

# Vaccine Sentimeter

## Global Monitoring of Vaccination Conversations

User Manual Version 2.0  
November 2015



Vaccine Sentimeter provides real-time surveillance and trend analyses of vaccination conversations in mainstream and social media. This monitoring is visualized in a web-based dashboard that displays volume of the conversation and the channels, sentiments, influencers, and networks. Vaccine Sentimeter pulls intelligence from 100,000+ online sources including social media, news aggregators, blogs, eyewitness reports, expert-curated discussions and validated official reports. It thereby achieves a unified and comprehensive view of the current global conversation on vaccination, allowing those who manage vaccination programs to listen, understand and engage with the public to build and sustain confidence in vaccination.

### Table of Contents

Objectives: ..... 2  
Dashboard Functionalities: ..... 2  
Data Collection:..... 5  
Automated Classification: ..... 5  
Data Curation: ..... 6  
Monthly Intelligence Reports: ..... 6  
Thank You..... 6


## Objectives:


Vaccine Sentimeter was developed by Epidemico, ProMED-mail, and Sanofi Pasteur as a practical tool for those who manage vaccination programs and public health communications to listen, understand, and engage with the public and media to build and sustain public confidence in vaccination. This tool captures data from various channels and hubs of influences such as online news, blogs, social media, and other informal and unofficial sources of information. By categorizing reports as positive, neutral, or negative sentiment; offering various visualizations both temporally and geographically; and analyzing trends and high risk alerts, Vaccine Sentimeter facilitates data-based listening, understanding, and engaging that is needed to build or sustain confidence in vaccination programs.

The partners envision Vaccine Sentimeter to become a routine tool for vaccine program coordinators and public health professionals to track and understand the local conversation around vaccination and vaccine-preventable diseases. Timely detection of shifts in the conversation will enable proactive intervention to sustain or build public confidence in immunization, maintain vaccination coverage rates, and prevent disease spread and outbreaks. The project also aims to help identify hubs of influence (people, blogs, websites) that drive opinions and perceptions, with hopes of engaging these influencers in public health campaigns.


Vaccine Sentimeter is an effort in progress, adding data sources, supported languages, sentiment categories, visualizations, and other features that would be of practical application to users. Partners welcome collaboration and sponsorship as well as user feedback and ideas to fulfill the vision of developing a practical open-access tool adopted across the globe to manage vaccination programs. Due to budget constraints, real-time processing and visualization was stopped in December 2014. However, data collection is still ongoing in real-time and the processing and visualization can be continued with additional support.

## Dashboard Functionalities:

1. Vaccination-related online articles can be seen on a map or list view (Figure 1). On the default map view, individual reports are colored by sentiment, and countries are also colored by percentage of negative reports (click  on top right for legend). Reports can be clicked to show the date, title, and link to the original article.

The list view can be accessed by clicking  on the top right. In the list view, reports are labeled with date, title, vaccine, category, sentiment, and location. The list can be sorted by any of these parameters. The filter function on the list view allows free text searches within

## Vaccine Sentimeter: Global Monitoring of Vaccination Conversations

the list. The list can be converted into a csv file using the export function (  icon on top right, login required).




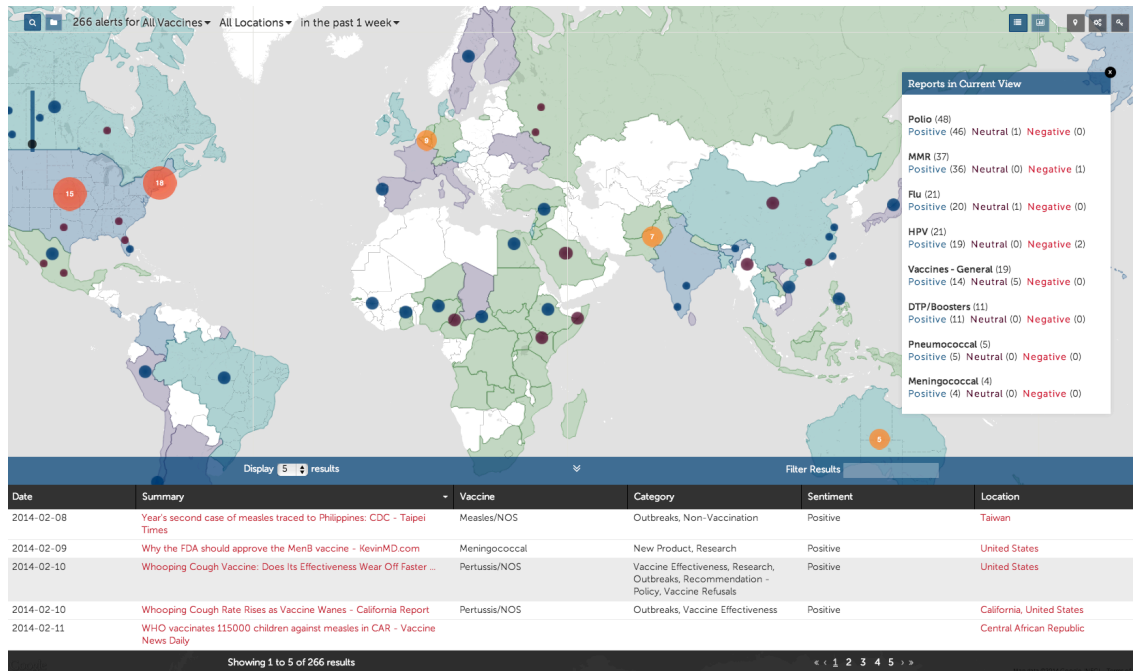
The search box (  ) on top left allows users to search by date, vaccine, category, and location, and to save searches (  , login required). The summary box shows a ranked list of vaccines by mentions and sentiment; this box can be turned on and off by clicking  . The map, list, and summary box adjust according to the search results.

Figure 1.

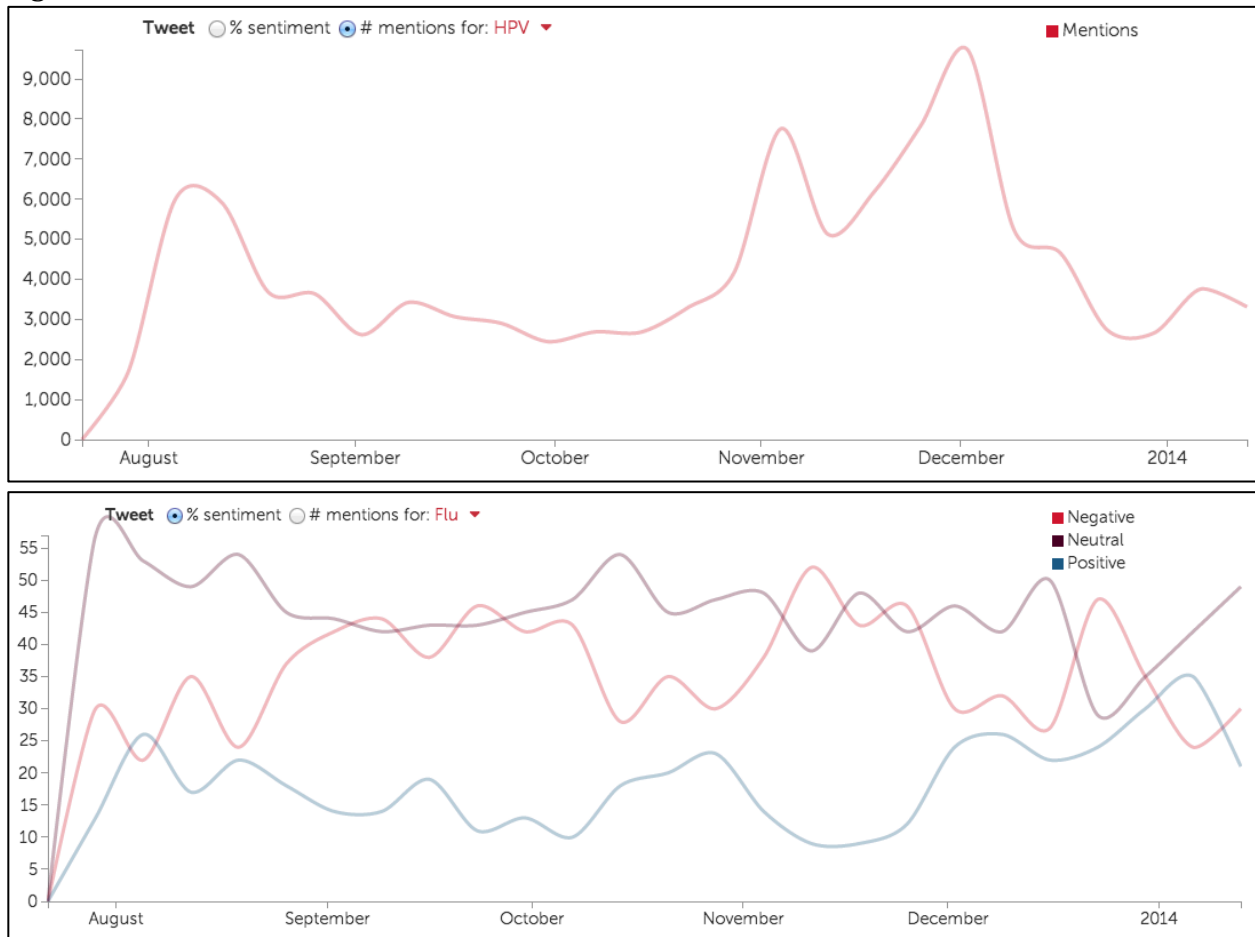


2. High-volume social media chatter on Twitter can be visualized graphically (Figure 2). The user can select the vaccine of interest then toggle between the overall volume of conversations over time (number of tweets mentioning selected vaccine) or fluctuations in sentiment (percentage of positive, neutral, and negative sentiment).

This visualization allows users to identify trends and signals in conversation volume and sentiment for particular vaccines (or vaccination in general). To understand what is driving these trends, the user can search for the time and topic of interest or peruse the monthly reports to see what media may have been the trigger. The original Twitter data and in-depth analyses can also be made available – please contact [info@epidemico.com](mailto:info@epidemico.com).

## Vaccine Sentimeter: Global Monitoring of Vaccination Conversations

Figure 2.



3. Monthly reports on trends and specific reports of interest are available. These reports summarize key metrics including mentions by location, categories, and vaccine, along with analysis of trends & high-risk signals. These can be accessed by clicking [Monthly Reports](#) on the bottom of the dashboard.

4. Additional functionalities to be added include but are not limited to:

- New visualizations of news and social media data from other channels
- New sentiment categories
- Improved search and navigation function
- Additional sources and languages

Feedback and suggestions are always welcome through a 5-minute survey (<http://bit.ly/Ne8ZeG>) or through email at [info@epidemico.com](mailto:info@epidemico.com). Thank you in advance! Please read on to learn more about our data collection, categorization, and curation processes.

## Vaccine Sentimeter: Global Monitoring of Vaccination Conversations

### Data Collection:

Vaccine Sentimeter is powered by the HealthMap platform ([www.healthmap.org](http://www.healthmap.org)), which collects data from over 100,000 sources including online news aggregators, ProMED-mail, expert-curated discussions, validated official reports, blogs, and mailing lists. Every hour, these sources are searched using vaccine-specific taxonomy in English, Spanish, and French. The taxonomy is composed of inclusion words spanning generic and brand names of vaccines, vaccine-related terminology (e.g. adjuvant, adverse event), and commonly used vernaculars for vaccines (e.g. jab, shot). There are exclusion words to filter out animal vaccines, financial reports regarding vaccine products, and other sources of noise.

The social media data of Vaccine Sentimeter is derived from the public Twitter API, searched real-time using vaccine-specific taxonomy. The search terms are generic and brand names of vaccines as well as commonly used terms for vaccination, in English.

### Automated Classification:

Through natural languages processing and machine learning algorithms, reports are classified to the appropriate vaccine, category, sentiment, and location. The list of vaccines, categories, and sentiments are shown in Table 1. Locations are geo-located onto a map at the country, province, or local level, depending on availability of information. Duplicates are grouped together based on a similarity score and unique reports are displayed on the dashboard.

Table 1.

Tag	Labels
<b>Vaccine</b>	Adenovirus, BCG, DTap, TDaP, DTPa, DTP, DPT, DTwP, IPV, MMR, MMR II, MMRV, OPV, Pentavalent, Hexavalent, Td, HPV, Pneumococcal, Influenza, H1N1, Rotavirus, Meningococcal, HAV, HBV, HEV, Hib, Shingles, Varicella, Anthrax, Diphtheria/NOS, Measles/NOS, Mumps/NOS, Pertussis/NOS, Polio/NOS, Rubella/NOS, Tetanus/NOS, Cholera, Yellow Fever, Japanese encephalitis, Rabies, Typhoid, Vaccines in general/All vaccines
<b>Category</b>	Additives, Adverse Event Following Immunisation (Autism, Death, General, HIV/AIDS, Immune System, Mild, Miscarriage, Narcolepsy, Neurological damage, Paralysis, Fertility, Fever, Seizures), Beliefs (Conspiracy theory, Philosophical, Religious, Socio-Cultural), Conflict/War, Contamination, Cost, Disease Burden, Drop Outs, High Profile Individual, Lack of Awareness/Information, Marginalised Populations, Mass Campaign, Motives (Business, Political), New Product, Outbreaks, Programmatic Error/Issue/Conflict, Public Demonstration, Recommendation – Policy, Recommendation (by HCW, by Health Organisation, by journalist, by lay person, Recommendation by parent, by religious leader), Research, Risk/Benefit, Strategy (Delivery, Funding/Fundraising, Outreach, Public Engagement, Investigation), Supply, Tampering, Vaccine Recall, Vaccine Refusals, Vaccine Schedule, Vaccine Suspension, Non-vaccination, Fertility, Low Vaccination Coverage, Vaccine Effectiveness

## Vaccine Sentimeter: Global Monitoring of Vaccination Conversations

<b>Sentiment</b>	Positive, Neutral/Unclear, Negative
------------------	-------------------------------------

Similarly, social media data is tagged with vaccine mentioned through natural language processing. Machine learning algorithms then classify tweets according to sentiment. Tags for social media data are simplified as shown below as information is often limited in any single social media post.

Table 2.

Tag	Labels
<b>Vaccine</b>	Chicken pox, DTP/Boosters, Endemic vaccines (Yellow Fever, Japanese encephalitis, Rabies, Typhoid, Cholera), Flu, Hepatitis, HPV, Meningococcal, Measles/Mumps/Rubella, Pneumococcal, Polio, Vaccines in general/All vaccines
<b>Sentiment</b>	Positive, Neutral/Unclear, Negative

### Data Curation:

Once the web and social media data are collected and automatically classified by the HealthMap system, human curators review the data to correct any misclassifications and provide additional detail if necessary. This provides interpretation of nuances in sentiment that are missed by the algorithm and allows valid insights to be distilled from noise. Public health experts at ProMED and Epidemico respectively curate the web and social media data. As more curated data is available, the algorithm is retrained to improve accuracy.

### Monthly Intelligence Reports:

Expert analysts at ProMED review the data from Vaccine Sentimeter on a monthly basis to draw out trends and possible signals in changing sentiments and vaccine confidence. Beyond a descriptive summary of sentiment, these reports tease out topics of high volume by country, analyze negative sentiment by country and vaccine/disease, and illustrate a time series of sentiments showing possible hubs of influence. The report also lists high-risk reports, which act as a starting point for practitioners in responding to rumors and loss of confidence in vaccines.

### Thank You

Thank you for using Vaccine Sentimeter. It would be greatly appreciated if you could please fill out a 5-minute survey (<http://bit.ly/Ne8ZeG>) sharing your experiences, and feel free to contact us directly at [info@epidemico.com](mailto:info@epidemico.com) with any ideas, suggestions, or feedback.

Current partners envision research collaborations through data sharing (already underway with London School of Hygiene and Tropical Medicine & Penn State University) and

## Vaccine Sentimeter: Global Monitoring of Vaccination Conversations

application of the tool to real world situations. We also actively seek collaborative grants for these satellite projects, and additional partners who are prepared to support, develop, and implement the tool across the world.