

Google Trends Data to Determine Vaping Prevalence and Predictability in Europe: Usability Study

Keivan Armani, Richard Anthony Powell, Huda Y Almomani, Ireneous N Soyiri, Emily Knight

Submitted to: Journal of Medical Internet Research
on: October 06, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
---------------------------------	----------

Preprint
JMIR Publications

Google Trends Data to Determine Vaping Prevalence and Predictability in Europe: Usability Study

Keivan Armani^{1,2} PhD, MCLIN, BPHARM; Richard Anthony Powell¹ MSc, MA, BA; Huda Y Almomani³ PhD, PharmD; Ireneous N Soyiri⁴ PhD, MPhil, BSc; Emily Knight¹

¹Department of Primary Care and Public Health Faculty of Medicine Imperial College London London GB

²Faculty of Pharmaceutical Sciences UCSI University Kuala Lumpur MY

³Department of Applied Pharmaceutical Sciences and Clinical Pharmacy Faculty of Pharmacy Isra University Amman JO

⁴Hull York Medical School University of Hull Hull GB

Corresponding Author:

Keivan Armani PhD, MCLIN, BPHARM

Faculty of Pharmaceutical Sciences
UCSI University
Jalan Menara Gading
Kuala Lumpur
MY

Abstract

Background: The increasing use, and negative health effects, of e-cigarettes across Europe warrants continuous surveillance. Google Trends (GT) is a promising surveillance tool, providing real-time data that contrasts positively with cross-sectional surveys. However, its applicability to monitor e-cigarette use across Europe is unknown.

Objective: To monitor changes in online interest in e-cigarettes from 2012 to 2020, to investigate the relationship between GT searches on e-cigarettes and prevalence data and to assess the predictive capabilities of GT in Europe.

Methods: GT data for "electronic cigarette" and "vaporizer" were extracted for 27 European countries, whilst e-cigarette prevalence data were sourced from Eurobarometer. Regression and change-point analysis identified temporal changes in online interest and cross-correlation investigated the relationship between the two datasets. Countries displaying the strongest correlation underwent time series forecasting analysis using simple exponential smoothing (SES), and autoregressive integrated moving average (ARIMA) models.

Results: Significant correlations between GT and e-cigarette prevalence emerged for ten and seven countries for the "electronic cigarette" and "vaporizer" topics, respectively, while ARIMA outperformed SES in all but one country. Boosted ARIMA improved forecasts further in the evaluated countries.

Conclusions: This study affirms GT's potential to supplant cross-sectional e-cigarette data in certain European countries and highlights its predictive ability. However, future research should determine factors driving forecasting performance differences between countries and models. This work has potential to inform preventative policies using a flexible and continuous monitoring tool.

(JMIR Preprints 06/10/2025:85355)

DOI: <https://doi.org/10.2196/preprints.85355>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [JMIR Publications](#), my manuscript will be made available to all.

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in [JMIR Publications](#), my manuscript will be made available to all.



Original Manuscript

Google Trends Data to Determine Vaping Prevalence and Predictability in Europe: Usability Study

ABSTRACT

Background: The increasing use, and negative health effects, of e-cigarettes across Europe warrants continuous surveillance. Google Trends (GT) is a promising surveillance tool, providing real-time data that contrasts positively with cross-sectional surveys. However, its applicability to monitor e-cigarette use across Europe is unknown.

Objectives: To monitor changes in online interest in e-cigarettes from 2012 to 2020, to investigate the relationship between GT searches on e-cigarettes and prevalence data and to assess the predictive capabilities of GT in Europe.

Methods: GT data for "electronic cigarette" and "vaporizer" were extracted for 27 European countries, whilst e-cigarette prevalence data were sourced from Eurobarometer. Regression and change-point analysis identified temporal changes in online interest and cross-correlation investigated the relationship between the two datasets. Countries displaying the strongest correlation underwent time series forecasting analysis using simple exponential smoothing (SES), and autoregressive integrated moving average (ARIMA) models.

Results: Significant correlations between GT and e-cigarette prevalence emerged for ten and seven countries for the "electronic cigarette" and "vaporizer" topics, respectively, while ARIMA outperformed SES in all but one country. Boosted ARIMA improved forecasts further in the evaluated countries.

Conclusions: This study affirms GT's potential to supplant cross-sectional e-cigarette data in certain European countries and highlights its predictive ability. However, future research should determine factors driving forecasting performance differences between countries and models. This work has potential to inform preventative policies using a flexible and continuous monitoring tool.

Key words: Google Trends; Vaping; Electronic cigarette; Prevalence; Forecasting

Introduction

Tobacco smoking is a critical global public health challenge, constituting a leading cause of preventable disease and mortality worldwide, particularly in non-communicable and preventable diseases (e.g., cardiovascular ailments, respiratory conditions, stroke, and lung cancer).^{1,2} Over the past two decades, studies from England and Wales have shown a rising respiratory disease

burden, with overall hospitalisations nearly doubling,³ marked increases in acute upper respiratory infections,⁴ and steady growth in asthma and chronic obstructive pulmonary disease admissions and prescribing.⁵ Smoking is a well-established risk factor for these conditions, contributing substantially to their incidence and progression.^{6–8} In Europe, smoking trends vary significantly by region and country, with the highest rates in Eastern and South-eastern Europe (e.g., Bulgaria, Turkey), and the lowest rates in Northern and Western Europe (e.g., Sweden, Netherlands).⁹

Evidence suggests battery-operated electronic cigarettes (e-cigarettes)—also known as vapes or vaping devices that enable vapour inhalation via the heating of an electronic liquid (e-liquid)^{1,10}—reduces entailed harm and can potentially aid smoking cessation.^{11,12} E-cigarette use is increasingly prevalent among young people.^{13–15} There is, however, an ongoing debate regarding the relative benefits versus problems associated with e-cigarette use.^{16,17} E-cigarettes result in less deleterious health effects compared to traditional tobacco cigarettes given they do not combust tobacco or tar, which produce carcinogens and other toxins.^{17,18}

However, e-cigarettes are not harm-free; they can produce aerosols containing nicotine, carbonyl compounds and heavy metal compounds.^{10,18} Indeed, although research on the long-term health effects of e-cigarettes is limited given the relatively short time the products have been available on the market, nicotine is a key component of the e-liquid they use, for which evidence exists of its adverse health effects.^{19–22} Moreover, e-cigarettes are associated with an increased intention to smoke tobacco cigarettes among adolescents,^{23,24}—although the hypothesis of e-cigarettes being a gateway to tobacco smoking is yet to be definitively proven^{24,25}—and evidence indicates that, for established tobacco smokers, nicotine e-cigarettes produce higher cessation rates than nicotine replacement therapy and non-nicotine e-cigarettes for at least 6 months.^{26,27}

Within the European Union (EU), the Tobacco Products Directive (TPD) regulations impose minimum restrictions on all EU countries, which have also been adopted by the United Kingdom.²⁸ These include a maximum nicotine strength of 20 mg/mL, a refill bottle size of 10 mL, and mandatory warning labels.²⁸ Within the TPD regulations, individual states are permitted to impose additional restrictions, with many officially recommending e-cigarettes as a tobacco smoking cessation strategy.²⁹ Against this legislative backdrop, it is imperative to monitor the prevalence and associated trends of e-cigarette use both at the EU and country levels accurately to inform policies on the optimal balance of e-cigarettes' benefits and drawbacks.

Cross-sectional surveys are the primary methodology for measuring e-cigarette prevalence. The approach provides the most cost-effective method for large populations when conducting follow-up research is challenging. However, the precise measurement of e-cigarette prevalence can be problematic methodologically, with challenges including self-report bias and non-disclosure.^{30,31} Moreover, cross-sectional surveys only capture one time point and can take a long time to collect data, especially at the national level, providing policy makers with historical information.³²

Google Trends (GT) is an established tool for monitoring interests and potential behaviour in healthcare research.^{33,34} GT is a publicly accessible website displaying data on the popularity of

Google search terms over time. This enables researchers to analyse online interest in multiple topics.^{33,35} It is updated in near real-time, with each query filterable by location, language, time (from 2004 onwards), category and/or platform (i.e., web, Images, news, YouTube or shopping), providing a flexible and accessible tool capable of capturing public interest in a subject.³⁴

Given the ubiquity of the internet in Europe—an estimated 94% of EU households had access in 2024³⁶—and Google being the dominant internet search engine, GT data can complement traditional data sources.²³ Additionally, the majority of European internet users are younger populations,³⁷ the same demographic group among whom vaping is highly prevalent.²⁴ Examples of healthcare research using GT include tracking public perceptions on COVID-19 prevention measures, abortions, infectious diseases, and surgeries.^{35,38,39} Research on smoking has also been explored using GT, especially on how search volumes are affected by regulatory changes. For example, one European study reported increased smoking cessation searches following the reintroduction of the smoking cessation support reimbursement by health insurance,⁴⁰ whilst another observed significant search increases following the 'Stoptober' programme, an intervention encouraging temporary smoking abstinence.⁴¹ Consequently, some evidence exists that GT data on smoking reflects real-world public health initiatives, and is also a reliable alternative tool to assess e-cigarette prevalence.^{42,43}

GT has become an important tool within infodemiology, the science of analysing information in electronic media to improve public health.⁴⁴ Within this field, the main advantages of GT concern its use as a monitoring and forecasting tool, and as a proxy for true prevalence data to better inform policy makers of measures to mitigate public health concerns, including the rising prevalence of e-cigarettes. Indeed, the anonymised nature of GT may reflect the actual behaviour of young, potentially underage, populations better than formal surveys,⁴⁵ which can be prone to social desirability bias.^{46,47}

From a public health perspective, future forecasting is critical to implementing pro-active, preventative policies. One of the first applications of GT forecasting was predicting influenza cases 1–2 weeks ahead of official Centers for Disease Control and Prevention data.⁴⁸ Subsequent studies have successfully harnessed GT's disease forecasting potential for Zika virus, HIV, acute febrile illnesses and, more recently, COVID-19.^{49–52} GT has also been used to predict potential behaviours; for example, a 2017 study of suicide rates based on GT search terms,⁵³ and a 2019 study, successfully forecast suicide occurrences in Ireland using autoregression and neural network approaches.⁵⁴ These findings highlight GT's value in anticipating future events and providing more granular insights compared to cross-sectional data.

Consequently, this study aimed to explore GT's applicability in investigating e-cigarette use in Europe. Specifically, it aimed to monitor changes in online interest in e-cigarettes from 2012 to 2020; investigate the relationship between GT searches on e-cigarettes and prevalence data; and assess the predictive capabilities of GT in Europe.

Methods

Eurobarometer data extraction

E-cigarette prevalence data were obtained for 27 European countries from four cross-sectional Eurobarometer surveys of people aged 15 and over: wave 77.1 (2012), wave 82.4 (2014), wave 87.1 (2017) and wave 93.2 (2020).^{55–58} The Eurobarometer is a series of public opinion surveys, conducted on behalf of the European Commission since 1974, that address a range of topical issues related to the EU via its member states.⁵⁹ All Eurobarometer surveys employ a multi-stage sampling procedure in each country, where the number of sampling points is proportional to population size and density. A total of 26,751, 27,801, 26,853 and 28,300 individuals were interviewed for 2012, 2014, 2017 and 2020, respectively. Data were downloaded as Excel files as both point estimates and percentages.

GT data extraction

GT monthly time series data were downloaded from the GT website (<https://trends.google.com/trends/>) as .csv files.³⁴ GT data were extracted as relative search term (RSV), a normalised measure from 0 to 100, where 100 represents the month with the most searches in a specified time period and location, and all other search volumes are relative to this peak search date. GT provides the option to query 'topics' or 'search terms', where search terms represent exact, literal words or phrases typed in by a user, and topics represent a group of search terms that share the same concept in any language. For this reason, topics were chosen to compare adequately searches between the languages spoken across the 27 European countries. The two most relevant topics were determined to be "electronic cigarette" and "vaporizer". Overall, separate datasets were downloaded for all 27 countries for each search topic specifying only web searches spanning January 2012 to December 2020, to align with the range of the Eurobarometer survey years. GT data were extracted from the website on the same day.

Data analysis

All analyses were performed in RStudio™ version 2023.06.01, an open-source language for statistical computing and graphics.⁶⁰

To investigate online interest in e-cigarettes across Europe, GT data were firstly smoothed using locally weighted scatterplot smoothing (LOWESS) for visual inspection.⁶¹ It was further chosen for its ability to give less weight to data points with large errors, producing a robust curve with limited outlier influence. A smoothing parameter of 0.25 was used for both search terms based on visual inspection, meaning that 25% of the nearest data points are used in the local regression for each point. To assess trends across all countries, median RSV was calculated and chosen for its reduced susceptibility to extreme outliers. LOWESS utilises a locally weighted regression of y (RSV) on x (time), making it adept for smoothing noisy and sparse data without assuming any particular data distribution (non-parametric).⁶²

Change point analysis (CPA) was used to detect statistically significant changes in online interest

over time. The Pruned Exact Linear Time (PELT) algorithm was chosen to identify change points, selecting the segments with the smallest penalised cost.⁶³ To determine the optimal number of change points in each time series, the change points for a Range of Penalties (CROPS) algorithm was implemented to search over a range of penalty values and the final penalty was chosen via diagnostic elbow plots.⁶⁴ Time series cross correlation was applied to investigate the relationship between GT and Eurobarometer data and to check whether GT was a robust proxy for cross-sectional prevalence data.

We identified countries with statistically significant correlations at negative or zero lags. Since GT data were monthly and Eurobarometer data provided only one estimate per survey year, GT data were aggregated to be yearly median RSV per country. To provide a Europe estimate, data were aggregated as the median across all 27 countries annually. Years with GT data but not Eurobarometer data (i.e. 2013, 2015, 2016, 2018 and 2019) were imputed to provide sufficient data points for correlation.⁶⁵ Cross-correlation was chosen to account for time dependency between two time series, x_{t+h} (GT) and y_t (Eurobarometer), where h is the time lag (in years) between the two series. A negative h value indicates a correlation when y (Eurobarometer) is at time t and x (GT) is at time $t - h$, and a positive h places x at time $t + h$. To illustrate, a correlation with lag -1 implies Eurobarometer data lags behind GT by 1 year, a lag of $+1$ implies GT lags behind Eurobarometer data by 1 year, and a lag of 0 implies the two series are contemporaneous. For the purposes of GT forecasting, the most relevant correlations would be at negative lags, implying online searches influenced later e-cigarette use, or potentially at a lag of 0. Cross-correlation analysis was performed for each country individually and for Europe as a whole for each search topic with a significance threshold of 0.95.

For each topic, countries that displayed significant correlations between GT and Eurobarometer data were taken forward for GT time series forecasting to determine whether GT data on e-cigarettes prevalence showed robust predictive ability. Three forecasting models were used for each series in each country in increasing order of complexity: simple exponential smoothing (SES),⁶⁶ autoregressive integrated moving average (ARIMA)⁶⁷ and boosted ARIMA, subject to comprehensive evaluation with residual diagnostics, and subsequently compared using the Mean Absolute Percentage Error (MAPE) forecasting metric. For all models, the training set consisted of all data points except the most recent 6 months of data (July to December 2020), which was reserved as the test set on which forecasts were made. Finally, all forecast models were evaluated with mean absolute percentage error (MAPE), a popular metric for time series modelling.^{66,68} To assess the predictive ability of GT searches, time series forecasting was performed on the three countries with the highest correlation coefficient at any lag for each search topic: Malta, the Netherlands and Spain for “electronic cigarette” and Czechia, Spain and Greece for “vaporizer”.

Results

The LOWESS regression was used to visualize the temporal changes in online e-cigarette interest (Figure 1). The median RSV for both topics (i.e., “electronic cigarette” and “vaporizer”) followed

similar trends over time (2012–2020), and showed a general increase, despite minor fluctuations in online popularity at certain times. Notably, search interest in both topics dropped considerably from mid-2019 to mid-2020 and the “vaporizer” topic also displayed a slight drop in early 2014. The median RSV of both topics was similar, alternating in popularity, until roughly early 2014; from there on, “vaporizer” had a greater consistent popularity than “electronic cigarette”, indicating that in Europe this terminology became considerably more searched online after this point.



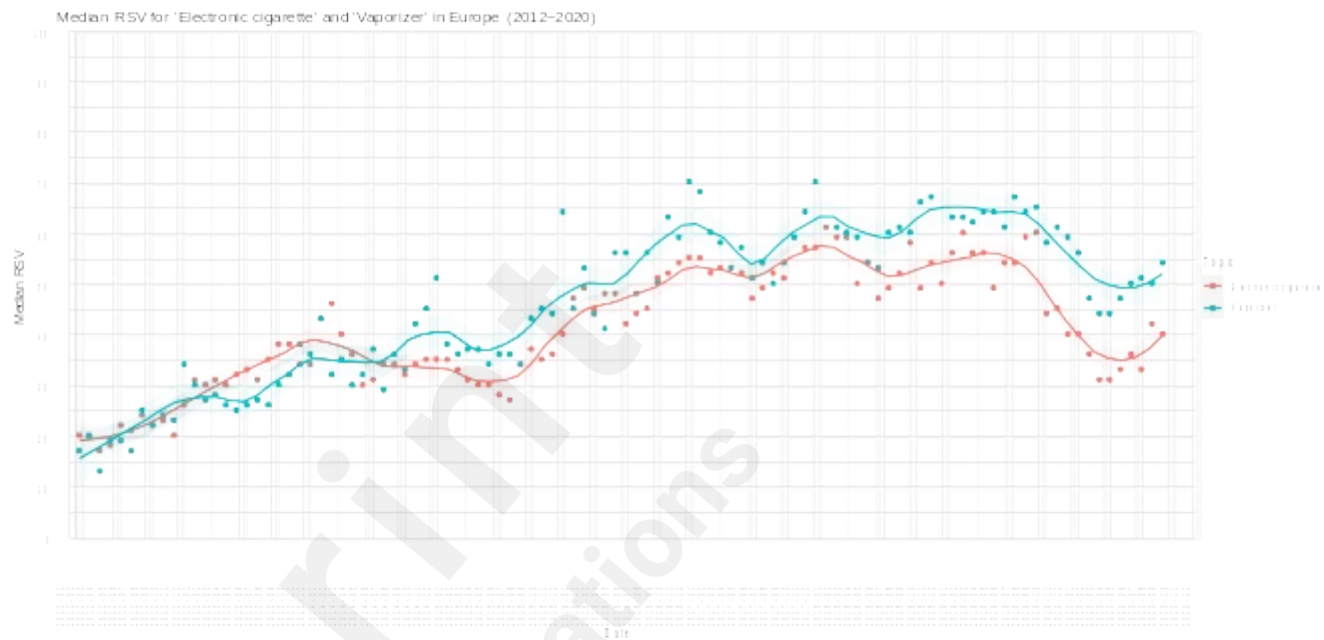


Figure 1LOWESS-smoothed curves showing the monthly median RSV across 27 European countries from January 2012 to December 2020 for the GT search topics “Electronic cigarette” (red) and “vaporizer” (blue)

The CPA diagnostic plot for “electronic cigarette” determined three significant change points in median RSV across Europe in December 2012, August 2016 and September 2019 (Figure 2). From visual inspection, only the latter date represented a considerable drop in median RSV. For “vaporizer”, the diagnostic plot determined that the optimal number of change points across Europe was four: October 2013, September 2016, January 2020 and May 2020 (Figure 2). Visually, all except January 2020 represented a surge in median RSV. Therefore, across all 27 countries, both search topics experienced a significant increase in median search volume around mid-to-late 2016.

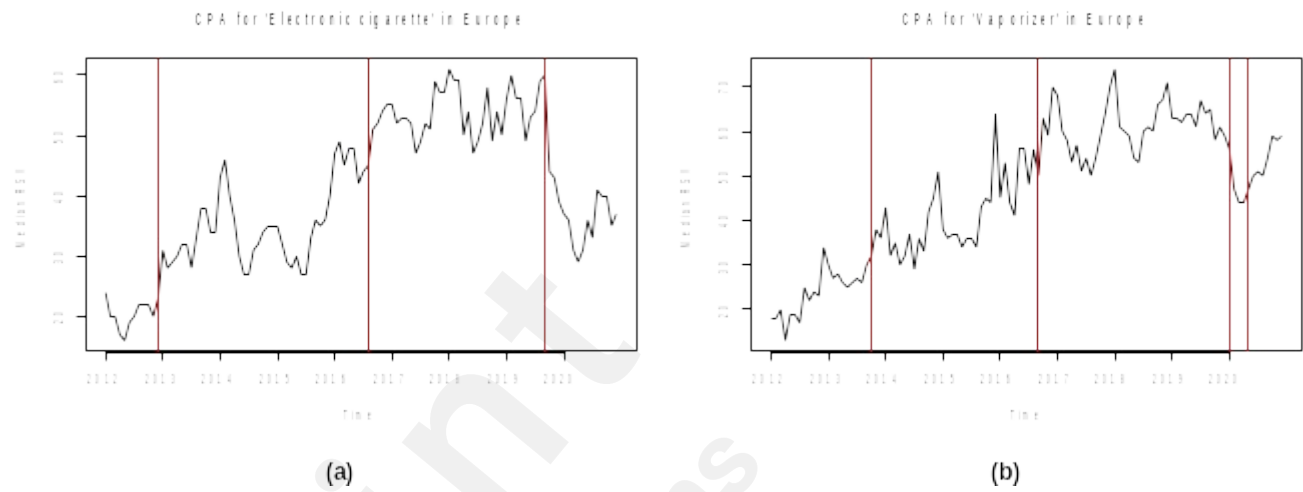


Figure 2: Time series plots indicating significant change points in median RSV across Europe for the GT search topics “electronic cigarette” and “vaporizer”.

Temporal relationship between prevalence and GT searches

Cross-correlation analysis was performed to ascertain the temporal relationship between yearly e-cigarette prevalence and median yearly GT searches for each of the 27 EU countries, as well as the EU as a whole. Eleven countries showed a significant positive correlation for the search topic “electronic cigarette” at lags of -2 to 0 years, where negative lags indicated that GT searches correlated with the e-cigarette prevalence data of the previous years, and lag 0 indicated GT searches correlated with the e-cigarette prevalence data of the same year. A positive correlation meant the e-cigarette prevalence data correlated with the GT searches of the previous years. Spain had the highest correlation coefficient of 0.82 at lag -1 ($P = .014$), where the -1 lag implies that GT preceded e-cigarette prevalence by 1 year. For the “vaporizer” topic, seven countries displayed significant positive correlations at lags of -2 to 0 years, with Czechia (i.e., the Czech Republic) producing the highest correlation coefficient of 0.86 at lag -1 ($P < .01$). For both topics, no significant correlations were detected for Europe as a whole, and no countries displayed significant negative correlations or positive lags. All countries with significant correlations are detailed in Appendix 1.

Predictive ability of GT searches

For the “electronic cigarette” search topic, the ARIMA models yielded better forecasting performances (as indicated by MAPE) than SES models for Malta and the Netherlands. Compared with SES, the ARIMA model reduced the MAPE by 2.14% and 0.05% for Malta and the Netherlands, respectively. Importantly, the residual diagnostic plots affirmed that no model assumptions were violated. The Ljung-Box test⁶⁹ further confirmed the absence of autocorrelated residuals ($P = .87$ for Malta and $P = .92$ for the Netherlands).

However, for Spain, the ARIMA (0,1,0) instead increased the forecasting error by 3.14%. The SES model produced significant residual autocorrelation (Ljung-box test $P < .001$) at lags of 1, 7, 8 and 9 months. Conversely, the ARIMA model showed a statistically non-significant Ljung-box test ($P = .48$)—indicating no autocorrelation of the residuals—and diagnostic plots confirmed no violated assumptions and a robust model fit. The ARIMA model’s forecast for Spain was not as robust as the SES model.

Both the ARIMA and SES models produced the highest MAPE for Malta, indicating less accurate RSV forecasting compared with the other two countries. Therefore, a boosted ARIMA model was applied to enhance modelling accuracy through gradient boosting. The ARIMA part of the model was specified with the ARIMA (1, 1, 1) parameters automatically selected by the previous ARIMA model, whilst the XGBoost⁷⁰ part of the model was tuned with time series cross validation that improved the forecasting accuracy. Forecasting plots for all “electronic cigarette” models are visualised in Figure 3. Only the most recent six months (test set) are shown.

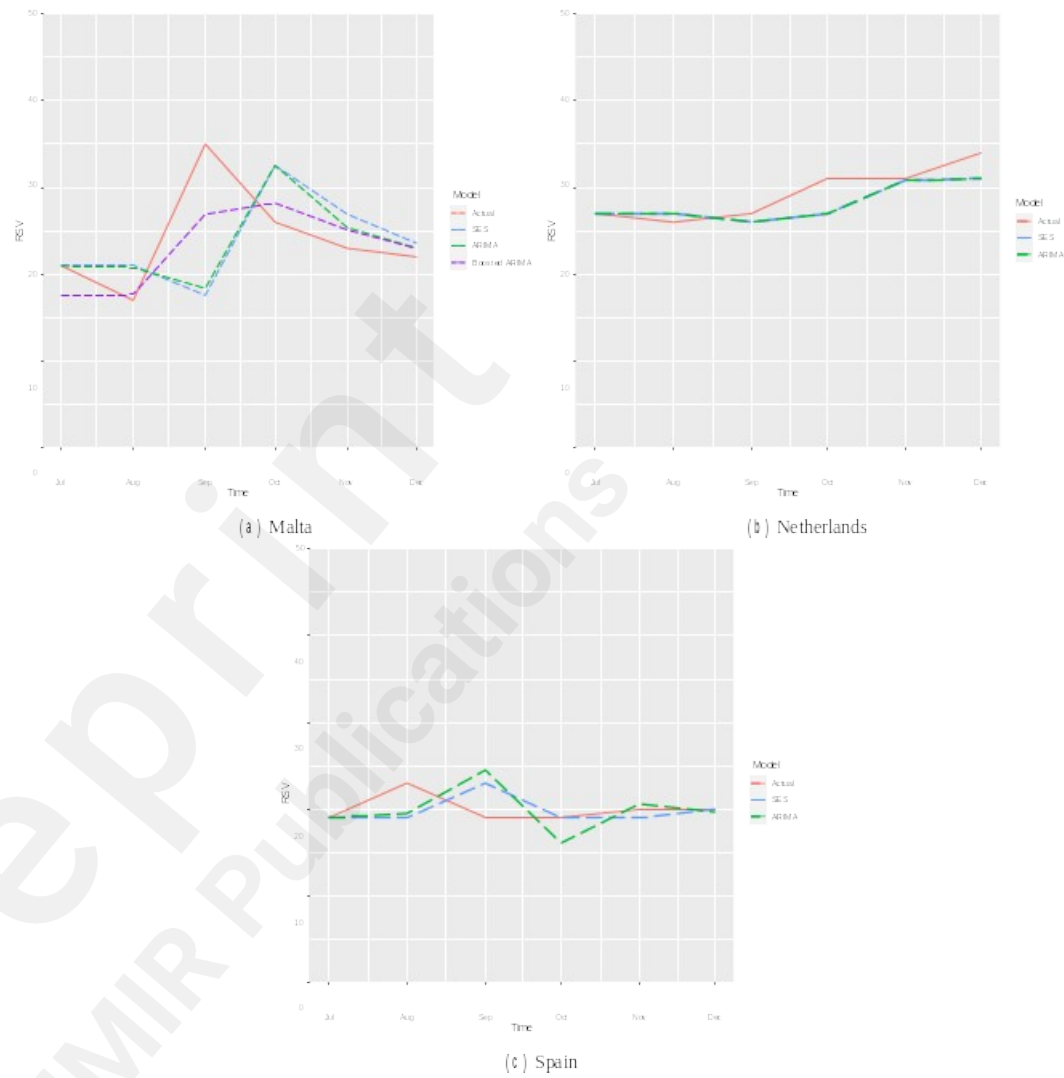


Figure 3Time-series forecasting models for the GT search topic "electronic cigarette" in a) Malta, b) the Netherlands and c) Spain

For the “vaporizer” topic, the automated algorithm to find the optimal ARIMA models selected a seasonal ARIMA (SARIMA) for all three countries, therefore capturing the intricate yearly seasonal dynamics in the RSV time series data. SARIMA models outperformed SES models (assessed by MAPE) for Greece and Czechia, suggesting that a seasonal component of the time series was better captured by SARIMA. However, for Spain the SES outperformed SARIMA, highlighting the importance of testing multiple forecasting models to obtain an optimal fit based on inherent data characteristics. Diagnostic plots confirmed model robustness, except for the Czechia SES model, where a significant Ljung-box test ($P < .001$) affirmed the significant lack of model fit.

For both SARIMA and SES models, the highest MAPE came from Greece, meaning that the models were unable to forecast RSV values as accurately as they did for Czechia and Spain. Boosted ARIMA was applied to improve modelling. The forecasted values for the SES, ARIMA and boosted ARIMA models of Czechia, Spain and Greece are visualised in Figure 4. Only the most recent six months (test set) are shown for visual purposes.

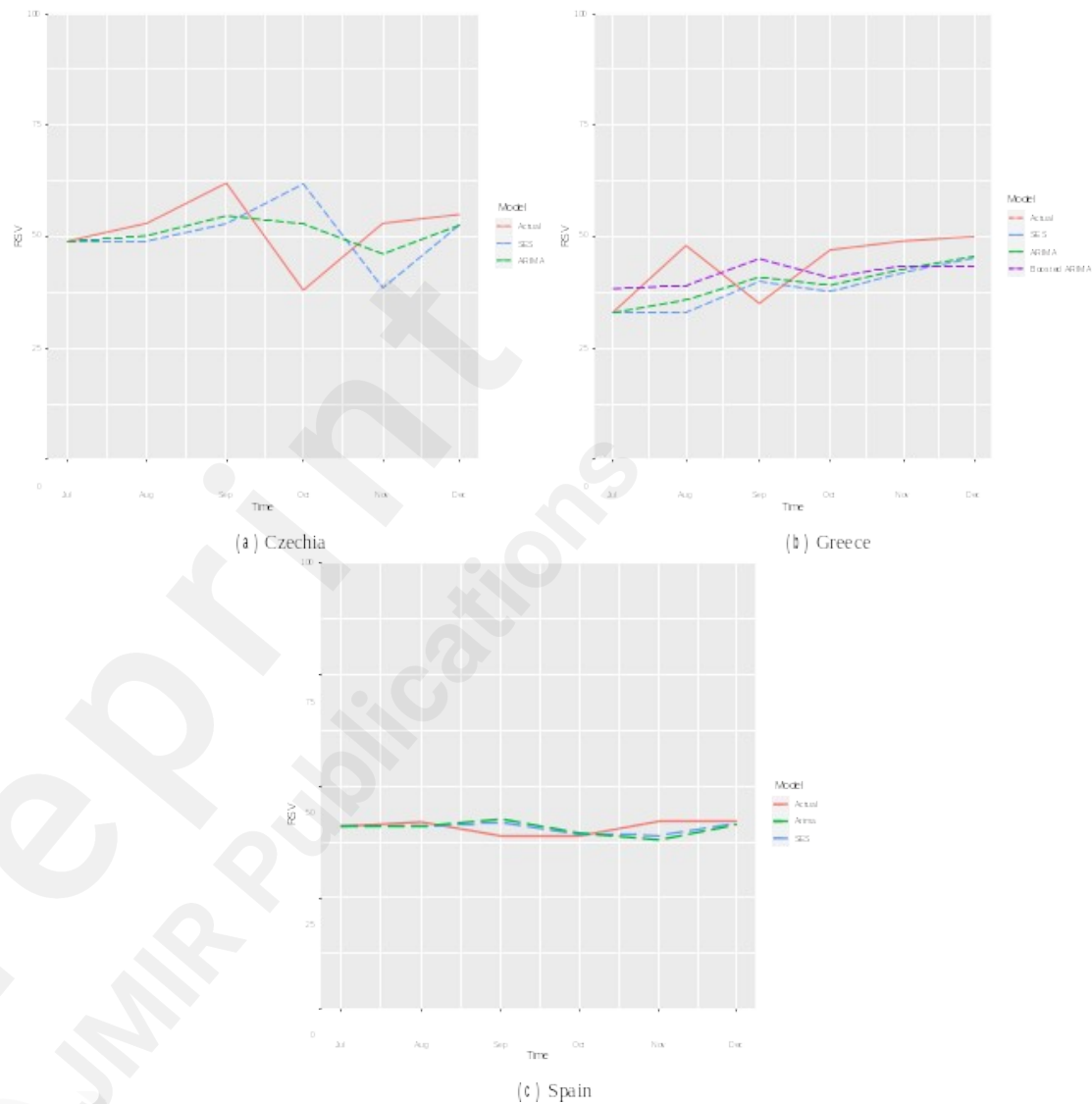


Figure 4Time-series forecasting models for the GT search topic "vaporizer" in a) Czechia, b) Greece and c) Spain

Discussion

The two GT search topics (i.e., “electronic cigarette” and “vaporizer”) showed positive correlation between GT and real prevalence data in some, European countries at lags between -2 and 0 years and, therefore, online interest appears to precede, or at least coincide with, e-cigarette use. In countries with the strongest correlation, GT displays substantial forecasting ability, although error metrics vary depending on the country and specific model used.

The search topic “vaporizer” became more popular than “electronic cigarette” after 2014, likely reflecting colloquial changes in speech when the word ‘vape’ more than doubled in usage.⁷¹ Such changes in popularity were captured with CPA, where seven different months were identified with a significant change in online interest across both topics, all of which displayed a surge in searches except for January 2020. This was similar to a separate study that observed significant change points in 2012 and 2013 for the “electronic cigarette” search term in the USA.⁷² The identified change points could be attributable to real-life policies and/or events; for example, the spike in December 2012 for “electronic cigarette” may correspond to the circulation of the draft EU TPD policy proposing e-cigarette restrictions across the EU for the first time.⁷³ The World Health Organization released a report in 2016 outlining the role of e-cigarettes in smoking cessation,⁷⁴ which may have contributed to the significant spike in online interest in August and September 2016. Furthermore, cases of e-cigarette or vaping use-associated lung injury EVALI could have exacerbated concerns and curiosity about the health effects of e-cigarettes; cases were first reported in April 2019 and peaked in September 2019,^{75,76} which was a significant change point in this analysis. Notably, it was at this time that the first case of EVALI was officially documented in Spain.⁷⁷ The rising popularity of social media platforms is reportedly associated with e-cigarette use in adolescents,⁷⁸ which may have further perpetuated spikes in search interest in recent years. Cross-correlation analyses suggest that GT can be used as a proxy for Eurobarometer data in over half of the countries investigated, similar to studies in the USA.^{72,79}

Only some countries showed a significant correlation. It is possible that countries lacking significant correlations had low internet penetration and therefore GT was unable to capture e-cigarette users’ online behaviour accurately. Internet penetration might have played a role in explaining why 11 of the 15 countries with significant correlation had an internet penetration over 90%, and Europe as a whole has a high internet penetration of 93%.^{36,80} Yet four countries had rates below 90%. Particularly, Bulgaria has the lowest internet penetration in the EU at 70%,⁸⁰ but still showed a significant correlation for the “electronic cigarette” search topic; therefore, perhaps it is not about how many people in a country have access to the internet, but rather the people who do also happen to be e-cigarette users. Age could be another factor as young people are the predominant group using both the internet and e-cigarettes.^{81,82} Data suggest that internet users are typically of higher socioeconomic classes, which has also been associated with higher

odds of e-cigarette use in current and ex-smokers.^{83,84} Therefore, it could be speculated that countries whose populations are of an older age and/or lower socioeconomic class may have lower e-cigarette prevalence and difficulty detecting a relationship between GT and prevalence data. The varying e-cigarette policies between European countries may help to explain the discrepancies in correlation. For example, lenient policies in countries such as Czechia and Bulgaria may reduce any hesitation to search online and also facilitate easier access to e-cigarettes.⁸⁵ The UK, which displayed a significant correlation in this analysis, actually encourages e-cigarette use for smoking cessation and provides them to tobacco smokers in hospitals.⁸⁶ On the other hand, strict restrictions could instead steer e-cigarette users towards online searches; for example, in Greece the internet and specialised vape shops are the two primary sources of e-cigarettes due to restrictions in shops and markets.⁸⁷ Other countries with heavy restrictions found to have a significant correlation in this analysis include Estonia and the Netherlands.^{88–90} However, strict restrictions do not necessarily preclude illicit, especially underage, e-cigarette use, and in stricter countries illegal use may be better reflected in anonymised GT data than formal cross-sectional surveys.^{91,92} Therefore, it is evident that the effects of country-specific policies are unclear.

In countries with highly correlated GT and Eurobarometer datasets, forecasting analysis revealed that GT is potentially a valuable tool for predicting future online interest in e-cigarettes, although model performance differed between countries. Nonetheless, more complex ARIMA models outperformed simpler SES models for all countries except Spain, possibly because SES does not detrend the data or capture more complex relationships, unlike the automated ARIMA procedure. These findings highlight the necessity of testing multiple models to find an optimal fit for each data structure.⁹³ Additionally, ARIMA boost—a relatively new method in time series forecasting—further improved forecasting performance. For example, one study applied long short-term memory networks (a machine learning method) to univariate GT data to predict COVID-19 cases and found the model outperformed ARIMA in terms of forecast metrics.⁹⁴ Additionally, using grey-wolf optimisation further improved forecast errors. Since this study is among the first to forecast GT e-cigarette data across Europe, it is challenging to determine if the forecast metrics in this analysis surpass or fall below other benchmarks; although one study created a predictive model of online e-cigarette interest in the UK,⁴³ this work did not reserve test set data to make forecasts on and in general the methodology is not comparable to the present study. Furthermore, there are papers forecasting other health topics, such as Dengue virus, COVID-19 vaccinations and pertussis, with forecast metrics either similar or worse than this analysis; but these comparisons only serve to confirm the forecasting ability of GT within healthcare, since the models were applied to entirely different datasets, topics, locations and/or forecast periods.^{95–98} It should also be emphasised that this research did not seek to create models with the highest forecasting accuracy, but rather to assess the general forecasting potential of GT for e-cigarettes.

Methodological considerations

From a methodological point of view, CPA provides a more robust method of detecting significant surges or collapses in online interest compared with visual inspection.^{99,100} Change points do not require knowledge of the underlying reason for the change, and the chosen PELT algorithm has been shown to be more accurate and faster than other change point algorithms.^{101,102} However, determining the number of change points with diagnostic plots is subjective, therefore the optimal number of change points could be inconsistent if the study was repeated. Furthermore, cross-correlation was harnessed for its ability to identify correlations at different lags between two time series, providing temporal insight into how one dataset later influences the other. A problem that was encountered during the analysis was that the Eurobarometer data was only available for four years, offering few datapoints to investigate. One explored solution was to use the Dirichlet distribution to obtain monthly estimates.¹⁰³ However, the substantial number of months requiring estimation posed a challenge and introduced considerable uncertainty into the analysis. For this reason, the problem was addressed by imputing the missing Eurobarometer years, in line with other reports, which compromised the granularity of correlation but reduced the number of estimated data points required.¹⁰⁴

The methodology of implementing forecasting models of varying complexity proved to be advantageous as not all time series were best captured by more complex models. Of these models, SES was chosen for its simplicity and ability to give more weight to recent observations. However, disadvantages of this method include the inability to detect seasonality and trends and the assumption that the future is a linear function of past observations.^{66,105} Furthermore, exponential smoothing forecasts can lag behind the actual trend, as seen in the present analysis, underestimating increasing trends and overestimating decreasing trends.¹⁰⁶ For these reasons, ARIMA models were chosen to capture more complex patterns in time series, including seasonality, as detected for the “vaporizer” ARIMA models.⁶⁷ Boosted ARIMA was used for its potential to improve modelling errors with gradient boosting and add further complexity to the fitting procedure. However, the disadvantages of this technique include increased computation expense and limited interpretability. As residuals are calculated on the training set and forecasts are calculated on the test set, it is likely that the ARIMA model captured temporal patterns in the training data well, but this did not guarantee accurate forecasts on new test data. This highlights the importance of considering inherent data characteristics and implementing multiple forecasting models to make optimal forecast.

General limitations

The Eurobarometer survey years are limited in comparability because the 2012 edition did not distinguish between current and former e-cigarette users and as detailed previously, cross-sectional data suffers from drawbacks such as self-report bias and only

capturing one snapshot in time. GT data are sourced from a random sample which can produce slightly different results for the same search on different days or devices. GT data only represents people with internet access, which may potentially exclude certain subpopulations and limit its applicability as a predictive method for countries with very limited internet coverage. Accordingly, it is possible that prevalence and/or GT search volume varies within each country, but only country-level information was available for both datasets. It is also not possible to extract RSV data representing 27 countries together from GT; therefore, median RSV was calculated to assess Europe as a whole. However, RSV represents search term popularity relative to other time points for a specified period and location; for example, an RSV of 50 only represents 50% of the peak search volume (denoted as RSV 100) for that specific country and period, and not the peak of any other country and period.

Moreover, the cross-correlation analysis is not causal, and there is no indication that searches for e-cigarettes cause a change in e-cigarette use. As discussed, granularity was lost when performing cross-correlation on imputed yearly data, and in the future, it would be ideal to obtain more frequent (e.g., every six months, quarter or monthly) e-cigarette prevalence estimates to investigate the relationship between GT and Eurobarometer data.

Conclusions

This study highlights the value of GT as a potential proxy for real e-cigarette prevalence data in select European countries and further demonstrates its utility as a forecasting tool for future e-cigarette use based on its correlative relationship with prevalence data. Furthermore, GT may provide a valuable tool that outputs near real-time, inexpensive and publicly available data, which could be particularly important in countries without the means to collect frequent cross-sectional data. However, additional work is required to delineate the reasons why GT did not demonstrate significant correlation in some European countries, and to validate the forecasting results. Should this be done, this could inform the extent to which GT can be co-opted as a public health tool, informing policy makers of appropriate, preventative policies that balance the benefits and drawbacks of e-cigarettes to, ultimately, navigate the rising prevalence of e-cigarettes across Europe.

The public health implications of forecasting e-cigarette trends are important. GT's correlative relationship with cross-sectional data and predictive ability for e-cigarette interest underscores its potential for public health policies regarding e-cigarette use in Europe, although some countries may benefit more from such models depending on the resulting forecast metrics. Particularly, forecasting future behaviours allows for preventative policies, precluding the need to act after the health effects of e-cigarettes have already affected the population.

References

1. Cao DJ, Aldy K, Hsu S, et al. Review of Health Consequences of Electronic Cigarettes and the Outbreak of Electronic Cigarette, or Vaping, Product Use-Associated Lung Injury. *J Med Toxicol.* 2020;16(3):295-310. doi:10.1007/s13181-020-00772-w
2. De Silva R, Silva D, Piumika L, et al. Impact of global smoking prevalence on mortality: a study across income groups. *BMC Public Health.* 2024;24(1):1786. doi:10.1186/s12889-024-19336-6
3. Naser AY, Mansour MM, Alanazi AFR, et al. Hospital admission trends due to respiratory diseases in England and Wales between 1999 and 2019: an ecologic study. *BMC Pulm Med.* 2021;21(1):356. doi:10.1186/s12890-021-01736-8
4. Al Rajeh AM, Naser AY, Siraj R, et al. Acute upper respiratory infections admissions in England and Wales. *Medicine (Baltimore).* 2023;102(21):e33616. doi:10.1097/MD.00000000000033616
5. Alwafi H, Naser AY, Ashoor DS, et al. Trends in hospital admissions and prescribing due to chronic obstructive pulmonary disease and asthma in England and Wales between 1999 and 2020: an ecological study. *BMC Pulm Med.* 2023;23(1):49. doi:10.1186/s12890-023-02342-6
6. Chen X, Zhou CW, Fu YY, et al. Global, regional, and national burden of chronic respiratory diseases and associated risk factors, 1990–2019: Results from the Global Burden of Disease Study 2019. *Front Med.* 2023;10. doi:10.3389/fmed.2023.1066804
7. Holtjer JCS, Bloemsma LD, Beijers RJHCG, et al. Identifying risk factors for COPD and adult-onset asthma: an umbrella review. *Eur Respir Rev.* 2023;32(168). doi:10.1183/16000617.0009-2023
8. Hammond A, Halliday A, Thornton HV, Hay AD. Predisposing factors to acquisition of acute respiratory tract infections in the community: a systematic review and meta-analysis. *BMC Infect Dis.* 2021;21(1):1254. doi:10.1186/s12879-021-06954-3
9. Smoking in Europe. Statista. Accessed September 21, 2025. <https://www.statista.com/topics/3767/smoking-habits-in-europe/>
10. Khaled Z, Dahmash EZ, Koner J, Al Ani R, Alyami H, Naser AY. Assessment of vaping devices as an alternative respiratory drug delivery system. *Drug Dev Ind Pharm.* 2022;48(9):446-456. doi:10.1080/03639045.2022.2123926
11. Lindson N, Butler AR, McRobbie H, et al. Electronic cigarettes for smoking cessation - Lindson, N - 2025 | Cochrane Library. Accessed September 21, 2025. <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010216.pub9/full>
12. Levett JY, Filion KB, Reynier P, Prell C, Eisenberg MJ. Efficacy and Safety of E-Cigarette Use for Smoking Cessation: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Am J Med.* 2023;136(8):804-813.e4. doi:10.1016/j.amjmed.2023.04.014
13. Wang TW, Neff LJ, Park-Lee E, Ren C, Cullen KA, King BA. E-cigarette Use Among Middle and High School Students — United States, 2020. *Morb Mortal Wkly Rep.* 2020;69(37):1310-1312. doi:10.15585/mmwr.mm6937e1

14. Oliver BE, Jones SE, Hops ED, Ashley CL, Miech R, Mpofu JJ. Electronic Vapor Product Use Among High School Students - Youth Risk Behavior Survey, United States, 2021. *MMWR Suppl.* 2023;72(1):93-99. doi:10.15585/mmwr.su7201a11
15. Mirbolouk M, Boakye E, Obisesan O, et al. E-cigarette use among high school students in the United States prior to the COVID-19 pandemic: Trends, correlates, and sources of acquisition. *Prev Med Rep.* 2022;29:101925. doi:10.1016/j.pmedr.2022.101925
16. Bareham D, Ahmadi K, Elie M, Jones AW. E-cigarettes: controversies within the controversy. *Lancet Respir Med.* 2016;4(11):868-869. doi:10.1016/S2213-2600(16)30312-5
17. Izquierdo-Condoy JS, Naranjo-Lara P, Morales-Lapo E, et al. Direct health implications of e-cigarette use: a systematic scoping review with evidence assessment. *Front Public Health.* 2024;12. doi:10.3389/fpubh.2024.1427752
18. Toledo EFV, Simões IF, Farias MT de, et al. A Comprehensive Review of the Harmful Compounds in Electronic Cigarettes. *Toxics.* 2025;13(4):268. doi:10.3390/toxics13040268
19. Gould TJ. Epigenetic and long-term effects of nicotine on biology, behavior, and health. *Pharmacol Res.* 2023;192:106741. doi:10.1016/j.phrs.2023.106741
20. Castro EM, Lotfipour S, Leslie FM. Nicotine on the developing brain. *Pharmacol Res.* 2023;190:106716. doi:10.1016/j.phrs.2023.106716
21. Sansone L, Milani F, Fabrizi R, et al. Nicotine: From Discovery to Biological Effects. *Int J Mol Sci.* 2023;24(19):14570. doi:10.3390/ijms241914570
22. Mao C, Li D, Zhou E, Zhang J, Wang C, Xue C. Nicotine exacerbates atherosclerosis through a macrophage-mediated endothelial injury pathway. *Aging.* 2021;13(5):7627-7643. doi:10.18632/aging.202660
23. Zhong J, Cao S, Gong W, Fei F, Wang M. Electronic Cigarettes Use and Intention to Cigarette Smoking among Never-Smoking Adolescents and Young Adults: A Meta-Analysis. *Int J Environ Res Public Health.* 2016;13(5):465. doi:10.3390/ijerph13050465
24. Conner M, Grogan S, Simms-Ellis R, et al. Do electronic cigarettes increase cigarette smoking in UK adolescents? Evidence from a 12-month prospective study. *Tob Control.* 2018;27(4):365-372. doi:10.1136/tobaccocontrol-2016-053539
25. Chan GCK, Stjepanović D, Lim C, et al. Gateway or common liability? A systematic review and meta-analysis of studies of adolescent e-cigarette use and future smoking initiation. *Addiction.* 2021;116(4):743-756. doi:10.1111/add.15246
26. Hajek P, Phillips-Waller A, Przulj D, et al. E-cigarettes compared with nicotine replacement therapy within the UK Stop Smoking Services: the TEC RCT. *Health Technol Assess.* 2019;23(43):1-82. doi:10.3310/hta23430
27. Levett JY, Filion KB, Reynier P, Prell C, Eisenberg MJ. Efficacy and Safety of E-Cigarette Use for Smoking Cessation: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Am J Med.* 2023;136(8):804-813.e4. doi:10.1016/j.amjmed.2023.04.014
28. Vardavas CI. European Tobacco Products Directive (TPD): current impact and future steps. *Tob Control.* 2022;31(2):198-201. doi:10.1136/tobaccocontrol-2021-056548
29. Brown R, Godwin JV, Page N, et al. Implementation of e-cigarette regulation through the EU

- Tobacco Products Directive (2016) in Wales, Scotland and England from the perspectives of stakeholders involved in policy introduction and enforcement. doi:10.18332/tpc/134370
30. Yang JJ, Ou TS, Lin HC, Kyung Nam J, Piper ME, Buu A. Retrospective and Real-time Measures of the Quantity of E-cigarette Use: An Ecological Momentary Assessment Study. *Nicotine Tob Res.* 2023;25(10):1667-1675. doi:10.1093/ntr/ntad094
 31. Adebisi YA, Lucero-Prisno DE, Ogunkola IO. Self-rated health differences between exclusive e-cigarette users and exclusive cigarette smokers: evidence from the 2017–2019 Scottish Health Survey. *Intern Emerg Med.* 2025;20(3):771-784. doi:10.1007/s11739-025-03873-y
 32. Hallingberg B, Maynard OM, Bauld L, et al. Have e-cigarettes renormalised or displaced youth smoking? Results of a segmented regression analysis of repeated cross sectional survey data in England, Scotland and Wales. *Tob Control.* 2020;29(2):207-216. doi:10.1136/tobaccocontrol-2018-054584
 33. Alibudbud R. Google Trends for health research: Its advantages, application, methodological considerations, and limitations in psychiatric and mental health infodemiology. *Front Big Data.* 2023;6:1132764. doi:10.3389/fdata.2023.1132764
 34. Google. Google Trends. Accessed September 22, 2025. <https://trends.google.com/trends?geo=JO&hl=en-US>
 35. Nuti SV, Wayda B, Ranasinghe I, et al. The Use of Google Trends in Health Care Research: A Systematic Review. Voracek M, ed. *PLoS ONE.* 2014;9(10):e109583. doi:10.1371/journal.pone.0109583
 36. Eurostat. Digital economy and society statistics - households and individuals. 2025. Accessed September 22, 2025. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals
 37. Eurostat. Young people - digital world. 2025. Accessed September 22, 2025. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Young_people_-_digital_world
 38. Wang D, Lang JC, Chen YH. Assessment of using Google Trends for real-time monitoring of infectious disease outbreaks: a measles case study. *Sci Rep.* 2024;14(1):1-9. doi:10.1038/s41598-024-60120-8
 39. Ito T. Global monitoring of public interest in preventive measures against COVID-19 via analysis of Google Trends: an infodemiology and infoveillance study. *BMJ Open.* 2022;12(8):e060715. doi:10.1136/bmjopen-2021-060715
 40. Troelstra SA, Bosdriesz JR, Boer MR de, Kunst AE. Effect of Tobacco Control Policies on Information Seeking for Smoking Cessation in the Netherlands: A Google Trends Study. *PLOS ONE.* 2016;11(2):e0148489. doi:10.1371/journal.pone.0148489
 41. Tieks A, Troelstra SA, Hoekstra T, Kunst AE. Associations of the Stoptober smoking cessation program with information seeking for smoking cessation: A Google trends study. *Drug Alcohol Depend.* 2019;194:97-100. doi:10.1016/j.drugalcdep.2018.08.040
 42. Jagomast T, Tangemann-Münstedt I, Lütthgen M, et al. Google Trends Analysis of Interest in Smoking Cessation, Treatment Options, and Electronic Nicotine Delivery Systems Following Tax Increase and Introduction of Pictorial Health Warnings in Germany. *Tob Use Insights.* 2025;18:1179173X251338447. doi:10.1177/1179173X251338447

43. Sindhoo Z, Sindhoo S, Ghosh A, Kaur S, Soyiri I, Ahmadi K. Online Interest for Electronic Cigarettes Using Google Trends in the UK: A Correlation Analysis. *Subst Use Misuse*. 2023;58(14):1791-1797. doi:10.1080/10826084.2023.2247056
44. Mavragani A, Ochoa G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. *JMIR Public Health Surveill*. 2019;5(2):e13439. doi:10.2196/13439
45. Alibudbud R. Google Trends for health research: Its advantages, application, methodological considerations, and limitations in psychiatric and mental health infodemiology. *Front Big Data*. 2023;6:1132764. doi:10.3389/fdata.2023.1132764
46. Balubaid M, Al-Husayni F, Alwafi H, et al. The Prevalence of Smoking Among Medical Residents in Saudi Arabia, A Cross Sectional National Survey. *Bahrain Med Bull*. 2024;46(1).
47. Pan L, Morton J, Mbulo L, Dean A, Ahluwalia IB. Electronic cigarette use among adults in 14 countries: A cross-sectional study. *eClinicalMedicine*. 2022;47. doi:10.1016/j.eclinm.2022.101401
48. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457(7232):1012-1014. doi:10.1038/nature07634
49. Teng Y, Bi D, Xie G, et al. Dynamic Forecasting of Zika Epidemics Using Google Trends. *PLOS ONE*. 2017;12(1):e0165085. doi:10.1371/journal.pone.0165085
50. Young SD, Zhang Q. Using search engine big data for predicting new HIV diagnoses. *PLOS ONE*. 2018;13(7):e0199527. doi:10.1371/journal.pone.0199527
51. Verma M, Kishore K, Parija PP, et al. Investigating Google Trends to forecast acute febrile illness outbreaks in North India reported through the Integrated Disease Surveillance Program. *BMC Infect Dis*. 2025;25(1):431. doi:10.1186/s12879-025-10801-0
52. Chu AMY, Tsang JTY, Chan SSC, Chan LSH, So MKP. Utilizing Google Trends data to enhance forecasts and monitor long COVID prevalence. *Commun Med*. 2025;5(1):179. doi:10.1038/s43856-025-00896-6
53. Parker J, Cuthbertson C, Loveridge S, Skidmore M, Dyar W. Forecasting state-level premature deaths from alcohol, drugs, and suicides using Google Trends data. *J Affect Disord*. 2017;213:9-15. doi:10.1016/j.jad.2016.10.038
54. Barros JM, Melia R, Francis K, et al. The Validity of Google Trends Search Volumes for Behavioral Forecasting of National Suicide Rates in Ireland. *Int J Environ Res Public Health*. 2019;16(17):3201. doi:10.3390/ijerph16173201
55. European Commission. Special eurobarometer on attitudes of Europeans towards tobacco - European Commission. 2012. Accessed September 22, 2025. https://health.ec.europa.eu/latest-updates/special-eurobarometer-attitudes-europeans-towards-tobacco-2012-05-30_en
56. European Commission. Attitudes of Europeans towards tobacco and electronic cigarettes. 2014. Accessed September 22, 2025. https://health.ec.europa.eu/publications/attitudes-europeans-towards-tobacco-and-electronic-cigarettes_en
57. European Commission. Attitudes of Europeans towards tobacco and electronic cigarettes. 2017. Accessed September 22, 2025. https://health.ec.europa.eu/publications/attitudes-europeans-towards-tobacco-and-electronic-cigarettes-0_en

58. Attitudes of Europeans towards tobacco and electronic cigarettes. 2020. Accessed September 22, 2025. <https://europa.eu/eurobarometer/surveys/detail/2240>
59. European Commission. About Eurobarometer – Eurobarometer – Public opinion in the European Union. Accessed September 22, 2025. <https://europa.eu/eurobarometer/about/eurobarometer>
60. R Core Team. R: The R Project for Statistical Computing. 2022. Accessed September 22, 2025. <https://www.r-project.org/>
61. Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots. *J Am Stat Assoc.* 1979;74(368):829-836. doi:10.1080/01621459.1979.10481038
62. Jacoby WG. Loess:: a nonparametric, graphical tool for depicting relationships between variables. *Elect Stud.* 2000;19(4):577-613. doi:10.1016/S0261-3794(99)00028-1
63. Killick R, Fearnhead P, Eckley IA. Optimal Detection of Changepoints With a Linear Computational Cost. *J Am Stat Assoc.* 2012;107(500):1590-1598. doi:10.1080/01621459.2012.737745
64. Haynes K, Eckley IA, Fearnhead P. Computationally Efficient Changepoint Detection for a Range of Penalties. *J Comput Graph Stat.* 2017;26(1):134-143. doi:10.1080/10618600.2015.1116445
65. Mayer M. missRanger: Fast Imputation of Missing Values. Published online August 25, 2017;2.6.1. doi:10.32614/CRAN.package.missRanger
66. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. 2nd Edition. Vancouver: Melbourne (Australia): OTexts; 2018. Accessed September 23, 2025. <https://otexts.com/fpp2/index.html>
67. Nelson BK. Time Series Analysis Using Autoregressive Integrated Moving Average (ARIMA) Models. *Acad Emerg Med.* 1998;5(7):739-744. doi:10.1111/j.1553-2712.1998.tb02493.x
68. Weigend AS. *Time Series Prediction: Forecasting The Future And Understanding The Past*. 1st edition. Routledge; 2018. doi:10.4324/9780429492648
69. LJUNG GM, BOX GEP. On a measure of lack of fit in time series models. *Biometrika.* 1978;65(2):297-303. doi:10.1093/biomet/65.2.297
70. Teodorescu V, Obreja Braşoveanu L. Assessing the Validity of k-Fold Cross-Validation for Model Selection: Evidence from Bankruptcy Prediction Using Random Forest and XGBoost. *Computation.* 2025;13(5):127. doi:10.3390/computation13050127
71. Ayres J, Williams R. 8.5 Million Google Searches Related To ‘Vaping’ In 2014. Medindia. November 2, 2016. Accessed September 23, 2025. <https://www.medindia.net/news/85-million-google-searches-related-to-vaping-in-2014-157656-1.htm>
72. Ghosh A, Kaur S, Roub F. Use and interest of electronic nicotine delivery systems (ENDS): Assessing the validity of Google Trends. *Am J Drug Alcohol Abuse.* 2021;47(4):414-421. doi:10.1080/00952990.2021.1944171
73. European Commission. Revision of the Tobacco Products Directive. Health.europa.eu. December 3, 2024. Accessed September 23, 2025. <https://health.ec.europa.eu/tobacco/product-regulation/implementing-tobacco-products-directive-directive-201440eu/revision-tobacco->

products-directive_en

74. World Health Organization. Electronic nicotine delivery systems and electronic non-nicotine delivery systems (ENDS/ENNDS). November 7, 2016. Accessed September 23, 2025. [https://www.who.int/publications/m/item/electronic-nicotine-delivery-systems-and-electronic-non-nicotine-delivery-systems-\(ends-ennds\)](https://www.who.int/publications/m/item/electronic-nicotine-delivery-systems-and-electronic-non-nicotine-delivery-systems-(ends-ennds))
75. Medicine TLR. The EVALI outbreak and vaping in the COVID-19 era. *Lancet Respir Med*. 2020;8(9):831. doi:10.1016/S2213-2600(20)30360-X
76. Layden JE, Ghinai I, Pray I, et al. Pulmonary Illness Related to E-Cigarette Use in Illinois and Wisconsin — Final Report. *N Engl J Med*. 2020;382(10):903-916. doi:10.1056/NEJMoa1911614
77. Casanova GS, Amaro R, Soler N, et al. An imported case of e-cigarette or vaping associated lung injury in Barcelona. *Eur Respir J*. 2020;55(2). doi:10.1183/13993003.02076-2019
78. Vassey J, Galimov A, Kennedy CJ, Vogel EA, Unger JB. Frequency of social media use and exposure to tobacco or nicotine-related content in association with E-cigarette use among youth: A cross-sectional and longitudinal survey analysis. *Prev Med Rep*. 2022;30:102055. doi:10.1016/j.pmedr.2022.102055
79. Foxon F. Estimating E-Cigarette Use Prevalence among US Adolescents Using Vaping-Related Online Search Trends. *Subst Use Misuse*. 2021;56(10):1559-1563. doi:10.1080/10826084.2021.1936054
80. Petrosyan A. Internet penetration rate in the European Union from 2019 to 2022, by country. Statista. 2023. Accessed September 23, 2025. <https://www.statista.com/statistics/1246141/eu-internet-penetration-rate/>
81. Zheng X, Li W, Wong SW, Lin HC. Social media and E-cigarette use among US youth: Longitudinal evidence on the role of online advertisement exposure and risk perception. *Addict Behav*. 2021;119:106916. doi:10.1016/j.addbeh.2021.106916
82. Dai L, He Y, Tan Y, Yu Z, Zhu J. Online E-Cigarette Information Exposure and Its Association with E-Cigarette Use among Adolescents in Shanghai, China. *Int J Environ Res Public Health*. 2022;19(6):3329. doi:10.3390/ijerph19063329
83. Kundu S, Shaw S, Khan J, Chattopadhyay A, Baptista EA, Paswan B. Age, gender and socioeconomic patterns of awareness and usage of e-cigarettes across selected WHO region countries: evidence from the Global Adult Tobacco Survey. *BMJ Open*. Published online January 1, 2023. doi:10.1136/bmjopen-2022-070419
84. Kock L, Shahab L, West R, Brown J. E-cigarette use in England 2014-2017 as a function of socioeconomic profile. *Addict Abingdon Engl*. 2019;114(2):294-303. doi:10.1111/add.14446
85. Glahn A, Kyriakos CN, Loghin CR, et al. Tobacco control achievements and priority areas in the WHO Europe Region: A review. *Tob Prev Cessation*. 2018;4(April):15. doi:10.18332/tpc/89925
86. Farrimond H, Abraham C. Developing E-cigarette friendly smoking cessation services in England: staff perspectives. *Harm Reduct J*. 2018;15:38. doi:10.1186/s12954-018-0244-8
87. Diamantopoulou E, Barbouni A, Merakou K, Lagiou A, Farsalinos K. Patterns of e-cigarette use, biochemically verified smoking status and self-reported changes in health status of a

- random sample of vapes shops customers in Greece. *Intern Emerg Med*. 2019;14(6):843-851. doi:10.1007/s11739-018-02011-1
88. Global Action to End Smoking Organization. State of Smoking and Health in Estonia. Global Action to End Smoking. 2024. Accessed September 25, 2025. <https://globalactiontoendsmoking.org/research/tobacco-around-the-world/estonia/>
 89. Government of the Netherlands. Government measures to discourage smoking. June 19, 2024. Accessed September 25, 2025. <https://www.government.nl/topics/smoking/government-measures-to-discourage-smoking>
 90. Klein DE, Chaiton M, Kundu A, Schwartz R. A Literature Review on International E-cigarette Regulatory Policies. *Curr Addict Rep*. 2020;7(4):509-519. doi:10.1007/s40429-020-00332-w
 91. Freitas-Lemos R, Stein JS, Tegge AN, et al. Illegal Experimental Tobacco Marketplace II: effects of vaping product bans — findings from the 2020 International Tobacco Control Project. *Tob Control*. 2022;31(Suppl 3):s214-s222. doi:10.1136/tc-2022-057515
 92. Ghosh A, Kaur S, Roub F. Use and interest of electronic nicotine delivery systems (ENDS): Assessing the validity of Google Trends. *Am J Drug Alcohol Abuse*. 2021;47(4):414-421. doi:10.1080/00952990.2021.1944171
 93. Petelin G, Cenikj G, Eftimov T. Towards understanding the importance of time-series features in automated algorithm performance prediction. *Expert Syst Appl*. 2023;213:119023. doi:10.1016/j.eswa.2022.119023
 94. Prasanth S, Singh U, Kumar A, Tikkiwal VA, Chong PHJ. Forecasting spread of COVID-19 using google trends: A hybrid GWO-deep learning approach. *Chaos Solitons Fractals*. 2021;142:110336. doi:10.1016/j.chaos.2020.110336
 95. Husnayain A, Fuad A, Lazuardi L. Correlation between Google Trends on dengue fever and national surveillance report in Indonesia. *Glob Health Action*. 2019;12(1):1552652. doi:10.1080/16549716.2018.1552652
 96. Rangarajan P, Mody SK, Marathe M. Forecasting dengue and influenza incidences using a sparse representation of Google trends, electronic health records, and time series data. *PLOS Comput Biol*. 2019;15(11):e1007518. doi:10.1371/journal.pcbi.1007518
 97. Maugeri A, Barchitta M, Agodi A. Using Google Trends to Predict COVID-19 Vaccinations and Monitor Search Behaviours about Vaccines: A Retrospective Analysis of Italian Data. *Vaccines*. 2022;10(1):119. doi:10.3390/vaccines10010119
 98. Shepherd T, Mallen C. Measles and Pertussis outbreaks in England and Wales: a time-series analysis. *NIHR Open Res*. 2024;4:56. doi:10.3310/nihropenres.13607.1
 99. Sharma S, Swayne DA, Obimbo C. Trend analysis and change point techniques: a survey. *Energy Ecol Environ*. 2016;1(3):123-130. doi:10.1007/s40974-016-0011-1
 100. Taylor W. Change-Point Analysis: A Powerful New Tool For Detecting Changes. *Taylor Enterp Inc*. Published online 2000.
 101. Biswas S, Kuna KK. Optimal Change-Point Analysis of Pavement Condition Data for Identification of Homogeneous Sections. *J Transp Eng Part B Pavements*. 2023;149(1):04022078. doi:10.1061/JPEODX.PVENG-1104

102. Wambui GD, Waititu GA, Wanjoya A. The Power of the Pruned Exact Linear Time(PELT) Test in Multiple Changeoint Detection. *Am J Theor Appl Stat.* 2015;4(6):581-586. doi:10.11648/j.ajtas.20150406.30
103. Gamma A, Schleifer R, Weinmann W, Buadze A, Liebrezn M. Could Google Trends Be Used to Predict Methamphetamine-Related Crime? An Analysis of Search Volume Data in Switzerland, Germany, and Austria. *PLOS ONE.* 2016;11(11):e0166566. doi:10.1371/journal.pone.0166566
104. Wutchiett D, Durand C. Multilevel and time-series missing value imputation for combined survey and longitudinal context data. *Qual Quant.* 2022;56(3):1799-1828. doi:10.1007/s11135-021-01186-8
105. Monfared MAS, Ghandali R, Esmaili M. A new adaptive exponential smoothing method for non-stationary time series with level shifts. *J Ind Eng Int.* 2014;10(4):209-216. doi:10.1007/s40092-014-0075-5
106. Gardner ES. Exponential smoothing: The state of the art—Part II. *Int J Forecast.* 2006;22(4):637-666. doi:10.1016/j.ijforecast.2006.03.005