MEASLES RESURGENCE IN THE UNITED STATES OF AMERICA: AN INFODEMIOLOGICAL STUDY FROM 2022 TO 2025
NÁVRAT SPALNIČEK VE SPOJENÝCH STÁTECH AMERICKÝCH: INFODEMIOLOGICKÁ STUDIE Z LET 2022 AŽ 2025

SANTANGELO Omar Enzo*, PIZZAMIGLIO Anna Sole, VELLA Carlotta

Regional Health Care and Social Agency of Lodi, ASST Lodi, 26900 Lodi, Italy

** korešpondenčný autor / corresponding author: Omarenzosantangelo@hotmail.it*

ABSTRACT

Introduction: A resurgence of measles has been observed in the United States between 2022 and 2025, despite its declared elimination in 2000. Vaccine hesitancy and misinformation are key factors. Infodemiology allows us to analyze the trends of online health information as a potential support for traditional surveillance.

Material and Methods: A cross-sectional study of weekly measles cases reported by the CDC (2022 – 2025) and Google Trends search volumes ("Measles") and Wikipedia views ("Measles," "Measles vaccine"). Correlation, cross-correlation, and linear regression analyses were performed; significance $p<0.05$.

Results: Raw data on CDC measles cases, Google Trends searches for "Measles," and views of the Wikipedia pages "Measles" and "Measles Vaccine" show statistically significant correlations ($p < 0.001$). Strong correlations emerged between CDC cases and Google searches, moderate correlations between CDC cases and Wikipedia views, and strong correlations between Google searches and Wikipedia pages. Linear regression models generally fit the data better than exponential models. All analyzed linear regressions were statistically significant across variables.

Conclusions: A significant temporal association emerged between reported cases and digital indicators. Infodemiological tools can complement epidemiological surveillance, improving the early identification of outbreaks and supporting communication strategies and combating vaccine hesitancy.

Key words: Measles. Infodemiology. Google Trends. Wikipedia. Digital epidemiology. United States.

ABSTRAKT

Úvod: V letech 2022 až 2025 byl ve Spojených státech pozorován opětovný výskyt spalniček, a to i přes jejich deklarovanou eliminaci v roce 2000. Klíčovými faktory jsou váhání s očkováním a dezinformace. Infodemiologie nám umožňuje analyzovat trendy online zdravotních informací jako potenciální podporu pro tradiční sledování.

Materiál a metody: Průřezová studie týdenních případů spalniček hlášených CDC (2022 – 2025) a objemů vyhledávání v Google Trends („Spalničky") a zobrazení na Wikipedii („Spalničky", „Vakcína proti spalničkám"). Byly provedeny korelační, křížová korelační a lineární regresní analýzy; významnost $p<0,05$.

Výsledky: Nezpracovaná data o případech spalniček v CDC, vyhledávání výrazu „Spalničky" v Google Trends a zobrazení stránek Wikipedie „Spalničky" a „Vakcína proti spalničkám" vykazují statisticky významné korelace ($p < 0,001$). Mezi případy CDC a vyhledáváním na Googlu se objevily silné korelace, mezi případy CDC a zobrazením Wikipedie mírné korelace a silné korelace mezi vyhledáváním na Googlu a stránkami Wikipedie. Lineární regresní modely obecně lépe odpovídaly datům než exponenciální modely. Všechny analyzované lineární regrese byly statisticky významné napříč proměnnými.

Závěry: Mezi hlášenými případy a digitálními ukazateli se objevila významná časová souvislost. Infodemiologické nástroje mohou doplňovat epidemiologický dohled, zlepšovat včasnou identifikaci ohnisek nákazy, podporovat komunikační strategie a bojovat proti váhání s očkováním.

Klíčová slova: Spalničky. Infodemiologie. Trendy Google. Wikipedie. Digitální epidemiologie. Spojené státy.

INTRODUCTION

The United States of America (USA) is facing a public health challenge with the return of measles especially from 2022 to 2025 [1]. Measles was once eliminated in the US thanks to vaccination efforts that were successful in 2000 [1, 2]. However, in years there has been a worrying rise in measles cases. Factors like hesitation to get vaccinated, misinformation and gaps in vaccination coverage are contributing to this resurgence putting both population health at risk [2]. To understand these factors and their role in public health crises we need to not track the disease but also explore how the public perceives it and how information about it spreads. This is where the field of infodemiology comes in. Infodemiology, a term introduced by Eysenbach in 2002 [3] is the study of how health-related information's distributed and what determines its spread within a population. It is a field that uses digital tools to analyze how health information moves, including through social media, search engines and news outlets. Infodemiology looks at not the amount of information but also its quality, credibility and potential impact on public behaviors and health-related decisions. In the case of measles infodemiological methods can track searches helping researchers see how concerns about measles change

how misinformation spreads and where public education is needed. By analyzing search data, such as how often people search for "measles vaccine " or "measles" we can gain insights into what the public thinks and knows about measles [4]. The period from 2022 to 2025, when measles cases rose again in the US is a time to study how public concerns about measles and vaccination have shifted and how these concerns relate to actual outbreaks. Infodemiology is a tool for public health professionals providing real-time data that can help inform targeted interventions and communication strategies to control measles and reduce vaccine hesitancy [5, 6]. 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 Centers for Disease Control and Prevention (CDC) USA. 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

MATERIALS AND METHODS

A cross-sectional study was conducted. Measles cases were selected between January 2, 2022 and December 21, 2025. The CDC publishes a weekly bulletin containing the number of measles cases reported [1]. Internet search data were obtained from GT [7]. Data was extracted on December 31, 2025 using the search terms in the “Health” category: “Measles” of GT searches carried out in the USA, in the time period between January 2, 2022 and December 27, 2025; data were aggregated on weekly 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 [8] it is possible to obtain information on the number of views of a specific page by users. The data were extracted as daily data and then aggregated on weekly basis. The following pages were therefore considered: “Measles” and “Measles vaccine”, for the period between January 2, 2022 and December 27, 2025. Data from Wikipedia and GT have been temporally aligned with the weekly incidence of cases reported in the CDC epidemiological bulletins. Then a cross-correlation analysis was done. 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). According to a commonly adopted rule of thumb, a correlation is considered strong if > 0.7, moderate if between 0.3 and 0.7 and weak if < 0.3 [9]. Linear regression analyses were also conducted, with Wikipedia searches for the “Measles” page, Google Trends RSV values for the term “Measles,” and CDC case counts treated as dependent variables; independent variables included the Google Trends search terms, the analyzed Wikipedia pages, and, where applicable, CDC case counts. 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 [10]. 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 statistical significance level was set at 0.05. Analyses were conducted using STATA statistical software, version 14.

RESULTS

The raw data for the number of CDC cases, RSV for the term “Measles” and the number of views of the “Measles” page and “Measles Vaccine” on Wikipedia are reported respectively in Figure 1. Table 1 shows the correlations between CDC cases, GT searches for the terms “Measles”” and the number of views of the Wikipedia pages “Measles”, and “Measles Vaccine”. All the results reported in Table 1 are statistically significant with p-value <0.001. Strong correlations emerged between CDC cases and GT search for the terms "Measles", moderate correlations between CDC cases and views of the Wikipedia pages "Measles" and "Measles Vaccine". Strong correlations emerged between GT search and Wikipedia pages. Figure 2 shows, both graphically

and mathematically, the linear and exponential regression models of the scatterplots between the CDC cases, the GT search term and the analysed Wikipedia pages. In most cases, linear models showed a better fit to the data than exponential models. Finally, Table 2 reports the associations estimated through linear regression, using as dependent variables, alternatively, CDC measles cases, GT term “Measles” and views of the "Measles" Wikipedia page. All the regressions analysed were statistically significant.

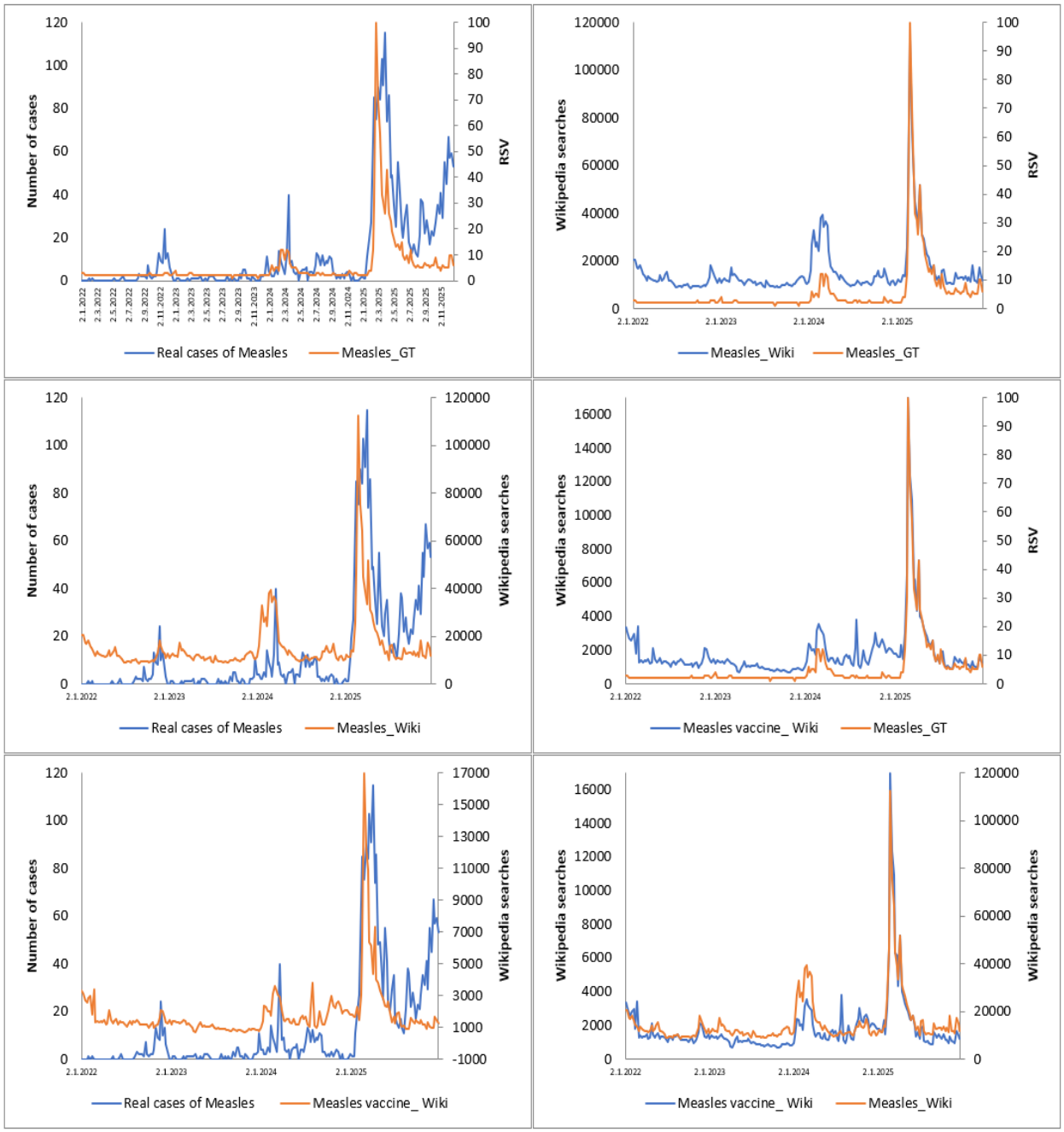

Figure 1 Comparison between raw data for the number of CDC cases, RSV measles from GT and the number of searches for the measles and measles vaccine Wikipedia pages

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

Parameter		CDC cases	Measles GT	Measles Wikipedia	Measles Vaccine Wikipedia
CDC cases	Correlation	1			
	p-value	-			
	observations	208			
Measles GT	Correlation	0.7338	1		
	p-value	<0.001	-		
	observations	208	208		
Measles Wikipedia	Correlation	0.6221	0.9277	1	
	p-value	<0.001	<0.001	-	
	observations	208	208	208	
Measles Vaccine Wikipedia	Correlation	0.6125	0.9378	0.9389	1
	p-value	<0.001	<0.001	<0.001	-
	observations	208	208	208	208

Table 2 Linear regression models

Parameter	Independent variable: Measles GT			
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
CDC case	1.51	1.32-1.71	<0.001	0.84
Independent variable: Measles Wikipedia				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
CDC case	0.001	0.0009-0.0014	<0.001	0.53
Independent variable: Measles Vaccine Wikipedia				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
CDC case	0.007	0.006-0.008	<0.001	0.54
Independent variable: CDC cases				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles GT	0.35	0.30-0.40	<0.001	1.04
Independent variable: Measles Wikipedia				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles GT	0.0008	0.0008-0.0009	<0.001	0.25
Independent variable: Measles Vaccine Wikipedia				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles GT	0.005	0.005-0.006	<0.001	0.44
Independent variable: CDC cases				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles Wikipedia	323.32	267.41-379.22	<0.001	0.71
Independent variable: Measles GT				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles Wikipedia	998.44	943.25-1053.62	<0.001	0.24
Independent variable: Measles Vaccine Wikipedia				
Dependent variable	Coef.	C.I. 95%	*p-value*	Durbin Watson
Measles Wikipedia	6.06	5.75-6.37	<0.001	0.45

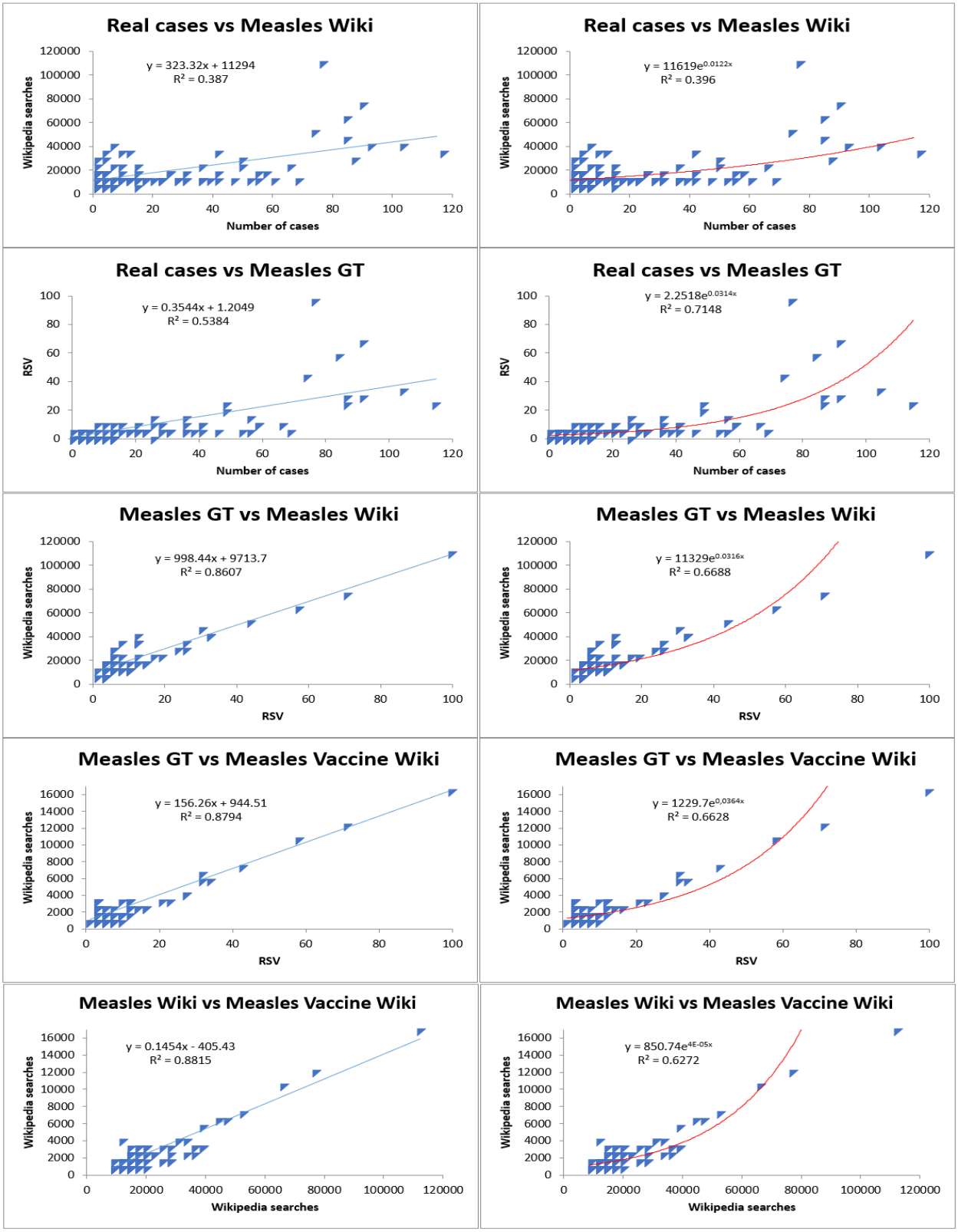

Figure 2 Exponential and linear regression models between Wikipedia searches, GT's RSV, and CDC cases. The lines shows the model. The equation and corresponding R^2 are shown in the figures.

DISCUSSION

The results confirm a significant correlation between CDC-reported measles cases and online search volumes from Google Trends and Wikipedia, supporting digital indicators as complementary tools to traditional surveillance. These findings align with the concept of infodemiology proposed by Eysenbach [3, 11], emphasizing the epidemiological value of online health information. Previous research has shown similar correlations, while noting methodological limitations such as reproducibility issues and the relative nature of search volumes [12, 13]. Wikipedia page views have also demonstrated strong associations with infectious disease trends and near-real-time predictive potential [14]. Our findings are consistent with this evidence, showing a temporal association between searches for "Measles" and "Measles vaccine" and CDC-reported cases. The resurgence of measles in the United States has been linked to declining vaccination coverage and vaccine hesitancy. The 2019 outbreak the largest since 1992 primarily affected communities with low vaccination rates [15]. In our analysis, search peaks often preceded or coincided with increases in reported cases, possibly reflecting both outbreak dynamics and media influence, consistent with evidence that digital platforms can amplify vaccine-related debates [16].

This study provides updated data for 2022 – 2025 and integrates Google Trends, Wikipedia, and official surveillance using cross-correlation and regression models with autocorrelation control, strengthening methodological rigor compared with descriptive analyses [12, 13]. The findings suggest that digital data may function as an early warning system, supporting timely interventions, vaccination campaigns, and risk communication strategies [1, 12]. Monitoring searches related to the "measles vaccine" may also offer indirect insight into vaccine confidence and early signs of hesitancy or misinformation [15, 16], improving evidence-based communication approaches [3, 11].

However, limitations must be considered. Google Trends data are expressed as Relative Search Volume (RSV) and do not represent absolute counts, limiting comparability across time and settings [4, 12, 13]. Online interest may also be influenced by media coverage, political events, algorithm changes, or internet penetration [5, 13, 15]. The observational design precludes causal inference, and increases in cases may drive search activity rather than the reverse. CDC data may be subject to reporting delays, and national-level analysis does not capture local heterogeneity, particularly in under-vaccinated communities [15]. The study also did not include qualitative analyses or other platforms such as social media, which are central to vaccine misinformation dynamics [16]. Despite these limitations, the findings strengthen the evidence supporting infodemiology as a complement to traditional surveillance and highlight the importance of rigorous, integrated use of digital tools in public health to improve outbreak detection, preparedness, and communication [6, 17].

CONCLUSIONS

Between 2022 and 2025, a significant temporal association was observed between CDC measles surveillance data and digital indicators from Google Trends and Wikipedia. Although measles was declared eliminated in the United States in 2000, recent years have shown renewed increases in cases, consistent with previous outbreaks [1, 12]. Search volumes for "Measles" and "Measles vaccine" demonstrated moderate to strong correlations with weekly CDC case incidence. Cross-correlation analysis suggests that peaks in digital interest may precede or accompany increases in reported cases, supporting a potential predictive role for infodemiological tools [3, 11]. These findings are consistent with the growing literature supporting the use of Google Trends and Wikipedia for digital disease surveillance [4–6]. While digital indicators reflect epidemiological dynamics, they must be interpreted with caution due to media influence, internet penetration variability, and the relative nature of RSV data [4–6]. Integrating traditional surveillance with infodemiological tools may improve outbreak detection, enhance communication strategies, and help monitor vaccine hesitancy – an important factor in the reemergence of measles [15, 16].

References

[1] CENTERS FOR DISEASE CONTROL AND PREVENTION. Measles cases and outbreaks. Available from: https://www.cdc.gov/measles/data-research/index.html (last access 10 January, 2026)

[2] GASTAÑADUY P.A., REDD S.B., FIEBELKORN A.P. et al. Division of Viral Disease, National Center for Immunization and Respiratory

Diseases, CDC. Measles - United States, January 1-May 23, 2014. *Morb Mortal Wkly Rep.* 2014; 63 (22): 496-499.

[3] EYSENBACH G. Infodemiology: The epidemiology of (mis)information. *Am J Med.* 2002; 113 (9): 763-765.

[4] SANTANGELO O.E., PROVENZANO S., PIAZZA D. et al. Digital epidemiology: assessment of measles infection through Google Trends mechanism in Italy. *Ann Ig.* 2019; 31 (4): 385-391.

[5] PROVENZANO S., SANTANGELO O.E., GIORDANO D. et al. Predicting disease outbreaks: evaluating measles infection with Wikipedia Trends. *Recenti Prog Med.* 2019; 110 (6): 292-296.

[6] SANTANGELO O.E., PROVENZANO S., GIANFREDI V. Infodemiology of flu: Google trends-based analysis of Italians' digital behavior and a focus on SARS-CoV-2, Italy. *J Prev Med Hyg.* 2021; 62 (3): E586-E591.

[7] GOOGLE TRENDS. Available online: https://trends.google.com/trends/ (last access 10 January, 2026)

[8] WIKIPEDIA. Available online: https://it.wikipedia.org/wiki/Wikipedia (last access 10 January, 2026)

[9] MUKAKA M.M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012; 24 (3): 69-71.

[10] WHITE K.J. 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.

[11] EYSENBACH G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med.* 2011; 40 (5 Suppl 2): S154-S158.

[12] NUTI S.V., WAYDA B., RANASINGHE I. et al. The use of google trends in health care research: a systematic review. *PLoS One.* 2014; 9 (10): e109583.

[13] MAVRAGANI A., OCHOA G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. *JMIR Public Health Surveill.* 2019; 5 (2): e13439.

[14] MCIVER D.J., BROWNSTEIN J.S. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS Comput Biol.* 2014; 10 (4): e1003581.

[15] PATEL M., LEE A.D., REDD S.B. et al. Increase in Measles Cases - United States, January 1-April 26, 2019. *MMWR Morb Mortal Wkly Rep.* 2019; 68 (17): 402-404.

[16] BRONIATOWSKI D.A., JAMISON A.M., QI S. et al. Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. *Am J Public Health.* 2018; 108 (10): 1378-1384.

[17] SANTANGELO O.E., PROVENZANO S., GRIGIS D. et al. Can Google Trends and Wikipedia help traditional surveillance? A pilot study on measles. *Acta Biomed.* 2020; 91 (4): e2020190.