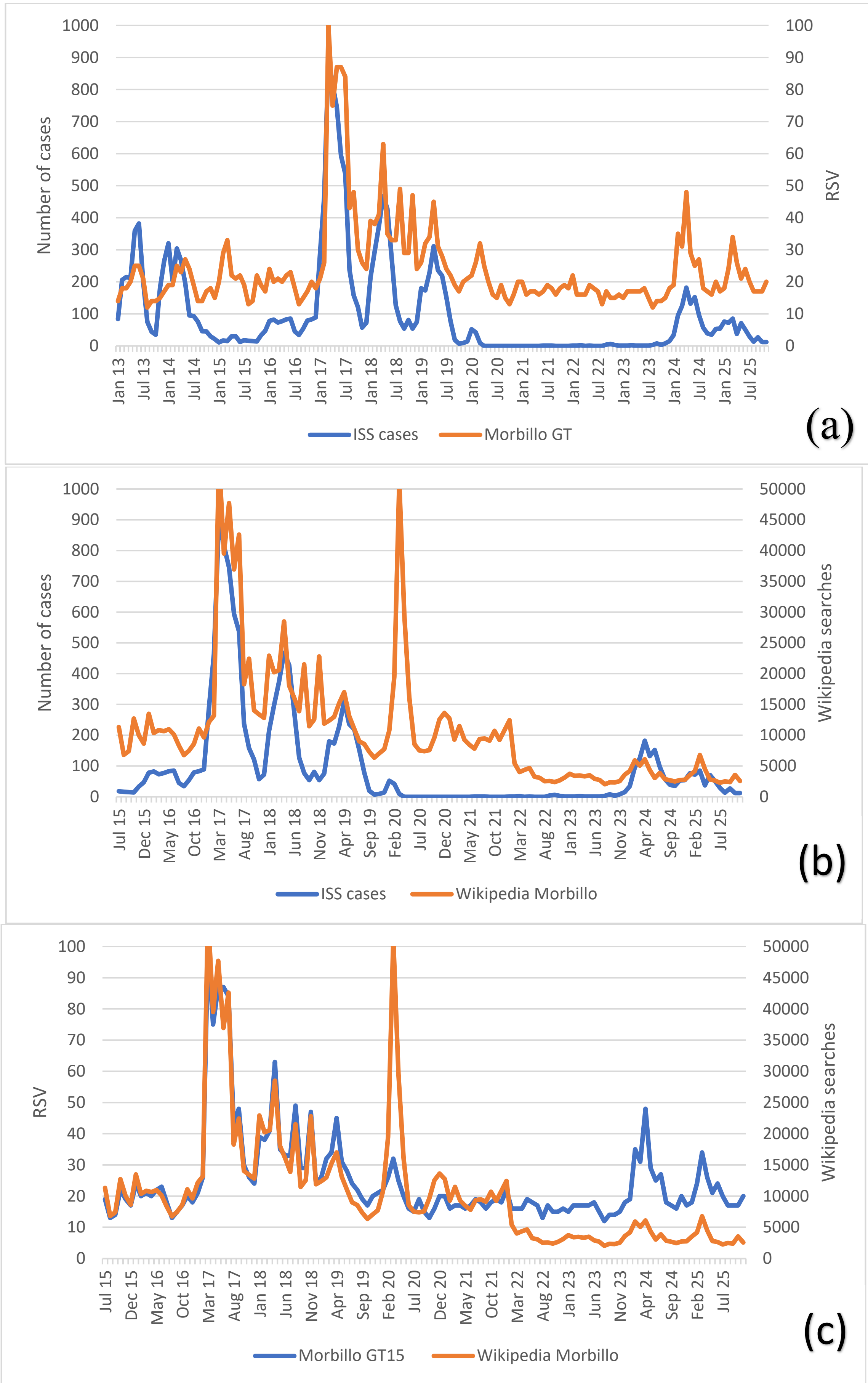

Figure 1. Raw data for the number of ISS and RSV measles cases from GT (a). Raw data for the number of ISS cases and the number of searches for the measles Wikipedia page (b). Raw RSV measles data from GT15 and the number of searches for the measles Wikipedia page (c).

Table 1. Correlations between GT terms, Wikipedia pages, and ISS cases. Strong correlations are shown in green, moderate correlations are shown in yellow.

		ISS cases	Morbillo GT	Morbillo GT15	Sintomi Morbillo GT	Sintomi Morbillo GT15	Morbillo Wikipedia	Vaccino del Morbillo Wiki	Vaccino MPR Wiki	Esantema Wiki	Macchie di Koplik Wiki
ISS cases	Correlation	1									
	p-value	-									
	Observations	155									
Morbillo GT	Correlation	0.8105	1								
	p-value	<0.001	-								
	Observations	155	155								
Morbillo GT15	Correlation	0.8879	1	1							
	p-value	<0.001	<0.001	-							
	Observations	125	125	125							
Sintomi Morbillo GT	Correlation	0.8247	0.8532	0.8738	1						
	p-value	<0.001	<0.001	<0.001	-						
	Observations	155	155	125	155						
Sintomi Morbillo GT15	Correlation	0.8856	0.8738	0.8738	1	1					
	p-value	<0.001	<0.001	<0.001	<0.001	-					
	Observations	125	125	125	125	125					
Morbillo Wikipedia	Correlation	0.7402	0.8312	0.8312	0.6634	0.6634	1				
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001	-				
	Observations	125	125	125	125	125	125				
Vaccino del Morbillo Wiki	Correlation	0.8131	0.8870	0.8870	0.7185	0.7185	0.9001	1			
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-			
	Observations	125	125	125	125	125	125	125			
Vaccino MPR Wiki	Correlation	0.7370	0.7920	0.7920	0.6331	0.6331	0.8375	0.9271	1		
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-		
	Observations	125	125	125	125	125	125	125	125		
Esantema Wiki	Correlation	0.4369	0.3965	0.3965	0.2870	0.2870	0.5974	0.5299	0.6108	1	
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-	
	Observations	125	125	125	125	125	125	125	125	125	
Macchie di Koplik Wiki	Correlation	0.7405	0.6834	0.6834	0.5807	0.5807	0.7547	0.7895	0.7770	0.8318	1
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-
	Observations	125	125	125	125	125	125	125	125	125	125

The moderate correlations observed between ISS cases and certain Wikipedia pages (e.g., "Esantema Wiki" and "Macchie di Koplik Wiki") may be attributed to the more specific nature of these search terms, which are less frequently consulted by the general population compared to broader topics such as "Morbillo" (Measles) or "Vaccino del Morbillo" (Measles Vaccine). This specificity results in relatively lower search volumes and less widespread public engagement. Furthermore, terms related to specific symptoms of measles, such as exanthem and Koplik spots, are likely to attract attention primarily during more apparent epidemic events or heightened media coverage. As such, online interest in these topics tends to be less consistent and more susceptible to seasonal variations or media-driven peaks, thereby influencing the strength of the correlation observed

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 March 2026 doi:10.20944/preprints202603.1557.v1

Table 2. Linear regression models.

	Independent variable: Morbillo GT			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
ISS cases	8.86	7.84-9.98	<0.001	0.56
	Independent variable: Morbillo GT15			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
ISS cases	9.34	8.48-10.21	<0.001	0.93
	Independent variable: Sintomi Morbillo GT			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	9.72	8.66-10.79	<0.001	0.67
	Independent variable: Sintomi Morbillo GT15			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	10.03	9.09-10.98	<0.001	0.92
	Independent variable: Morbillo Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	0.012	0.010-0.014	<0.001	0.43
	Independent variable: Vaccino del Morbillo Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	0.18	0.16-0.21	<0.001	0.44
	Independent variable: Vaccino MPR Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	0.04	0.03-0.05	<0.001	0.37
	Independent variable: Esantema Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	0.02	0.01-0.03	<0.001	0.21
	Independent variable: Macchie di Koplik Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
ISS cases	0.22	0.18-0.26	<0.001	0.25
	Independent variable: ISS cases			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT	0.07	0.06-0.08	<0.001	0.82
	Independent variable: Sintomi Morbillo GT			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT	0.92	0.83-1.01	<0.001	1.06
	Independent variable: ISS cases			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	0.08	0.07-0.09	<0.001	1.22
	Independent variable: Sintomi Morbillo GT15			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	7.93	0.94-1.03	<0.001	1.05
	Independent variable: Morbillo Wikipedia			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson

 doi:10.20944/preprints202603.1557.v1

Morbillo GT15	0.001	0.0011-0.0015	<0.001	0.51
	Independent variable: Vaccino del Morbillo Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	0.019	0.017-0.021	<0.001	0.52
	Independent variable: Vaccino MPR Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	0.004	0.003-0.004	<0.001	0.55
	Independent variable: Esantema Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	0.002	0.001-0.003	<0.001	0.51
	Independent variable: Macchie di Koplik Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo GT15	0.019	0.015-0.023	<0.001	0.66
	Independent variable: ISS cases			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	44.62	37.39-51.86	<0.001	0.73
	Independent variable: Morbillo GT15			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	527.67	464.67-590.67	<0.001	0.52
	Independent variable: Sintomi del Morbillo GT15			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	453.25	362.01-544.49	<0.001	0.61
	Independent variable: Vaccino del Morbillo Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	12.38	11.31-13.45	<0.001	0.84
	Independent variable: Vaccino MPR Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	2.92	2.58-3.26	<0.001	0.81
	Independent variable: Esantema Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	2.34	1.78-2.91	<0.001	0.64
	Independent variable: Macchie di Koplik Wiki			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin-Watson
Morbillo Wikipedia	13.78	11.64-15.92	<0.001	0.83

4. Discussion

The study of the application of GT and Wikipedia in the surveillance of communicable diseases in Italy has a high relevance for public health, as suggested by numerous previous studies [7,8]. Measles outbreaks remain a global health problem, and surveillance is an essential tool for their control. However, traditional monitoring approaches based on conventional reporting systems present some critical issues, including time lags in data availability and potential underestimation of cases due to underreporting.

In 2009, Eysenbach defined infodemiology and infoveillance as "the science of the distribution and determinants of information in an electronic medium, especially the Internet, or in a population,

 doi:10.20944/preprints202603.1557.v1

with the ultimate goal of informing public health and public policy.” [1]. This emerging approach has been widely used to explore public interest in communicable diseases, such as influenza. [3], and for the early detection of infectious epidemics [22]. Previous studies have shown that digital platforms, such as Google Trends, can effectively anticipate influenza activity and provide early signals of epidemic dynamics [3,23]. During the COVID-19 pandemic, infodemiology and infoveillance approaches successfully captured population concern, risk perception, and health-seeking behaviors, complementing conventional surveillance and informing timely public health responses [24]. These observations highlight the potential of online data to reflect both epidemic trends and public information-seeking behavior. Consistent with these findings, our study demonstrates that Google Trends and Wikipedia can capture public attention toward measles and correlate with official epidemiological data in Italy [25,26]. Searches for general terms, such as “morbillo” and “sintomi morbillo,” showed strong correlations with ISS-reported cases, while symptom-specific searches, such as “esantema” and “macchie di Koplik,” displayed moderate correlations. This pattern aligns with prior work suggesting that search term specificity and frequency influence the reliability of digital surveillance signals [18,25]. Integrating multiple digital sources may enhance infoveillance by capturing complementary dimensions of public interest. While Google Trends provides quantitative measures of search activity, Wikipedia page views may reflect deeper engagement with health information, offering additional insights into public awareness and concern. This multi-source approach is particularly useful where traditional surveillance is delayed or incomplete, providing near real-time proxies to guide timely public health actions [18,24]. Overall, our results underscore the broad applicability of infodemiology and infoveillance for monitoring communicable diseases. Beyond rapidly spreading infections like influenza and COVID-19, digital surveillance can complement traditional systems for vaccine-preventable diseases such as measles, particularly in regions with suboptimal vaccination coverage and recurrent outbreaks. By integrating digital and conventional data, early warning capabilities, public health communication, and disease control strategies can be strengthened [23–26].

Measles is one of the most contagious vaccine-preventable diseases (R0 = 12–18). The inherent delays in traditional surveillance systems can limit the ability of health authorities to intervene promptly. Data from Google Trends and Wikipedia are collected and processed in near real time, allowing for significantly faster information availability than traditional systems [9]. In this study, statistical analyses highlighted a clear temporal correlation between ISS surveillance data and search trends on Google and Wikipedia. In particular, a strong correlation was observed on a monthly basis, suggesting that this approach could provide a useful time frame for more timely public health interventions. However, it cannot be ruled out that a weekly or daily analysis could allow for even earlier case reporting, as reported in other studies [26,27]. Although the Durbin-Watson test suggests some potential autocorrelation in the residuals of the regression models (DW values <2), the impact on the model’s validity is limited, as the correlations between digital indicators and measles cases remain statistically significant.

This study has some limitations. First, the influence of mass media (television, radio, and print) may lead to an increase in online searches regardless of actual case trends [28]. Furthermore, Google Trends does not provide sufficiently granular data to accurately identify the geographic location of outbreaks. Further limitations arise from changes over time in the GT interface and algorithms, which are not fully documented and could influence the results. Therefore, the interpretation and generalization of the results should be made with caution.

Previous experiences, such as Google Flu Trends, have shown how the use of big data for epidemiological forecasting can lead to significant overestimations, as occurred in the 2012–2013 influenza season [29]. In the case of measles, online search volume is likely influenced more by media attention than by direct patient experience, given the relative rarity of cases. This makes digital surveillance potentially more sensitive, but also more susceptible to bias. A prudent and integrated use of these tools, alongside traditional surveillance systems, is therefore necessary.

 doi:10.20944/preprints202603.1557.v1

5. Conclusions

In summary, the Italian healthcare system operates in a context characterized by limited resources, which necessitates the adoption of innovative strategies to improve efficiency and sustainability. Investing in technological innovation represents a strategic lever to strengthen the National Health Service and improve its capacity to respond to the population's needs. The results of this study suggest that surveillance systems based on Google Trends and Wikipedia could play a complementary role in public health by providing dynamic, near-real-time indicators of the spread of infectious diseases. The integration of these tools with traditional surveillance systems could support policymakers in implementing preventive measures and targeted communication strategies [30,31]. Although this field of research is still relatively young, its rapid expansion is expected, facilitated by the increasing availability of data and increased computational power [32]. Online searches, conducted in real time, could serve as an early warning system, signalling sudden changes in the population's interests and therefore deserving attention from health authorities [33]. Looking ahead, this approach could be used as a supplementary tool for epidemic forecasting and for a more efficient allocation of health resources [34–36].

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Exponential and linear regression models between Wikipedia searches, GT's RSV, and ISS cases. The dashed line shows the model. The equation and corresponding R^2 are shown in the figures.

Author Contributions: Conceptualization, O.E.S.; methodology, O.E.S.; software, O.E.S.; validation, O.E.S.; formal analysis, O.E.S.; investigation, O.E.S.; resources, O.E.S.; data curation, O.E.S.; writing—original draft preparation, O.E.S., A.S.P., C.V., D.G., S.P, A.F.; writing—review and editing, O.E.S., A.S.P., C.V., D.G., S.P, A.F.; visualization, O.E.S.; supervision, O.E.S.; project administration, O.E.S.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: the data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: None.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Eysenbach G. Infodemiology: The epidemiology of (mis)information. Am J Med. 2002;113(9):763-765. doi:10.1016/s0002-9343(02)01473-0
2. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res. 2009;11(1):e11. Published 2009 Mar 27. doi:10.2196/jmir.1157
3. Mavragani A. Infodemiology and Infoveillance: Scoping Review. J Med Internet Res. 2020;22(4):e16206. Published 2020 Apr 28. doi:10.2196/16206
4. Strzelecki A, Meinzenbach A, Zieger M. Infodemic and infodemiology in public health: Similarities and differences. Healthcare Analytics. 2023; 4:100243. doi: 10.1016/j.health.2023.100243
5. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. Health Info Libr J. 2018;35(2):91-120. doi:10.1111/hir.12216
6. World Health Organization. 2020, 1st WHO Infodemiology Conference – Call for Papers and Recommendations.

 doi:10.20944/preprints202603.1557.v1

7. Milinovich GJ, Williams GM, Clements AC, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. Lancet Infect Dis. 2014;14(2):160-168. doi:10.1016/S1473-3099(13)70244-5.
8. Pollett S, Boscardin WJ, Azziz-Baumgartner E, et al. Evaluating Google Flu Trends in Latin America: Important Lessons for the Next Phase of Digital Disease Detection. Clin Infect Dis. 2017;64(1):34-41. doi:10.1093/cid/ciw657
9. Teng Y, Bi D, Xie G, et al. Dynamic Forecasting of Zika Epidemics Using Google Trends. PLoS One. 2017;12(1):e0165085. Published 2017 Jan 6. doi:10.1371/journal.pone.0165085
10. Giordano D, Provenzano S, Santangelo OE, et al. Active immunization status against measles, mumps, rubella, hepatitis B in internationally adopted children, surveyed at the university hospital of Palermo, Sicily. Ann Ig. 2018;30(5):431-435. doi:10.7416/ai.2018.2243.
11. Google Trends. Available online: https://trends.google.com/trends/ [last access December 31, 2025]
12. Moss WJ. Measles. Lancet. 2017;390(10111):2490-2502. doi:10.1016/S0140-6736(17)31463-0
13. Coughlan S, Connell J, Cohen B, Jin L, Hall WW. Suboptimal measles-mumps-rubella vaccination coverage facilitates an imported measles outbreak in Ireland. Clin Infect Dis. 2002;35(1):84-86. doi:10.1086/340708
14. Probert WS, Glenn-Finer R, Espinosa A, et al. Molecular Epidemiology of Measles in California, United States-2019. J Infect Dis. 2021;224(6):1015-1023. doi:10.1093/infdis/jiab059
15. Istituto Superiore di Sanità. Available online: https://www.epicentro.iss.it/morbillo/bollettino [last access, January 10, 2026]
16. Hilbert M, López P. The world’s technological capacity to store, communicate, and compute information. Science. 2011;332(6025):60-65. doi:10.1126/science.1200970.
17. Alicino C, Bragazzi NL, Faccio V, et al. Assessing Ebola-related web search behaviour: insights and implications from an analytical study of Google Trends-based query volumes. Infect Dis Poverty. 2015;4:54. Published 2015 Dec 10. doi:10.1186/s40249-015-0090-9
18. Velasco E, Agheneza T, Denecke K, Kirchner G, Eckmanns T. Social media and Internet-based data in global systems for public health surveillance: a systematic review. Milbank Q. 2014;92(1):7–33.
19. Istituto Superiore di Sanità. Available online: https://www.epicentro.iss.it/morbillo/bollettino [last access, January 10, 2026]
20. Wikipedia. Available online: https://it.wikipedia.org/wiki/Wikipedia [last access, December 31, 2025]
21. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J. 2012;24(3):69-71.
22. White KJ. The Durbin-Watson Test for Autocorrelation in Nonlinear Models 370 The Review of Economics and Statistics the Durbin-Watson Test for Autocorrelation in Nonlinear Models. Rev Econ Stat. 1992; 74:370–373
23. Lu FS, Hou S, Baltrusaitis K, et al. Accurate Influenza Monitoring and Forecasting Using Novel Internet Data Streams: A Case Study in the Boston Metropolis. JMIR Public Health Surveill. 2018;4(1):e4. Published 2018 Jan 9. doi:10.2196/publichealth.8950
24. Nuti SV, Wayda B, Ranasinghe I, et al. The use of google trends in health care research: a systematic review. PLoS One. 2014;9(10):e109583. Published 2014 Oct 22. doi:10.1371/journal.pone.0109583
25. Siani A. Measles outbreaks in Italy: A paradigm of the re-emergence of vaccine-preventable diseases in developed countries. Prev Med. 2019;121:99-104. doi:10.1016/j.ypmed.2019.02.011
26. Santangelo OE, Provenzano S, Piazza D, Giordano D, Calamusa G, Firenze A. Digital epidemiology: assessment of measles infection through Google Trends mechanism in Italy. Ann Ig. 2019;31(4):385-391. doi:10.7416/ai.2019.2300
27. Gianfredi V, Bragazzi NL, Mahamid M, et al. Monitoring public interest toward pertussis outbreaks: an extensive Google Trends-based analysis. Public Health. 2018;165:9-15. doi:10.1016/j.puhe.2018.09.001
28. Provenzano S, Santangelo OE, Giordano D, et al. Predicting disease outbreaks: evaluating measles infection with Wikipedia Trends. Recenti Prog Med. 2019;110(6):292-296. doi:10.1701/3182.31610
29. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science. 2014;343(6176):1203-1205. doi:10.1126/science.1248506
30. Odone A, Signorelli C. When vaccine hesitancy makes headlines. Vaccine. 2017;35(9):1209-1210. doi:10.1016/j.vaccine.2015.11.051

doi:10.20944/preprints202603.1557.v1

31. Niu Q, Liu J, Zhao Z, et al. Explanation of hand, foot, and mouth disease cases in Japan using Google Trends before and during the COVID-19: infodemiology study. BMC Infect Dis. 2022;22(1):806. Published 2022 Oct 29. doi:10.1186/s12879-022-07790-9
32. Provenzano S, Santangelo OE, Gianfredi V. Infodemiology and infoveillance: framework for contagious exanthematous diseases, of childhood in Italy. Pathog Glob Health. 2024;118(4):317-324. doi: 10.1080/20477724.2024.2323844
33. Santangelo OE, Gentile V, Pizzo S, Giordano D, Cedrone F. Machine Learning and Prediction of Infectious Diseases: A Systematic Review. Machine Learning and Knowledge Extraction. 2023; 5(1):175-198. https://doi.org/10.3390/make5010013
34. Lan R, Campana F, Tardivo D, Catherine JH, Vergnes JN, Hadj-Saïd M. Relationship between internet research data of oral neoplasms and public health programs in the European Union. BMC Oral Health. 2021;21(1):648. Published 2021 Dec 17. doi: 10.1186/s12903-021-02022-z
35. Steinert Y, Naismith L, Mann K. Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review: BEME Guide No. 19. Med Teach. 2012;34(6):483-503. doi:10.3109/0142159X.2012.680937
36. Santangelo OE, Gianfredi V, Provenzano S. Wikipedia searches and the epidemiology of infectious diseases: A systematic review. Data & Knowledge Engineering. 2022; 142: 102093. Published 2022 15 Oct. doi: 10.1016/j.datak.2022.10209