Multimedia Appendix 1. Supplementary material.

Relevance screening form

Question	Options	Definitions/additional notes
RefID		
,	1) Yes, general review or brief overview or commentary 2) Yes, within the context of primary research 3) None of the above If 3) "none of the above" is selected, abstract will be EXCLUDED	Yes, general review or overview or commentaries: Includes general reviews or commentaries (from peer-reviewed articles to lay magazine or newspaper articles) describing, discussing and/or promoting the use of social media in support of disease surveillance (within the context of broad disease prevention and control defined below), with or without specific contextual examples. Yes, within the context of primary research Any original research describing the use of social media in support of disease surveillance (within the context of broad disease prevention and control defined below), and reporting study (or program) design methods and results in replicable fashion of conventional, primary research. Social media are defined as a
		Social media are defined as a group of Internet-based online and mobile applications that allow the creation and exchange of user-generated content and data. (Examples

include Twitter, Facebook, Ushahidi, Google groups, data exhaust from Google searches, etc.).

<u>Disease surveillance</u> is defined as the ongoing systematic collection and analysis of data and the provision of information which leads to action being taken to prevent and control a disease. Within the context of this study, disease prevention and control is defined broadly as any activity(ies) designed to protect public health (including human, animal *and/or environmental health)* from actual or potential health threats and their harmful consequences. This would include activities related to (early) detection of disease (including new disease, sporadic or outbreaks, endemic and epidemic, infectious and chronic), prevention, control and eradication. Threats are limited to biological and chemical. Articles describing the use of social media within the context of educational or risk communication campaigns or strategies will be excluded.

Characterization form

Question	Options	Definitions/additional notes
RefID	Will be retrieved automatically from Distiller	notes
RefID Does the abstract investigate the use of social media in disease surveillance? (Check all that apply)	Will be retrieved automatically from Distiller. Yes, general review or commentary. Yes, general review also describing case studies or detailed examples of social media use in support of disease prevention and control. Yes, primary research. None of the above, specify reason for exclusion: If "None of the above" is selected, article will be EXCLUDED	notes All definitions from RS form are applicable-CONFIRMATION OF RELEVANCE (full-article level). Additional definitions: Yes, general review or commentary: Includes general review or commentaries of social media use in support of disease prevention and control, and only mentions examples without providing a detailed description. Yes, general review with case studies or examples: Includes general review or commentaries as above but also illustrates the concepts through provision of detailed case-studies or examples. Yes, primary research: The authors provide a detailed description of one or more social
		or more social media tools within the context of prevention and

			control of one or more diseases or issues.
Within the context of this article, what is the main objective of social media in support of disease prevention and control? Check all that applies	outbroinfect Infect Output Detect disease	warning, detection or control of disease eaks, epidemics or pandemics (due to tious agents or toxic substances) etion, control or management of chronic see et, specify:	Articles describing the use of social media within the context of educational or risk communication campaigns or strategies will be excluded. Note: Data extraction
			will initially be conducted for 1). Articles coded as 2) and 3) will proceed as time permits.
What type of document is the citation?	☐ Lay n☐ News☐ Gove	al article magazine article spaper article rnment or research report r, specify:	OtherE.g. conference proceedings, thesis, book, web site, etc.
In what sector(s) is the article written? (Check all that apply)	specif	rulture and agri-food, fy: h, specify disease: r, specify:	Agriculture and agrifood E.g., field crops, forestry, fishery, livestock Health Specify the specific sub-field (e.g. public health, eHealth, infectious diseases, chronic diseases, mental health, foodborne disease) and specific disease, if applicable (e.g. dengue)
			Other: social sciences, business, education,

		ecology, biology etc.
In the context of what population/target user group is the article written? (Check all that apply)	□ Practitioners, clinicians, or service providers □ Policy- and/or decision-makers □ Researchers/academics □ Other, specify :	Practitioners, clinicians, or service providers • Article is targeted at clinical stakeholders (e.g. doctors), service providers or other practitioners Policy/decision makers • Article is targeted at public or government policy and/or decision makers Other • E.g. researchers/acade mics, general public, etc.
Is a definition of social media provided in this article? What definitions are used to describe social media in this article? (Check all that apply)	☐ Provided by authors, specify: ☐ Other reference, specify: ☐ No definition provided	 Provided by authors Please type or copy and paste sentences used to define social media. Referenced Please list the sentences used to define social media and the author/year of the definition's reference source.
What definitions of terms related to social media are provided in this article? (Check all that apply)	☐ Provided by authors, specify: ☐ Other reference, specify: ☐ No definitions provided	Provided by authors • Please list the term(s) or type or copy and paste sentences used to describe the social media. Referenced

		• Please list the term(s) or sentences used to describe the social media and the author/year of the definition's reference source.
Within this article are there any references/citations that should be procured to obtain additional information about social media use in support of disease prevention and control?	[Textbox]	
What is the main study (review) objective as stated by the authors? Within the context of this article, what is the jurisdiction level of social	 [Textbox] ☐ International, specify: ☐ Regional (greater than 1 country), specify: 	Type or copy and paste as reported by authors.
media-based surveillance? Check all that applies	☐ National, specify: ☐ Other, specify:	
What specific social media tools are described or discussed by the authors? Check all that applies	Google Twitter Other, please specify:	Examples of Google as a tool include data exhaust (data on searches conducted), or tools such as Correlate (not just simple Google searches conducted by the researchers).
Is there a description of the	☐ Yes ☐ No	Note: May not be applicable for review-

method as stated		type papers
by the author(s)?	If "Yes" is selected, then:	
		Type or copy and
	What is the name of the social media tool?	paste as reported by
		authors. Under 1), if
	What is the data source?	description is not provided please write
	What is the data source.	"not described".
		If more than one social
	What is the data time period coverage?	media tools or case
		studies or examples
	Wilest is the sector was a sector with a sec	are reported within
	What is the outcome measurement?	one article, please extract these data
		using separate form.
	Is the model validated?	Only questions for this
	☐ Yes	section need to be
	□ No	completed more than
	☐ Not applicable	once.
	Is there an additional social media tool described?	
	Yes	
	□ No	
	If "Yes" is selected then:	
	What is the name of the 2 nd social media tool?	
	what is the name of the 2 social media toor:	
	What is the data source for this 2 nd social media	
	tool?	
	What is the data time period coverage for this 2 nd	
	social media tool?	
	What is the outcome measurement for this 2 nd	
	social media tool?	
	Is the model validated?	
	☐ Yes ☐ No	
	☐ Not applicable	
	Tr	

Do the authors compare (or correlate) the accuracy of social media-based surveillance concurrently with other surveillance systems?	 □ Not applicable □ Yes, it was compared to an existing sentinel or surveillance system, specify: □ No, it was not compared to another surveillance system 	Note: May not be applicable for review-type papers Under 'specify': type or copy and paste as reported by authors. If description is not provided please write "not described".
Do the authors report that the social media-based surveillance is more accurate than other concurrently used surveillance systems? Check all that applies	 Yes, social media-based surveillance is more accurate, specify:	Note: May not be applicable for review-type papers Type or copy and paste as reported by the authors. Accuracy with respect to sensitivity (properly categorizing positives), specificity (properly categorizing negatives), and/or correctly predicting patterns in time and space
Do the authors report that the social media-based surveillance is effective in general within the context of disease prevention and control?	☐ Yes, please specify: ☐ Not reported	Type or copy and paste as reported by the authors.
Do the authors report that the social media-based surveillance is cost-effective within the context of disease prevention and control?	Yes, please specify: Not reported	Type or copy and paste as reported by the authors.

Do the authors report the strengths of the social-media based surveillance? Check all that applies	☐ Ease of use ☐ Ease of access ☐ Time effective (less human labor) ☐ Faster ☐ Cost-effective ☐ Other, please specify: ☐ No, not reported	
Do the authors report the weakness of the social-media based surveillance? Check all that applies	 □ Information overload □ Lack of validated methods (sensitivity and specificity) □ Other, please specify: □ No, not reported 	Type or copy and paste as reported by the authors.
What is overall recommendation by the authors regarding the social media-based surveillance within the context of disease prevention and control? Check all that applies	 □ Primarily in support of existing surveillance systems, please specify: □ Primarily in the context of future, when the methods are better validated or evaluated □ Other, please specify: 	Type or copy and paste as reported by the authors.
Do the authors report or comment on the human or professional-based social networks (outside of technology) within the context of disease prevention and control?	[Textbox]	Type or copy and paste as reported by the authors.
Do the authors report or comment on implications for the end user Please provide a	[Textbox]	Type or copy and paste as reported by the authors.
summary of any		

additional	
important detail in	
the article that has	
not already been	
captured in this	
form. (If	
applicable)	

Thematic analysis form and code list

Question	Options	Definitions/additional notes
Does this article contain sufficient information for coding?	1) Yes 2) No, specify reason: If #2 is selected above, please submit the form without completing the remaining questions	Reasons for exclusion could include short abstracts or articles with little to no relevant information about social media surveillance
Please indicate the key themes that are reviewed or described in detail in this review. Check all that apply and indicate others as necessary.	Please see code list below for list of possible code options and their definitions	Only include themes that are discussed in detail or depth in the article – and not those that are only briefly mentioned.
Please indicate 5-10 key messages from the article		Please include a brief summary of 5-10 key messages from the article in point form.
Are there any other details or aspects of this article that should be noted? (E.g. key quotes)		If there are any additional aspects or details about the article that should be noted, please enter those here.

Code List

Couc List	
Code	Definition
Accessibility	Data is publicly reported/available/accessible for free
Accuracy/validity/reliability	Discussion of sensitivity and/or specificity or validity/reliability of
	the programs (e.g. false alarms)
Beneficiaries	Discussion of who benefits from social media surveillance
Collaborative	Collaborations to integrate data across multiple
	jurisdictions/agencies/networks/countries
Confidentiality/privacy	Discussion of confidentiality of information or privacy as an issue
Costs/funding	Discussion of costs, cost-effectiveness of funding for programs
Data mining	Discussion of methods and processes to synthesize and extract
	necessary data
Data noise/volume	Discussion of need to filter data to reduce background "noise" and
	prevent information overload
Diverse end users	Discussion of different populations of end-users (e.g. researchers,
	practitioners, general public)
Evaluation/efficacy	Discussion of need to evaluate and verify the efficacy of the
	programs
Government role	Discussion of role of government agencies
Infrastructure/capacity	Discussion of infrastructure or capacity (e.g. skills, training) to
	implement social media surveillance

Informal/public data	Informal sources for data collection from general public
	information (e.g. blogs, web searches, RSS feeds); reflects public
	"collective intelligence"
International	Program is global or international in nature
Legal implications	Discussion of legal implications or responsibilities of social media
	surveillance (e.g. outbreak reporting obligations)
Mapping/visualization	Includes capabilities to map/visualize trends
Moderators	Discussion of need for moderators to manage information flow
Multiple data sources	Data is aggregated from multiple sources (e.g. blogs, searches,
	wikis, RSS feeds)
Networking/connections	Programs facilitate network building and connections among users
New technology tools	Use of mobile phones (e.g. iPhone), "texting", Google, Twitter,
	Wikipedia, RSS feeds, Facebook, etc. for surveillance
Ownership	Discussion of who owns or is responsible for the data/program
Population/jurisdiction	Discussion of populations or jurisdictions where program is or can
	be implemented (e.g. developed vs. developing countries)
Detection of emerging	Programs can be used to detect emerging diseases, outbreaks, or
diseases, outbreaks,	pandemics
pandemics	
Rapid response/real-time	Importance of timely recognition and response of outbreaks/events
Supplementary role	Social media surveillance discussed in role supplementary or
	supportive of existing official/government programs
User-friendly	Discussion of need to ensure programs are easy to use/user
	friendly

List of search terms

The search strategy is a combination of key terms developed within the following domains: "social media/internet-based tools" AND "disease" AND "surveillance."

Broad category	Search terms		
1. Social	social media OR blog* OR microblog* OR crowdsourc* OR twitter		
media/internet-	OR facebook OR wiki* OR search engine OR (google OR internet		
based tools	OR web OR online) AND (social network OR social networking O		
	search OR query)		
2. Disease	disease OR illness* OR outbreak OR epidemic OR pandemic		
3. Surveillance	surveillance OR track* OR forecast* OR monitor* OR predict*		

Limits: Starting in 2002

List of relevant articles from article characterization (n=51)

Case study articles (n=8)			
Author	Title	Source	Year
Brownstein JS, Freifeld	Surveillance Sans Frontieres:	PLoS Med.	2008
CC, Reis BY and Mandl	Internet-based emerging		
KD.	infectious disease intelligence and the HealthMap project		
Carneiro HA and	Google trends: A web-based tool	Clin. Infect. Dis.	2009
Mylonakis E.	for real-time surveillance of disease outbreaks		
Collier N, Doan S,	BioCaster: detecting public health	Bioinformatics	2008
Kawazoe A, Goodwin RM,	rumors with a Web-based text		
Conway M, Tateno Y, Ngo	mining system		
QH, Dien D, Kawtrakul A,			
Takeuchi K, Shigematsu M			
and Taniguchi K.			
Eke PI.	Using social media for research and public health surveillance	J. Dent. Res.	2011
Eysenbach G.	Infodemiology and infoveillance:	J. Med. Internet	2009
	framework for an emerging set of	Res.	
	public health informatics methods		
	to analyze search, communication		
	and publication behavior on the		
77 ID 1 MOI	Internet.		2010
Kamel Boulos MN,	Social Web mining and	Comput. Methods	2010
Sanfilippo AP, Corley CD and Wheeler S.	exploitation for serious	Programs Biomed.	
and wheeler S.	applications: Technosocial		
	Predictive Analytics and related technologies for public health,		
	environmental and national		
	security surveillance		
Madoff LC, Fisman DN	A new approach to monitoring	PLoS Negl Trop.	2011
and Kass-Hout T.	dengue activity	Dis.	2011
Schein R, Wilson K and	Literature Review on	Report	2010
Keelan J.	Effectiveness of the Use of Social	-r	
	Media: a report for Peel Public		
	Health		

General review articles (n=11)			
Author	Title	Source	Year
Anonymous	Google Flu Trends includes 14	Euro Surveill.	2009
	European countries.		
Brownstein JS, Freifeld	Digital disease detection -	New Engl. J.Med.	2009
CC and Madoff LC.	Harnessing the web for public health surveillance		
Brownstein JS.	Using social media for disease surveillance	CNN.com Blogs	2011
Cook AR, Chen MIC and	Internet search limitations and	Emerg. Infect. Dis.	2010
Lin RTP.	pandemic influenza, Singapore	Emerg. Infect. Dis.	2010
Dreesman J and Denecke	Challenges for signal generation	Stud. Health	2011
K.	from medical social media data	Technol.Inform.	2011
Greene M.	Constructing paths through social	Proc SPIE Int Soc	2011
	networks for disease surveillance	Opt Eng	
Greene M.	Epidemiological monitoring for	Proc SPIE Int Soc	2010
	emerging infectious diseases	Opt Eng	
Khan AS, Fleischauer A,	The next public health revolution:	Am. J. Public	2010
Casani J and Groseclose SL.	Public health information fusion and social networks	Health	
Lopes LF, Silva FAB,	Epidemic marketplace: An	Lect. Notes	2010
Couto F, Zamite J,	information management system	Comput. Sci.	2010
Ferreira H, Sousa C and	for epidemiological data	compout sen	
Silva MJ.	1		
Morozov E.	Swine flu: Twitter's power to misinform	Foreign Policy	2009
Strongin R.	Health reform in 140 characters	Med Device Diagn Ind	2010

Primary research articles (n=32)			
Author	Title	Source	Year
Achrekar H, Gandhe A, Lazarus R, Yu S- and Liu B.	Predicting flu trends using twitter data	IEEE Conf. Comput. Commun. Workshops	2011
Althouse BM, Ng YY and Cummings DA.	Prediction of dengue incidence using search query surveillance	PLoS Negl Trop.Dis.	2011
Aramaki E, Maskawa S and Morita M.	Twitter Catches the Flu: Detecting influenza epidemics using Twitter	Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing	2011;
Boyle JR, Sparks RS, Keijzers GB, Crilly JL, Lind JF and Ryan LM.	Prediction and surveillance of influenza epidemics	Med. J. Aust.	2011
Chan EH, Sahai V, Conrad C and Brownstein JS.	Using web search query data to monitor dengue epidemics: A new model for neglected tropical disease surveillance	PLoS. Negl. Trop. Dis.	2011
Chester TLS, Taylor M, Sandhu J, Forsting S, Ellis A, Stirling R and Galanis E.	Use of a Web Forum and an Online Questionnaire in the Detection and Investigation of an Outbreak	Online Journal of Public Health Informatics	2011
Chew C and Eysenbach G.	Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak	PLoS ONE	2010
Corley CD, Cook DJ, Mikler AR and Singh KP.	Using Web and social media for influenza surveillance.	Adv. Exp. Med. Biol.	2010
Culotta A.	Detecting influenza outbreaks by analyzing Twitter messages		2010
de Quincey E and Kostkova P.	Early warning and outbreak detection using social networking websites: the potential of twitter	Electronic healthcare, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunicatio ns Engineering	2010
Dukic VM, David MZ and Lauderdale DS.	Internet queries and methicillin- resistant staphylococcus aureus surveillance	Emerg. Infect. Dis.	2011
Dukic VM, Lopes LF and Polson NG.	Tracking flu epidemics using Google Flu Trends	AD41 - 2	2011
Eysenbach G.	Infodemiology: tracking flu-	AMIA Annu Symp	2006

	related searches on the web for syndromic surveillance.	Proc	
Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS and Brilliant L.	Detecting influenza epidemics using search engine query data	Nature	2009
Goel S, Hofman JM, Lahaie S, Pennock DM and Watts DJ.	Predicting consumer behavior with web search	Proc. Natl. Acad. Sci. U.S.A.	2010
Gomide J, Veloso A, Meira W, Almeida V, Benevenuto F, Ferraz F and Teixeira M.	Dengue surveillance based on a computational model of spatio-temporal locality of Twitter	Proceedings of the ACM WebSci'l1	2011
Hooten MB, Anderson J and Waller LA.	Assessing North American influenza dynamics with a statistical SIRS model	Spat.Spatio- temporal Epidemiol.	2010
Hulth A and Rydevik G.	GET WELL: An automated surveillance system for gaining new epidemiological knowledge	BMC Public Health	2011
Krieck M, Dreesman J, Otrusina L and Denecke K.	A New Age of Public Health: Identifying Disease Outbreaks by Analyzing Tweets	Health Web Science Workshop, ACM Web Science 2011 3rd International Conference on Web Science	2011
Ku Y, Chiu C, Zhang Y, Fan L and Chen H.	Global disease surveillance using social media: HIV/AIDS content intervention in web forums	ISI - IEEE Int. Conf. Intell. Secur. Informatics: Public Saf. Secur.	2010
Lampos V and Cristianini N.	Tracking the flu pandemic by monitoring the social web	Int. Workshop Cognitive Inf. Process.	2010
Lampos V, De Bie T and Cristianini N.	Flu detector-tracking epidemics on twitter	Lect. Notes Comput. Sci.	2010
Leffler CT, Davenport B and Chan D.	Frequency and seasonal variation of ophthalmology-related internet searches	Can. J. Ophthalmol.	2010
Luo Y, Zeng D, Cao Z, Zheng X, Wang Y, Wang Q and Zhao H.	Using multi-source web data for epidemic surveillance: A case study of the 2009 Influenza A (H1N1) pandemic in Beijing	Proc. IEEE Int. Conf. Serv. Oper. Logist., Informatics, SOLI	2010
Ortiz JR, Zhou H, Shay DK, Neuzil KM, Fowlkes AL and Goss CH.	Monitoring Influenza activity in the United States: A comparison of traditional surveillance systems with Google Flu Trends	PLoS ONE	2011

Pelat C, Turbelin C, Bar- Hen A, Flahault A and Valleron A	More diseases tracked by using google trends	Emerg. Infect. Dis.	2009
Polgreen PM, Chen Y, Pennock DM and Nelson FD.	Using internet searches for influenza surveillance	Clin. Infect. Dis.	2008
Seifter A, Schwarzwalder A, Geis K and Aucott J.	The utility of "Google Trends" for epidemiological research: Lyme disease as an example	Geospatial Health	2010
Signorini A, Segre AM and Polgreen PM.	The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic	PLoS ONE	2011
Valdivia Pérez A, Miguel Benito A, Escortell Mayor E and Monge Corella S.	Can the influenza epidemic be predicted by data from the Internet?	Gaceta Sanit.	2010
Valdivia A, López-Alcalde J, Vicente M, Pichiule M, Ruiz M and Ordobas M.	Monitoring influenza activity in Europe with Google Flu Trends: Comparison with the findings of sentinel physician networks - results for 2009-10	Eurosurveillance	2010
Zhou X- and Shen H	Notifiable infectious disease surveillance with data collected by search engine	J. Zhejiang Univ. Sci. C	2010