



## The use of web analytics combined with other data streams for tailoring online vaccine safety information at global level: The Vaccine Safety Net's web analytics project



Francesco Gesualdo<sup>a,1,\*</sup>, Francesco Marino<sup>a</sup>, Jas Mantero<sup>a</sup>, Andrea Spadoni<sup>a</sup>, Luigi Sambucini<sup>a</sup>, Giammarco Quaglia<sup>a</sup>, Caterina Rizzo<sup>a</sup>, Isabelle Sahinovic<sup>b</sup>, Patrick L.F. Zuber<sup>b</sup>, Alberto E. Tozzi<sup>a</sup>, the VSN-WAP study group

<sup>a</sup> Bambino Gesù Children's Hospital, Rome, Italy

<sup>b</sup> World Health Organization, Geneva, Switzerland

### ARTICLE INFO

#### Article history:

Received 3 December 2019

Received in revised form 22 July 2020

Accepted 31 July 2020

Available online 9 August 2020

#### Keywords:

Vaccine communication

Vaccine safety

Web analytics

News monitoring

Internet

### ABSTRACT

The Vaccine Safety Net's Web Analytics Project (VSN-WAP) was launched in October 2017 to monitor the behavior of users visiting websites belonging to the VSN, a global network of websites providing science-based information on vaccine safety. Participating websites could provide web metrics in two ways: through a Google Analytics (GA) script, which automatically forwarded metrics to a central account and through manual input (MI) of a reduced subset of metrics (Sessions, Page Views, New Users, Bounce Rate, Views/Session and Average Session Duration), which were pooled with the metrics obtained through GA. Additional metrics were obtained from websites providing data through Google Analytics (Country, Age, Sex, Device). We report results from February 2018 to March 2019. In March 2019, 32 websites were participating in the project (21 through GA, 11 through MI). From February 2018 to March 2019 we recorded 22,471,535 sessions, with 38,307,349 page views. Sessions, New Users and Page views progressively increased, Views/Session, Bounce Rate and Average Session Duration remained stable. Most users were female (68%) and belonged to the 25–34 age range (37%), followed by 35–44 (22%) and 18–24 (19%). Fifty-four percent of users connected from a mobile device, 42% from a desktop and 4% from a tablet.

Digital media monitoring techniques can provide insights on the characteristics of users with a specific interest in vaccines. These data can be exploited to improve the performance of websites providing information on vaccines to the general public.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Concern about vaccine safety is universal and directly impacts on vaccine hesitancy in upper middle-income and high-income countries [1,2], and more recently, has become a crucial driver for vaccine hesitancy in lower middle-income countries [3].

*Abbreviations:* GA, Google Analytics; SEO, Search Engine Optimization; VPD, vaccine preventable disease; VSN, Vaccine Safety Net; WAP, web analytics project; WHO, World Health Organization; AF, African WHO Region; EM, Eastern Mediterranean WHO Region; EU, European WHO Region; SEA, South East Asia WHO Region; PAHO, WHO Region of the Americas; WP, Western Pacific WHO Region.

\* Corresponding author.

E-mail address: [francesco.gesualdo@opbg.net](mailto:francesco.gesualdo@opbg.net) (F. Gesualdo).

<sup>1</sup> Address: Bambino Gesù Children's Hospital, Piazza S. Onofrio 4, 00165 Rome, Italy.

<https://doi.org/10.1016/j.vaccine.2020.07.070>

0264-410X/© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

In the digital world, the internet plays a pivotal role in the circulation of information (and misinformation) on vaccine safety. It has also opened a gateway for monitoring users' interest and behavior with the use of monitoring softwares - a strategy commonly adopted by for-profit companies to improve their revenues through increased public interest in their websites, social network profiles and products [4]. Health authorities, researchers and non-governmental organizations are beginning to use these monitoring tools for better understanding the vaccine hesitant and vaccine refusers [5]. Nevertheless, these tools have never been used to measure the success, reach and impact of websites offering evidence-based information on vaccine safety.

In 2003, the World Health Organization (WHO) established the Vaccine Safety Net (VSN): a network of diverse websites from a range of low-to-high income countries that provide science-

based information on vaccine safety in various languages. VSN candidates are evaluated based on the good information practices criteria [6] put forth by the Global Advisory Committee on Vaccine Safety, and membership is assigned based on credibility, content (quality and quantity), design and accessibility criteria.

The VSN Web Analytics Project (VSN-WAP) was subsequently launched in October 2017 [7], on the basis that VSN member websites would be monitored as a single, global information channel.

The VSN-WAP monitors: 1) the performance of the VSN through analyzing pooled web analytics; 2) the network's geographic outreach; 3) VSN users' profiles (at the global and local levels); 4) the public's interest towards vaccine preventable diseases (VPD); and 5) the visibility of the VSN on the Google search engine. Combining data about visitors' interest with other data sources (i.e., demographic data, occurrence of vaccine preventable diseases outbreaks, vaccine coverage, data from media monitoring) has provided the VSN with novel insights and a more complete picture of vaccine confidence within a global context, and has allowed to analyze and understand the successes and challenges of websites in conveying vaccine safety information on the Internet.

The aim of the present article is to describe the methods adopted in the VSN-WAP and to highlight insights on users' characteristics and behavior on VSN websites, as obtained through web analytic monitoring.

## 2. Methods

In May 2017, participation in the VSN-WAP was offered to all VSN members through an e-mail invitation, which also included details on the project's protocol. Web analytics monitoring started in October 2017. The first months of the projects were dedicated to setting up technological and logistic issues, and most of the websites that had agreed to participate in the project had set up the system by the end of January 2018. All results reported in the present article are related to the time period from February 2018 to March 2019.

### 2.1. VSN websites' metrics

Monitoring web analytics for participating VSN members was the core of the VSN-WAP. We obtained web metrics through 2 different methods: a) Google Analytics (GA) and b) manual input of a limited set of metrics on the project's website.

Google Analytics is a web tool allowing the automatic collection of web metrics. Websites using GA were provided with a customized code that allowed to automatically forwarding a large set of metrics to a central account. Metrics from different websites were aggregated.

The VSN members who were not using GA provided a reduced set of metrics through manual data input on the project's website:

- Sessions: a group of interactions one user takes within a given time frame on a website.
- Page Views: number of pages accessed by all users.
- New Users: number of first-time users during a selected date range.
- Bounce Rate: % of visitors who enter the website and then leave (in other words, % of single page views)
- Views/Session: mean number of pages viewed in a single session
- Average Session Duration: mean duration of all sessions.

Metrics obtained through GA and those obtained through manual input were pooled, to obtain metrics for the whole network participating in the project.

We report in the present article the trend for each of the metrics included in the reduced set, for the considered time period, taking into account two groups of websites: a) all websites participating in the project, independently from the month in which they started to provide metrics and b) only websites that contributed to the project during the entire study period. Thus, for the latter group of websites, metric variations were only attributable to a variation in the actual performance of the websites, and not to a variation in the number of websites included in the group.

### 2.2. Geographic outreach

The geographic outreach of the network, i.e. the countries from which visits to the participating websites originated, was obtained from the aggregated metric Session by country, representing the number of sessions from each single world country on the group of websites participating through the GA script. The number of sessions was normalized by 2 different variables: country population [8] and country Internet penetration [9]. Obtained data were used to create maps that were available on the project's website.

### 2.3. Users' profile

The users' profile was obtained by aggregating the following metrics: age group, sex and device. Age group and sex are calculated by Google through a proprietary, undisclosed algorithm, and are available only for a limited set of users, i.e. users that are logged-in with their personal Google accounts. In addition, for users that are not logged-in, Google is able to calculate these metrics through cookies or through other undisclosed techniques. The device used to navigate the websites (desktop, mobile or tablet) is recorded through a specific metric and is therefore available for all users accessing the VSN websites participating in the project through GA.

### 2.4. Ethical and privacy issues

Most data collected and analyzed in Google Analytics originate from first-party cookies, i.e. cookies obtained by websites monitored with Google Analytics [10]. Cookies are information on visitor's interactions on websites. Web users allow websites to collect and use this information when they agree to the website's cookie policy. Users may disable cookies or delete any individual cookie, or can disable measurements by Google Analytics for any site through a browser add-on. Metrics are displayed by the Google Analytics system in an aggregated way, and no information on single users is available. Information about the country of origin is automatically collected by Google through IP address and additional data from user's activity [11]. Moreover, all VSN websites comply with local privacy regulations, e.g. GDPR for Europe and HIPAA for the US.

The study was approved by the WHO Research Ethics Review Committee (Prot. EC.00028201).

### 2.5. Other web data

#### 2.5.1. Global interest in VPDs through Google Trends and Medisys

Global interest in VPDs was studied through Google Trends and Medisys.

We used Google Trends as a tool for monitoring the use of VPD-related keywords or search strings in Google Search. Google Trends allows search interest to be analyzed and mapped by search term or topic, through proprietary algorithms. Search volumes are normalized by time and location. Google Trends topic analysis takes into account different terms referring to the same concept, using automatic translation in different languages. Data are indexed to

100, where 100 is the maximum search interest for the selected keywords/topic, time period and location [12,13].

We mapped the frequency of searches for each VPD keyword by country. The VPD interest monitoring on Google Trends was performed on the first day of each month, selecting the “previous month” time frame on the Google Trends platform.

Obtained data were interpreted based on information derived from different sources, including news articles on specific VPDs in different countries, epidemics, and vaccination campaigns.

Data interpretation was facilitated by the use of MediSys, an open fully automatic web-based tool developed by the European Commission (EC) and the EC Joint Research Centre (EC JRC) to follow in almost real time online public health information [14].

### 2.5.2. VSN visibility on Google search

To ascertain the visibility of VSN websites on search engines with regards to searches on VPDs, we simulated searches on Google Search, with the following criteria:

- search performed on the specific Google search engine of the country of interest (e.g. google.co.uk, google.it, google.com.br, etc.)
- language setting based on the official language (or languages) of the country of interest
- incognito mode: this option allows to obtain results that are not influenced by the user's profile or search history. Although most users receive results based on their profile, incognito mode has allowed us to obtain a relatively reliable information on the probability that a user finds a VSN website on the first Google result page, when using keywords relative to VPDs

For each search, we recorded the number and rank of VSN websites on the first search result page.

Taking into account the way Google Trends and Google Search data were collected, we will present in the Results section an example of a case study exploring the interest about measles in Ukraine in February 2019 and the visibility of VSN websites on Google Search for searches on measles.

## 2.6. Web platform and monthly report

We developed a dedicated website with restricted access to participating members for manual data input and for customizable visualization of the project's pooled data. Moreover, we circulated a monthly report among VSN members, including a synthesis of the web monitoring data, together with additional comments and analyses. The monthly report also included suggestions for improving the VSN websites' search engine optimization (SEO, i.e. the visibility of each website on search engines), and two examples of the Google Trends and MediSys based analysis of users' interests.

## 3. Results

### 3.1. Metrics

The number of participating websites have varied through the study period (see Table 1). In February 2018, 58 websites from 30 countries were part of the VSN and a subset of 29 websites were participating in the VSN-WAP (16 through GA, 13 through manual data input). As of March 2019, the VSN network included 68 websites from 32 countries and 32 websites were participating in the project (21 through GA, 11 through manual data input). Of the above figures, 28 websites provided data during the whole study period (16 through GA, 12 through manual data input).

From February 2018 to March 2019 we recorded 22,471,535 sessions, with 38,307,349 page views. A total of 13,397,057 new users were recorded. Within this timeframe, the average session duration was 92.9, the average bounce rate was 68.25 and the average views/session was 1.8. All data on metrics derived from the pooling of GA metrics and those obtained through manual input.

Fig. 1 reports the trends of the metrics in the study period. Sessions, New Users and Page views show an increase during the study period, while Views/Session Bounce Rate and Average Session Duration remained stable. Nevertheless, in these graphs, observed variations are due to the instability of the number of websites participating in the study.

Fig. 2 reports the trends of the metrics obtained from websites providing monthly data for the entire study period. In this subgroup of websites, we observed a continuous increase in the metrics Sessions, New Users and Page views. Average Session Duration slightly decreased; Bounce Rate showed an overall increase, with a decreasing trend in the last couple of months; Views/Session showed different oscillations but remained around a value of 1.7 during the whole study period.

### 3.2. Geographic outreach

The geographic outreach of the VSN was analyzed based on data provided through GA. The VSN outreach was very high (data not shown). Considering the participating members as a single information channel, the network was accessed by all world countries, with few exceptions (Libya, Western Sahara, Cote d'Ivoire, South Sudan, Myanmar, Vietnam, Democratic People's Republic of Korea, Taiwan, Czech Republic, North Macedonia, Kosovo and French Guyana). The highest number of visits originated by New Zealand, Italy, Australia, Spain and Canada. Since the outreach was measured base only on a subset of VSN websites, we assume that the reported outreach underestimates the real outreach of the VSN.

### 3.3. Users' profile

Fig. 3 shows the age distribution of VSN users, calculated on 5,429,524 sessions (24% of the total sessions recorded in the study period, 45% of sessions recorded from GA). While the 25–34 age range is the most represented (37%), 18–24 and 35–44 are almost equally represented (19% and 22% respectively), followed by 45–54 (11%), 55–64 (7%), 65+ (4%). Regional differences are observed, with a higher prevalence of 25–34 in the African WHO Region (AF) (47%), in the Eastern Mediterranean WHO Region (EM) (47%) and the South East Asia WHO Region (SEA) (53%). The European WHO Region (EU) shows a lower percentage of 18–24 year-old users (11%) and a higher percentage of 35–44 (28%).

Fig. 4 shows the sex distribution by WHO Region, calculated on 5,806,689 sessions (26% of the total sessions recorded in the study period, 48% of sessions recorded from GA). At a global level, a predominance of female users is observed (68%). Most regions are in line with the global figures. A lower proportion of female users (who, though, are still the majority) is observed in SEA (62%), AF (60%) and EM (57%).

Fig. 5 shows the distribution of the device from which users connected to VSN websites, calculated on 12,044,003 sessions (54% of the total sessions recorded in the study period, 99% of sessions recorded from GA). Global figures show that slightly more than a half of the sessions originate from a mobile phone (6,440,391); 4,469,972 (42%) originate from a desktop, and a minor proportion (4%) from a tablet. Only in the West Pacific WHO Region (WP) did users mostly connect from a desktop (60%). A high predominance of connections from mobile devices is observed in SEA (76%) and in AF (66%).

**Table 1**  
List of websites participating in the VSN-WAP.

Country	Institution	URL	Months	GA/MI
Australia	National Centre for Immunisation Research and Surveillance of Vaccine Preventable Diseases (NCIRS)	<a href="http://www.ncirs.edu.au/">http://www.ncirs.edu.au/</a>	14	GA
Belgium	Network for Education and Support in Immunisation (NESI)	<a href="http://www.nesi.be/">http://www.nesi.be/</a>	3	GA
Belgium	HPV Prevention and Control Board	<a href="http://www.hpvboard.org">www.hpvboard.org</a>	14	MI
Belgium	Viral Hepatitis Prevention Board (VHPB)	<a href="http://www.vhpb.org/">http://www.vhpb.org/</a>	14	MI
Brazil	Brazilian Society of Immunizations (SBIm)	<a href="https://sbim.org.br/">https://sbim.org.br/</a>	9	GA
Brazil	Brazilian Society of Immunizations (SBIm)	<a href="http://www.familia.sbim.org.br">www.familia.sbim.org.br</a>	9	GA
Canada	Immunize Canada	<a href="https://immunize.ca/">https://immunize.ca/</a>	7	GA
Canada	I Boost Immunity	<a href="https://iboostimmunity.com/">https://iboostimmunity.com/</a>	14	GA
Canada	Kids Boost Immunity	<a href="https://kidsboostimmunity.com/">https://kidsboostimmunity.com/</a>	14	GA
Canada	Vaccines411	<a href="https://vaccines411.ca/">https://vaccines411.ca/</a>	14	MI
European Union	Vaccines Today	<a href="http://www.vaccinestoday.eu/">http://www.vaccinestoday.eu/</a>	14	GA
France	Infovac-France	<a href="https://www.infovac.fr/">https://www.infovac.fr/</a>	14	GA
France	MesVaccins.Net	<a href="https://www.mesvaccins.net">https://www.mesvaccins.net</a>	14	MI
Germany	Impfbrief.de	<a href="http://www.impfbrief.de/">http://www.impfbrief.de/</a>	5	MI
Germany	Impfkontrolle.de	<a href="http://www.impfkontrolle.de/">http://www.impfkontrolle.de/</a>	14	GA
Hungary	VACSATC - Oltásbiztonság	<a href="http://www.oltasbiztonsag.hu/">http://www.oltasbiztonsag.hu/</a> ; <a href="http://www.vacsatc.hu">www.vacsatc.hu</a>	14	GA
Ireland	HSE National Immunisation Office Website, Ireland	<a href="http://www.immunisation.ie/">http://www.immunisation.ie/</a>	14	MI
Italy	Io Vaccino	<a href="http://www.iovaccino.it/">http://www.iovaccino.it/</a>	14	GA
Italy	VaccinarSi	<a href="http://www.vaccinarsi.org/">http://www.vaccinarsi.org/</a>	14	GA
Italy	EpiCentro	<a href="http://www.epicentro.iss.it/">http://www.epicentro.iss.it/</a>	14	GA
New Zealand	The Immunisation Advisory Centre (IMAC)	<a href="http://www.immune.org.nz/">http://www.immune.org.nz/</a>	14	GA
South Africa	Vaccines for Africa (VACFA)	<a href="http://www.vacfa.uct.ac.za/">http://www.vacfa.uct.ac.za/</a>	14	GA
Spain	Asociación Española de Vacunología (AEV)	<a href="http://www.vacunas.org/">http://www.vacunas.org/</a>	14	GA
Spain	Vacunas. Murciasalud	<a href="http://www.murciasalud.es/vacunaciones">http://www.murciasalud.es/vacunaciones</a>	14	MI
Switzerland	Global Advisory Committee on Vaccine Safety (GACVS)	<a href="http://www.who.int/vaccine_safety/committee/en/">http://www.who.int/vaccine_safety/committee/en/</a>	14	GA
The Netherlands	Rijksvaccinatieprogramma (RVP)	<a href="https://rijksvaccinatieprogramma.nl">https://rijksvaccinatieprogramma.nl</a>	14	MI
UK	Vaccine Knowledge Project	<a href="http://vk.ovg.ox.ac.uk/">http://vk.ovg.ox.ac.uk/</a>	14	GA
UK	Immunisation Scotland (of the NHS in Scotland)	<a href="http://nhsinform.scot/immunisation">http://nhsinform.scot/immunisation</a>	14	MI
UK	Public Health England - Immunisation section	<a href="https://www.gov.uk/government/collections/immunisation">https://www.gov.uk/government/collections/immunisation</a>	14	MI
US	PATH's Vaccine Resource Library	<a href="http://www.vaccineresources.org">www.vaccineresources.org</a>	14	MI
US	The History of Vaccines - La Historia de lasVacunas	<a href="http://www.historyofvaccines.org/">http://www.historyofvaccines.org/</a>	14	GA
US	Vaccine Education Center (VEC), Children's Hospital of Philadelphia	<a href="http://www.vaccine.chop.edu/">http://www.vaccine.chop.edu/</a>	14	MI
US	Institute for Vaccine Safety, Johns Hopkins Bloomberg School of Public Health	<a href="http://www.vaccinesafety.edu/">http://www.vaccinesafety.edu/</a>	14	MI

GA = Google Analytics; MI = manual data input. Months during which the websites participated are intended with respect to the period February 2018-March 2019.

### 3.4. Measles in Ukraine: A case study monitoring global interest in a VPD and VSN visibility on Google search

A specific case study, presented in the February 2019 report, was put together to give insights on the interest regarding measles in Ukraine and on the visibility of VSN on Google Search regarding measles.

Initially, we found through Google Trends a significant interest for the topic “measles” in February 2019 in Ukraine (76 out of a 100 score, based on the Google Trends algorithm, where 100 is the maximum search interest for any VPD in February 2019 at the global level). Then, we explored through MediSys possible explanations for this specific interest.

We filtered all Ukrainian articles triggered by MediSys during February 2019 for the topic “measles” and recorded an unusually high level of news volume in comparison to previous months. Subsequently, we analyzed the contents of the articles to identify potential events and/or debates occurred in the country in relation to the ongoing measles outbreak, which might have caused the increase of interest. We then complemented the information collected with more details extracted from relevant sources of information, as online statements and data published by local health authorities.

Finally, we simulated a search for the term “measles” on the Ukrainian Google Search engine in order to measure the visibility of VSN websites (i.e. the presence of VSN websites in the first result page).

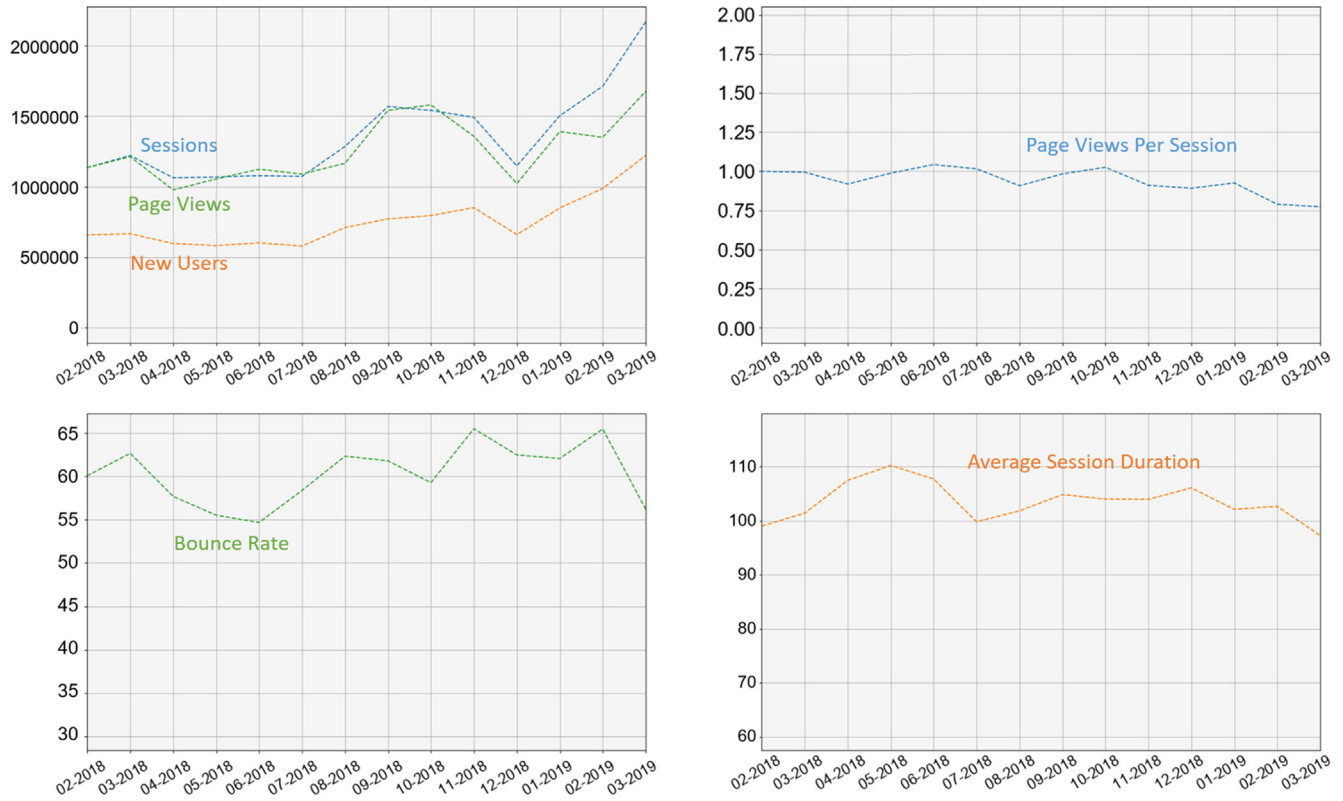
A summary of the findings related to this analysis is reported in Supplementary Material 1.

## 4. Discussion

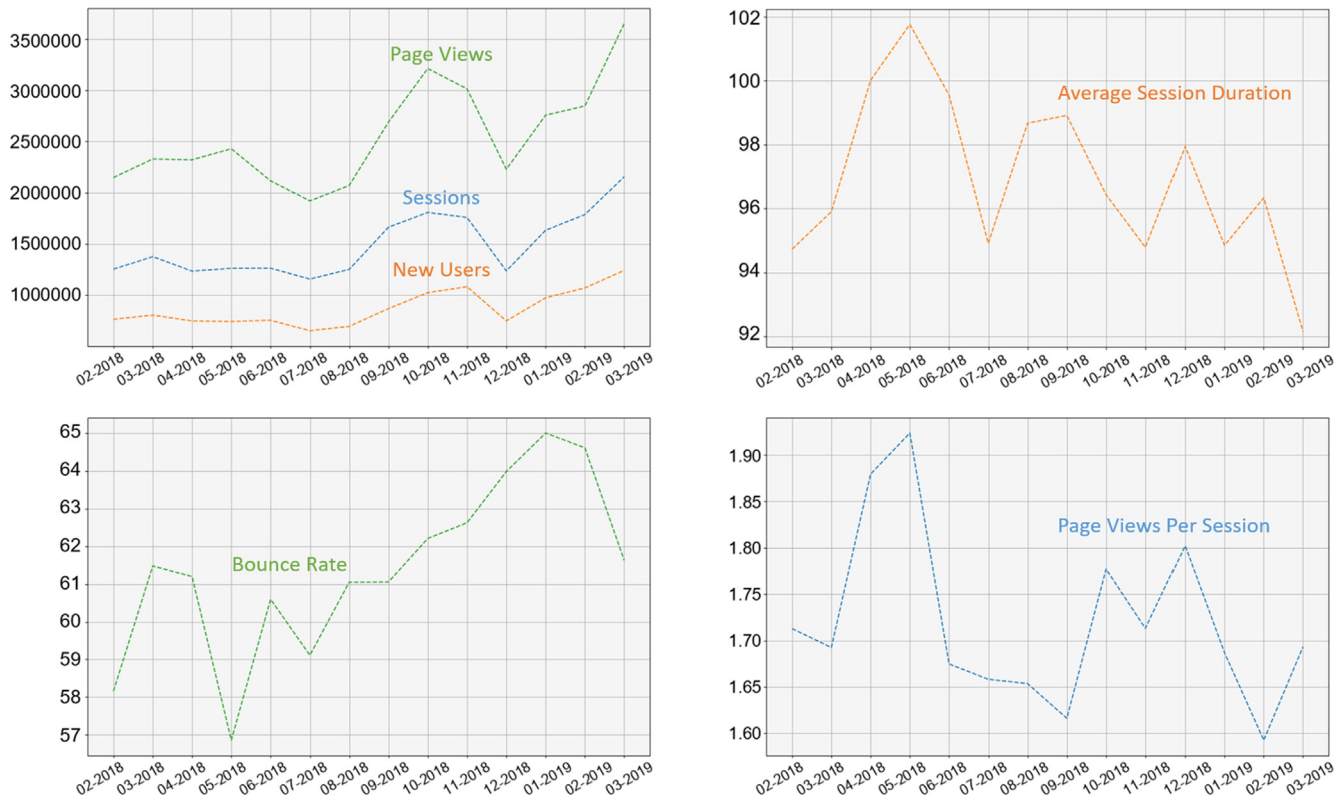
We present a multifaceted system aimed at monitoring user behavior on a network of websites dedicated to vaccine safety. Through the VSN-WAP, it was possible to gain insights on general interest towards the VSN network (through the metrics Sessions and Page Views), the ability of the network to attract new visitors (New Users), and the ability of the network to satisfy users' information needs (Bounce Rate, Average Session Duration and Views/Sessions).

Given the relative diversity of participating websites in terms of scope, structure and target audience, it is not possible to comment on single values of pooled analytic metrics, or to compare results to “normal” values. On the other hand, the variation of metric trends during time, in particular for the subgroup of websites providing data in the entire study period, can provide interesting insights on the performance of the network. We observed a progressive increase in Sessions, Page Views and New Users. The growth in new visitors could be indicative of a progressive growth of the public interest towards the network's websites. It is not possible to understand if this improvement is entirely due to the activity of the WAP, although, based on the results of an internal survey, almost half of the participating members stated that the VSN-WAP had a direct impact on their activities, especially on SEO (data





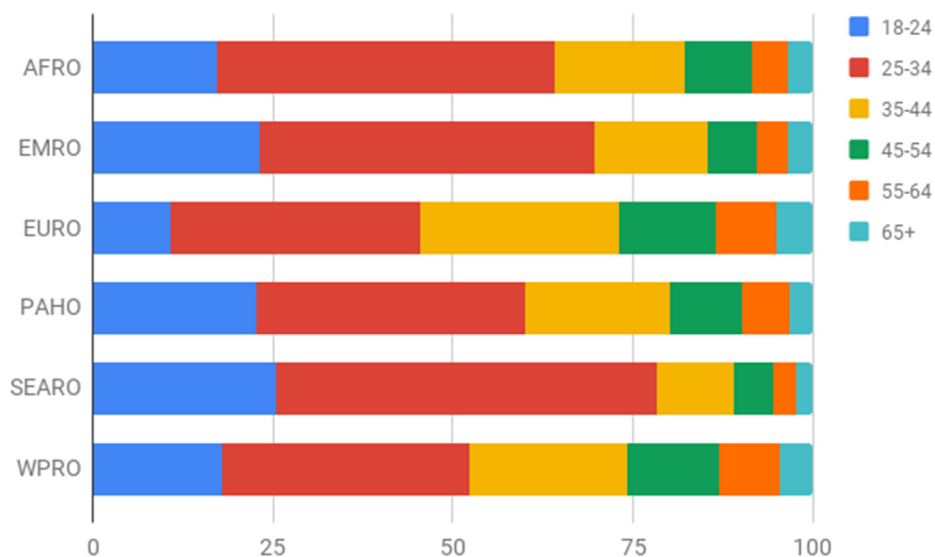
**Fig. 1.** Trend of the monitored web metrics from February 2018 to March 2019. All metrics from all participating websites were included in the trend, irrespective of the participation period.



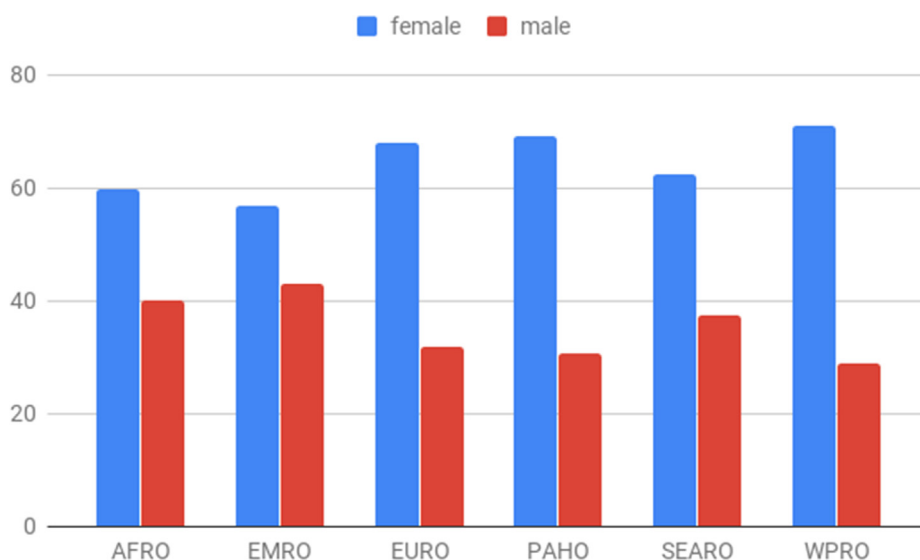
**Fig. 2.** Trend of the monitored web metrics from February 2018 to March 2019 for websites providing data during the whole study period.

not published). While Average Session Duration remained constant during the study period, Bounce Rate (i.e. the percentage of users visiting just one page of the website) and Views/Session slightly

worsened. This is an expected trend when the number of sessions increases [15]: typically, when the audience of a website grows (usually due to an improvement of the SEO), it is expected that a



**Fig. 3.** Age distribution of VSN users by WHO Region. Data obtained from websites providing data through Google Analytics. The total sessions on which Google Analytics was able to infer the age range were as follows: AF 74,556 users, EM 58,912 users, EU 1,729,431 users, PAHO 2,904,153 users, SEA 231,418 users, WP 431,054 users.



**Fig. 4.** Sex distribution of VSN users by WHO Region. Data obtained from websites providing data through Google Analytics. The total sessions on which Google Analytics was able to infer the sex were as follows: AF 76,914 users, EM 68,607 users, EU 1,784,160 users, PAHO 3,183,917 users, SEA 237,300 users, WP 461,791 users.

large proportion of the new users will not show a high level of interest towards the website, but will rather visit the website for satisfying a single, limited information need.

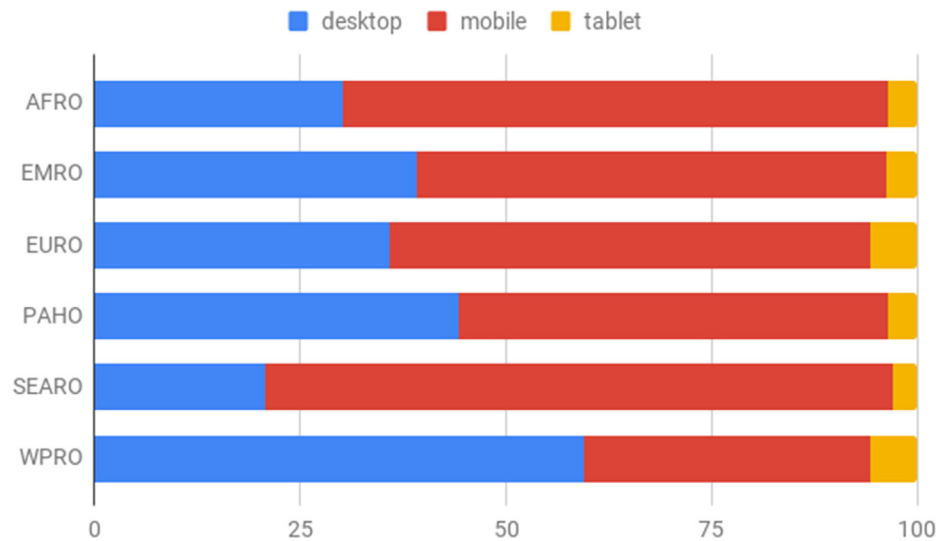
The VSN network is accessed by almost every world country. Nevertheless, these results underestimate the real coverage of the network, as the system does not take into account the origin of visits to websites participating in the VSN-WAP with manual data, or those not participating in the project. Countries from which the highest number of visits originated are those where the most active websites were based (e.g. New Zealand), or countries that had more than one participating website (e.g. Italy). The lower outreach of the network in specific countries or world areas should be better investigated, and strategies to improve the outreach should be implemented, eg. focusing on language or scouting for websites active in those countries but not included in the VSN network.

The average VSN user is a female of reproductive age (25–34 years). This confirms previous observation showing that, in

some countries, users searching for online information on health are mainly female [16–19]. Moreover, given that a large number of Google searches for vaccines are related to childhood vaccinations [20], we could hypothesize that most visitors might be mothers of young children. The proportion of males accessing the network was higher in AF and EM.

Regional differences with regards to age distribution are observed. The variability of interest with age is probably influenced by the age distribution in different areas. Moreover, based on the hypothesis that most VSN visitors are mothers of young children, these differences could be explained by the mean age of women at childbirth across different world regions. In 2016, mean age at childbirth in European countries was 30.6 years [21]; in US, in 2014, it was slightly lower (26.3 years) [22], and even lower in Latin American countries [23].

We observed a frequent use of mobile phones for searching for information on vaccines, confirming previous observations [24]. This result strengthens the indication to give priority to mobile-



**Fig. 5.** Device distribution of VSN users by WHO Region. Data obtained from websites providing data through Google Analytics. The total sessions on which Google Analytics was able to infer the device were as follows: AF 151,526 users, EM 128,744 users, EU 3,610,009 users, PAHO 6,581,945 users, SEA 421,520 users, WP 1,150,259 users.

friendly websites, also taking into account that Google gives a higher rank to websites with a mobile optimization [25].

Access to the internet is heterogeneous and dependent upon sociodemographic factors. Smartphone use varies by age - young adults have higher use than older adults and young children; and by gender - males have high internet usage in all regions globally, other than the Americas [26]. A rural-urban divide is present in some contexts [27], and, in settings where smartphone penetration is low [28], it is likely that sharing of devices in families and communities will occur. If the primary handset user was not the website user, it is possible that user demographics were incorrectly allocated. Sociocultural factors, particularly male-female dynamics within the family unit could affect this allocation. This could explain the higher representation of males in the AF and EM regions.

Monitoring the performance of a network of websites dedicated to vaccine safety holds potential in several areas.

Firstly, assessing the performance of websites with different characteristics but with a similar aim can allow improvements on each website, based on the experience of the whole network, in terms of ability to attract new visitors, and to retain a high level of interest, as indicated by returning users.

Secondly, data on the geographic outreach of the network, users' profiles and devices could be used to inform the development of tailored communication strategies. These could target groups that are particularly sensitive to vaccine safety issues (e.g. 25–34 year-old mothers), or segments of the population that have a lower interest or access to vaccine information online and therefore require more attention (e.g., elderly individuals, who are the target of various vaccination campaigns). Tailoring communication to regional characteristics and needs can also allow issues that are specific to different geographic areas to be effectively addressed.

Thirdly, as reported in the case study about measles in Ukraine, we assessed the performance of each website and of the whole network on search engines. Concurrently, we were able to analyse different websites to understand their level of SEO optimization. This kind of assessment has the potential to improve the visibility of the network's websites on search engines, and therefore to expand the role of the VSN as a reference on vaccine safety information at a global level.

A further potential function of the monitoring system is assessing the performance of future global information campaigns, coordinated across the network's websites.

Coordinated campaigns have the potential to reinforce messages on vaccine safety across the network, thus allowing VSN to deliver evidence-based information on vaccines with an unprecedented strength.

Monitoring metrics for single pages of each website is a function that will be implemented in the future. This will allow the tracking of users' interest towards specific topics, e.g. VPDs or vaccine safety issues. Mapping users' interests in different world areas would enable the network to have a deeper understanding of regional and global information needs, and therefore better tailor information campaigns.

Integrating new data sources, exploiting the monitoring of topic-specific metrics, expanding the detection of users' interests on search engines and exploring the information circulating on social networks will allow to finely tuning communication activities to users' information needs.

This study has a number of limitations.

First, only a portion of VSN websites participated in the VSN-WAP, therefore, as previously discussed, reported data may overestimate or underestimate the actual figures related to the whole network. Nevertheless, the adopted methods were transparent and can be easily applied to the whole group of websites in the future.

Secondly, demographic and geographic data are collected through Google's proprietary, undisclosed algorithms. Only in a few studies, geographic and demographic data obtained through Google services have been used for research [29,30]. In a recent article, Tschantz et al. showed that accuracy of the Google algorithm to correctly characterise age and sex was 67% and 66% on average, respectively [31]. Even if the accuracy of this method is far from being perfect, this kind of estimation deserves further investigation, as it could allow inferring age and sex for large groups of web users with an acceptable approximation. This characteristic makes it a promising possibility for research, also taking into account that Google continuously improves the algorithms in order to achieve a better performance.

Furthermore, both demographic and geographic characteristics are based on the number of sessions rather than on the number of users. We chose to display session-based data in order to include data obtained through manual input in the denominator of some of the reported figures - the metric Users was not listed among those collected through manual input. Nevertheless, the average

number of Sessions by Users, for users whose demographic characterization was available, was relatively low (between 1.2 and 1.3), therefore we hypothesize that sessions-based calculation are a good approximation of user-based ones.

Finally, even if, based on the project's data, advices for improving the SEO were delivered to participating websites through different channels (a brief SEO guideline for the network, video-calls with SEO experts), the intervention was not conducted systematically and therefore we cannot assess its impact on the websites' performance.

In conclusion, the VSN-WAP project represents the first study in which web analytic metrics have been exploited and integrated with other web-based data to understand and inform how to increase the performance of a global network of websites dedicated to vaccine safety. Applying digital media monitoring techniques to websites providing health-related content has the potential to provide insights on the characteristics of users with a specific interest in vaccines. Further research should focus on the development of a framework to facilitate the use of this data to inform and tailor vaccine communication, and to improve the experience of users on websites dedicated to vaccines. Furthermore, the impact of these activities on vaccine confidence and, ultimately, on vaccine uptake, deserves a deeper assessment.

## Funding

The VSN-WAP was supported by funds provided by GAVI and CDC.

## CRedit authorship contribution statement

**Francesco Gesualdo:** Conceptualization, Methodology, Writing - original draft, Supervision. **Francesco Marino:** Methodology. **Jas Mantero:** Writing - original draft. **Andrea Spadoni:** Software. **Luigi Sambucini:** Software. **Giammarco Quaglia:** Formal analysis. **Caterina Rizzo:** Writing - review & editing. **Isabelle Sahinovich:** Funding acquisition, Writing - review & editing. **Patrick LF Zuber:** Funding acquisition, Writing - review & editing. **Alberto E Tozzi:** Conceptualization, Methodology, Project administration, Writing - review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The VSN-WAP Study Group consists of the following members, all of whom gave an important contribution to this study: Veronica Bartolucci (Bambino Gesù Children's Hospital, Italy), Lucie Marisa Bucci (Immunize Canada, Canada), Diamantis Klimentidis (Clinical Pharmacist, Greece), Eleonora Grassucci (Bambino Gesù Children's Hospital, Italy), Smaragda Lamprianou (World Health Organization), Elisabetta Pandolfi (Bambino Gesù Children's Hospital, Italy), Tina Purnat (World Health Organization), Simonetta Zezza (Bambino Gesù Children's Hospital, Italy).

## References

- [1] Marti M, de Cola M, MacDonald NE, Dumolard L, Duclos P. Assessments of global drivers of vaccine hesitancy in 2014—Looking beyond safety concerns. *PLoS ONE* 2017;12:. <https://doi.org/10.1371/journal.pone.0172310>
- [2] Lane S, MacDonald NE, Marti M, Dumolard L. Vaccine hesitancy around the globe: analysis of three years of WHO/UNICEF Joint Reporting Form data-2015–2017. *Vaccine* 2018;36:3861–7. <https://doi.org/10.1016/j.vaccine.2018.03.063>.
- [3] Kochhar S, Rath B, Seeber LD, Rundblad G, Khamesipour A, Ali M, et al. Introducing new vaccines in developing countries. *Expert Rev Vaccines* 2013;12:1465–78. <https://doi.org/10.1586/14760584.2013.855612>.
- [4] Chaffey D, Patron M. From web analytics to digital marketing optimization: increasing the commercial value of digital analytics. *J Direct Data Digit Mark Pract* 2012;14:30–45.
- [5] Koukaras P, Rousidis D, Tjortjis C. Forecasting and prevention mechanisms using social media in health care. In: Maglogiannis I, Brahnam S, Jain L, editors. *Advanced computational intelligence in healthcare-7. Studies in computational intelligence*. Berlin: Springer; 2020. p. 121–37.
- [6] Join VSN. Vaccine Safety Network 2017. <https://www.vaccinesafetynet.org/vsn/join-vsn> [accessed July 13, 2020].
- [7] World Health Organization. Weekly epidemiological record Relevé épidémiologique hebdomadaire. Global Advisory Committee on Vaccine Safety, 30 November – 1 December 2016. World Health Organization; 2017. N. 2, 2017, 92, 13–20.
- [8] United Nations, Department of Economic and Social Affairs, Population Division (2019). <https://population.un.org/wpp/Download/Standard/Population/> [accessed July 13, 2020].
- [9] Internet Usage & Social Media Statistics. Internet Live Stats Web Site. <https://www.internetlivestats.com/> [accessed July 13, 2020].
- [10] Analytics Help - Safeguarding your data. <https://support.google.com/analytics/answer/6004245?hl=en> [accessed July 13, 2020].
- [11] GooglePrivacy & Terms - How Goggle uses location information. <https://policies.google.com/technologies/location-data?hl=en-US> [accessed July 13, 2020].
- [12] Trends Help. FAQ about Google Trends data. [https://support.google.com/trends/answer/4365533?hl=en-GB&ref\\_topic=6248052](https://support.google.com/trends/answer/4365533?hl=en-GB&ref_topic=6248052) [accessed July 13, 2020].
- [13] Mavragani A, Ochoa G. Google trends in infodemiology and infoveillance: methodology framework. *JMIR Public Health Surveill* 2019;5(2):. <https://doi.org/10.2196/13439>.
- [14] MediSys <https://medisys.newsbrief.eu/medisys/homeedition/it/home.html> [accessed July 13, 2020].
- [15] Bounce rate - Analytics Help <https://support.google.com/analytics/answer/1009409?hl=en> [accessed July 13, 2020].
- [16] Umefjord G, Sandström H, Malker H, Petersson G. Medical text-based consultations on the Internet: a 4-year study. *Int J Med Inform* 2008;77(2):114–21. <https://doi.org/10.1016/j.ijmedinf.2007.01.009>.
- [17] Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. *J Med Internet Res* 2011;13(1):. <https://doi.org/10.2196/jmir.1600>.
- [18] Beck F, Richard JB, Nguyen-Thanh V, Montagni I, Parizot I, Renahy E. Use of the internet as a health information resource among French young adults: results from a nationally representative survey. *J Med Internet Res* 2014;16(5):e128.
- [19] Pew Researcher Center Science 2015. Science and Health Interest Linked to Americans' Gender, Age and Personality. <https://www.pewresearch.org/science/2015/12/11/public-interest-in-science-and-health-linked-to-gender-age-and-personality/> [accessed August 2, 2019].
- [20] Bragazzi NL, Barberis I, Rosselli R, Gianfredi V, Nucci D, Moretti M, et al. How often people google for vaccination: qualitative and quantitative insights from a systematic search of the web-based activities using Google Trends. *Hum Vaccines Immunother* 2017;13:464–9. <https://doi.org/10.1080/21645515.2017.1264742>.
- [21] Mean age of women at childbirth across EU regions. Eurostat Web Site. <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20180323-1?inheritRedirect=true> [accessed July 13, 2020].
- [22] Matthews TJ, Hamilton BE. Mean age of mothers is on the rise: United States, 2000–2014. *NCHS Data Brief* 2016;232:1–8.
- [23] OECD - Social Policy Division - Directorate of Employment, Labour and Social Affairs. [https://www.oecd.org/els/soc/SF\\_2\\_3\\_Age\\_mothers\\_childbirth.pdf](https://www.oecd.org/els/soc/SF_2_3_Age_mothers_childbirth.pdf) [accessed August 2, 2019].
- [24] Digital 2019: Global Internet Use Accelerates. We Are Social Web Site. <https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates> [accessed August 2, 2019].
- [25] Google Search. Mobile-first indexing best practices. <https://developers.google.com/search/mobile-sites/mobile-first-indexing?hl=en> [accessed July 13, 2020].
- [26] Global internet usage rate by gender and region 2017. Statista Web Site. <https://www.statista.com/statistics/491387/gender-distribution-of-internet-users-region/> [accessed July 13, 2020].
- [27] Bhattacharya A. Rural Indian women are lagging far behind their urban counterparts in mobile phone usage. Quartz Web Site. <https://qz.com/india/969565/rural-indian-women-are-lagging-far-behind-their-urban-counterparts-in-mobile-phone-usage/> [accessed July 13, 2020].
- [28] Pew Researcher Center Global Attitudes Project 2018. Internet Seen as Having Positive Impact in Sub-Saharan Africa. <https://www.pewresearch.org/global/2018/10/09/internet-connectivity-seen-as-having-positive-impact-on-life-in-sub-saharan-africa/> [accessed July 13, 2020].
- [29] Davidson B, Alotaibi NM, Hendricks BK, Cohen-Gadol AA. Popularity of online multimedia educational resources in neurosurgery: insights from the neurosurgical atlas project. *J Surg Educ* 2018;75(6):1615–23. <https://doi.org/10.1016/j.jsurg.2018.05.001>.



- [30] Utunen H, Ndiaye N, Piroux C, George R, Attias M, Gamhewage G. Global reach of an online COVID-19 course in multiple languages on OpenWHO in the first quarter of 2020: analysis of platform use data. *J Med Internet Res* 2020;22(4):. <https://doi.org/10.2196/19076>e19076.
- [31] Tschantz MC, Egelman S, Choi J, Weaver N, Friedland G. The accuracy of the demographic inferences shown on Google's Ad Settings. In: *Proceedings of the 2018 workshop on privacy in the electronic society*; 2018. p. 33–41.