Digital Participatory Surveillance: the ten-years experience of Influenzanet and what’s next

Daniela Paolotti, Ph.D.
@danielapaolotti

Research Leader
ISI Foundation, Turin, Italy

Workshop on Infectious Disease Surveillance
Institute of Social and Preventive Medicine (University of Bern)
November 25th, 2019

ISI Foundation

- basic and applied research
- 30+ years of history
- 40+ researchers
- Turin, Italy & New York, USA
- international network

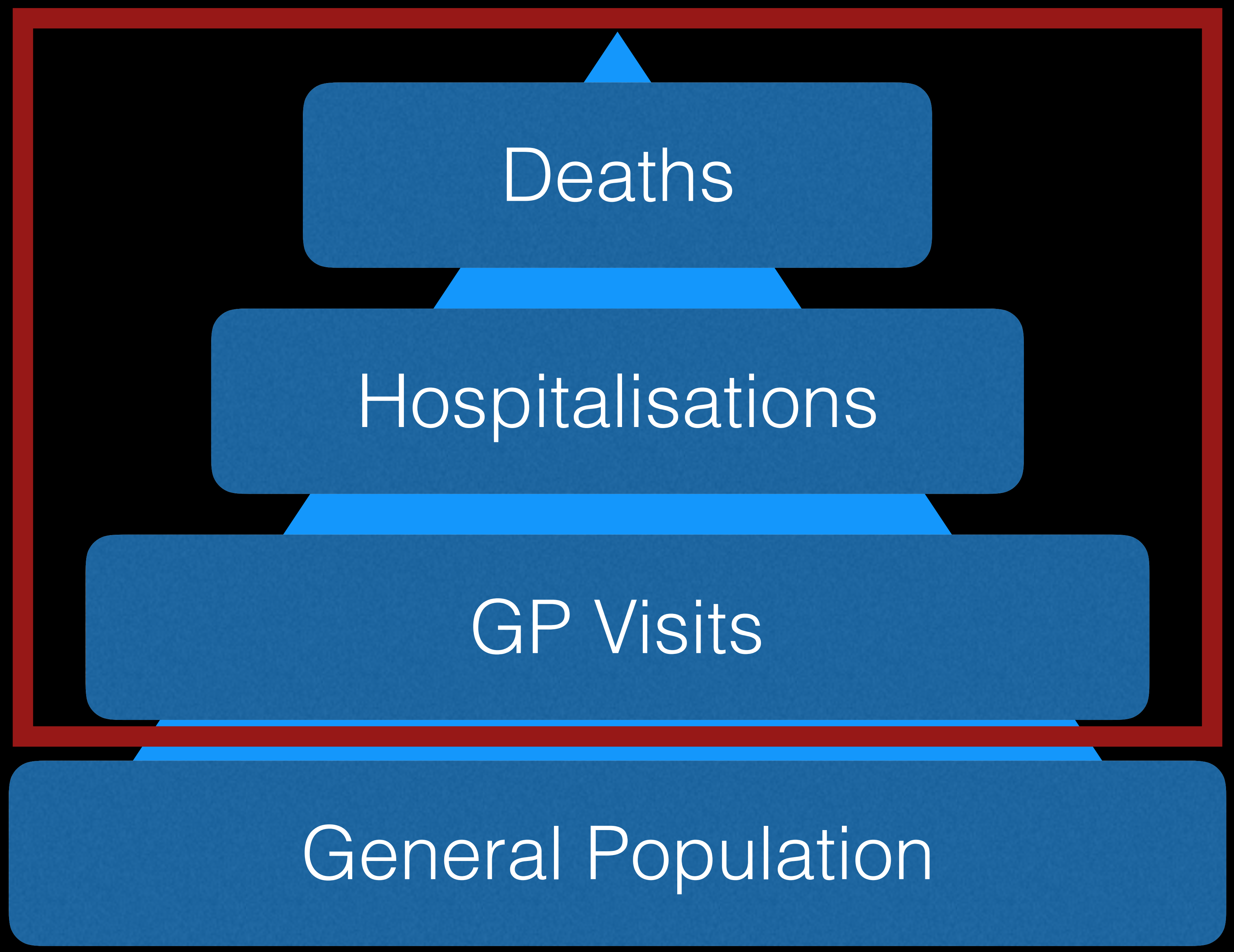
- focus on:
  - data & network science
  - computational and digital epidemiology
  - computational social science

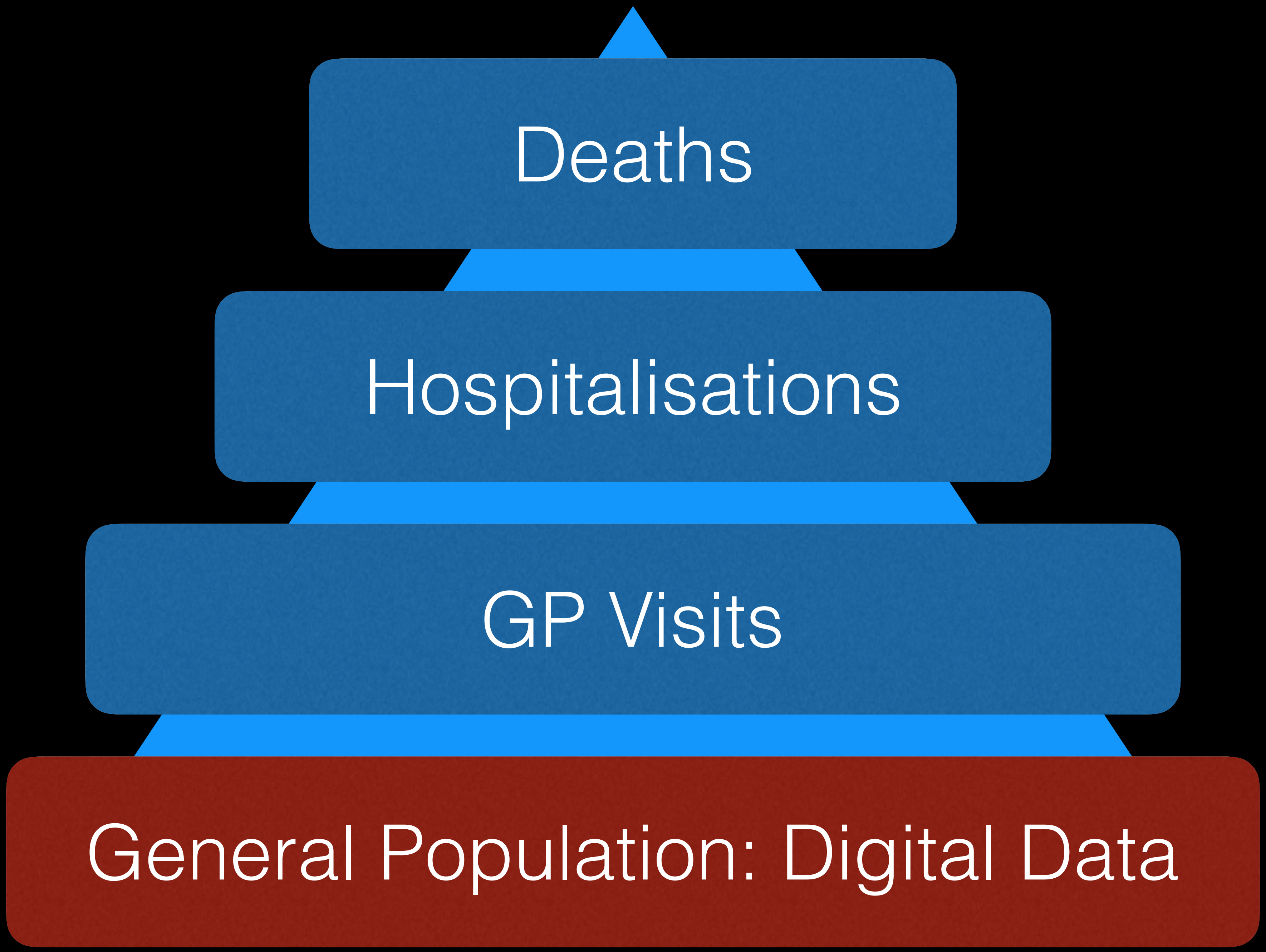
- supported by:
  - bank foundations
  - EC research grants
  - industrial partnerships

www.isi.it


Surveillance
Disease Burden

Surveillance




Deaths
Hospitalisations
GP Visits
General Population: Digital Data

RESEARCH ARTICLE

The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic

Alessio Signorini, Alberto Maria Segre, Philip M. Polgreen

Published: May 4, 2011 • http://dx.doi.org/10.1371/journal.pone.0019467

Abstract

Twitter is a free social networking and micro-blogging service that enables its millions of users to send and read each other's "tweets," or short, 140-character messages. The service has more than 190 million registered users and processes about 55 million tweets per day. Useful information about news and geopolitical events lies embedded in the Twitter stream, which embodies, in the aggregate, Twitter users' perspectives and reactions to current events. By virtue of sheer volume, content embedded in the Twitter stream may be useful for tracking or even forecasting behavior if it can be extracted in an efficient manner. In this study, we examine the use of information embedded in the Twitter stream to (1) track rapidly-evolving public sentiment with respect to H1N1 or swine flu, and (2) track and measure actual disease activity. We also show that Twitter can be used as a measure of public interest or concern about health-related events. Our results show that estimates of influenza-like illness derived from Twitter chatter accurately track reported disease levels.

OPEN ACCESS Freely available online

PLOS COMPUTATIONAL BIOLOGY

Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time

David J. McIver*, John S. Brownstein

Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America

Abstract

Circulating levels of both seasonal and pandemic influenza require constant surveillance to ensure the health and safety of the population. While up-to-date information is critical, traditional surveillance systems can have data availability lags of up to two weeks. We introduce a novel method of estimating, in near-real time, the level of influenza-like illness (ILI) in the United States (US) by monitoring the rate of particular Wikipedia article views on a daily basis. We calculated the number of times certain influenza- or health-related Wikipedia articles were accessed each day between December 2007 and August 2013 and compared these data to official ILI activity levels provided by the Centers for Disease Control and Prevention (CDC). We developed a Poisson model that accurately estimates the level of ILI activity in the American population, up to two weeks ahead of the CDC, with an absolute average difference between the two estimates of just 0.27% over 294 weeks of data. Wikipedia-derived ILI models performed well through both abnormally high media coverage events (such as during the 2009 H1N1 pandemic) as well as unusually severe influenza seasons (such as the 2012–2013 influenza season). Wikipedia usage accurately estimated the week of peak ILI activity 17% more often than Google Flu Trends data and was often more accurate in its measure of ILI intensity. With further study, this method could potentially be implemented for continuous monitoring of ILI activity in the US and to provide support for traditional influenza surveillance tools.

Citation: McIver DJ, Brownstein JS (2014) Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time. PLoS Comput Biol 10(4): e1003581. doi:10.1371/journal.pcbi.1003581

Editor: Marcel Salathé, Pennsylvania State University, United States of America

Received December 20, 2013; **Accepted** March 11, 2014; **Published** April 17, 2014

Funding: This work was funded by the National Institutes of Health and National Library of Medicine 1R01LM010812-03. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: david.mciver@childrens.harvard.edu

RESEARCH ARTICLE

Monitoring Influenza Epidemics in China with Search Query from Baidu

Qingyu Yuan, Elaine O. Nsoesie, Benfu Lv, Geng Peng, Rumi Chunara, John S. Brownstein

Published: May 30, 2013 • http://dx.doi.org/10.1371/journal.pone.0064323

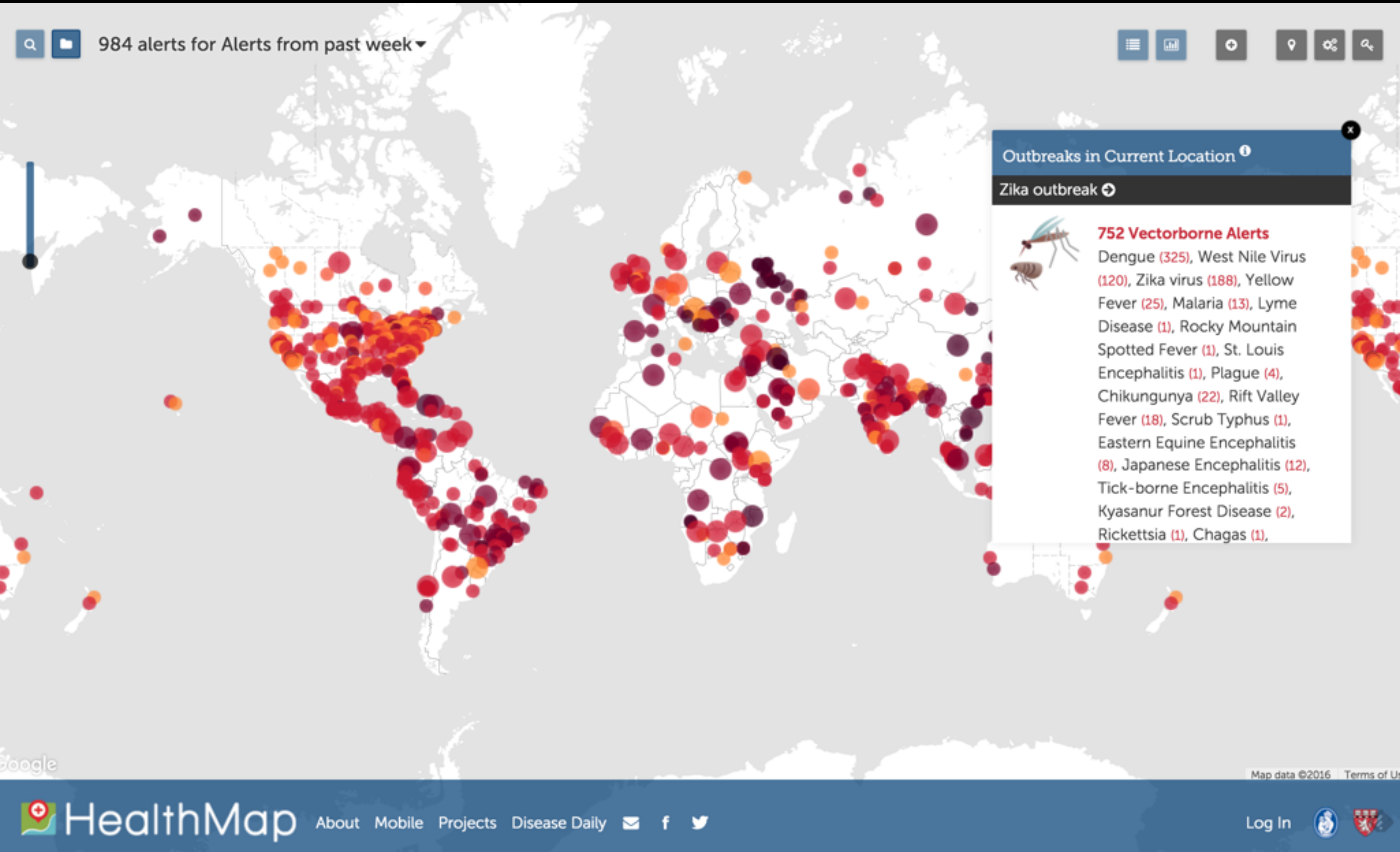
Abstract

Several approaches have been proposed for near real-time detection and prediction of the spread of influenza. These include search query data for influenza-related terms, which has been explored as a tool for augmenting traditional surveillance methods. In this paper, we present a method that uses Internet search query data from Baidu to model and monitor influenza activity in China. The objectives of the study are to present a comprehensive technique for: (i) keyword selection, (ii) keyword filtering, (iii) index composition and (iv) modeling and detection of influenza activity in China. Sequential time-series for the selected composite keyword index is significantly correlated with Chinese influenza case data. In addition, one-month ahead prediction of influenza cases for the first eight months of 2012 has a mean absolute percent error less than 11%. To our knowledge, this is the first study on the use of search query data from Baidu in conjunction with this approach for estimation of influenza activity in China.

Subject Areas

Influenza
Swine influenza
Influenza viruses
H1N1
China
Influenza A virus
Infectious disease s...

healthmap.org



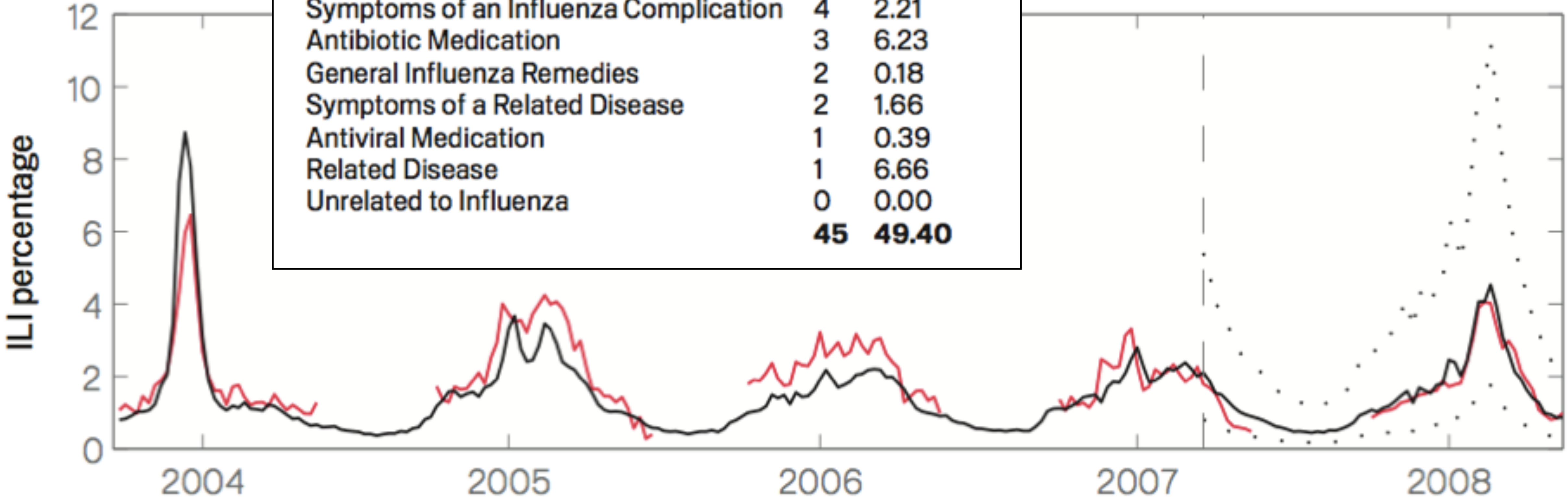
Google

Detecting influenza epidemics using search engine query data

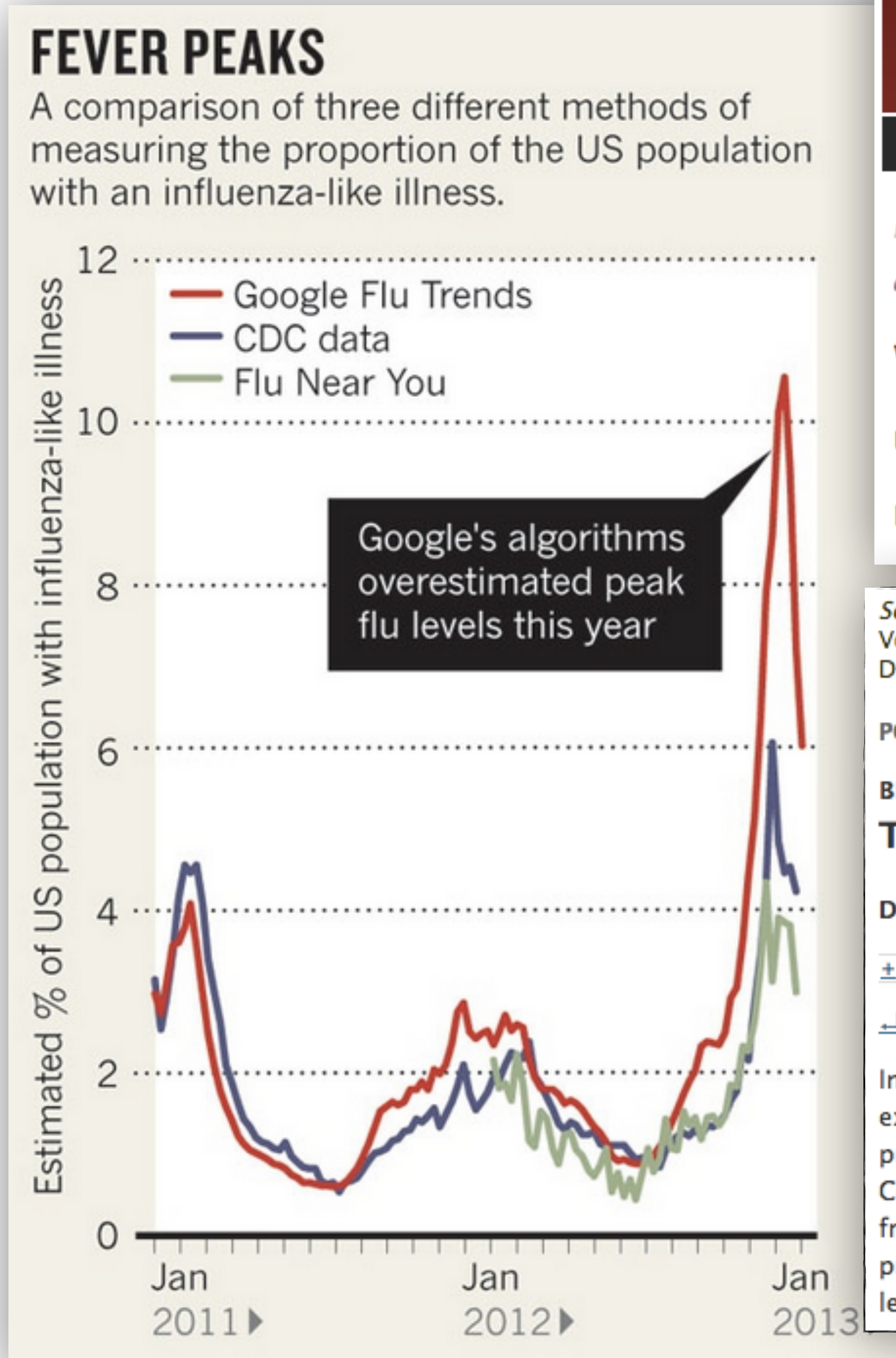
Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

[1]Google Inc. [2]Centers for Disease Control and Prevention

	Top 45 Queries	
Search Query Topic	N	Weighted
Influenza Complication	11	18.15
Cold/Flu Remedy	8	5.05
General Influenza Symptoms	5	2.60
Term for Influenza	4	3.74
Specific Influenza Symptom	4	2.54
Symptoms of an Influenza Complication	4	2.21
Antibiotic Medication	3	6.23
General Influenza Remedies	2	0.18
Symptoms of a Related Disease	2	1.66
Antiviral Medication	1	0.39
Related Disease	1	6.66
Unrelated to Influenza	0	0.00
	45	**49.40**



J. Ginsberg *et al.*, Nature 457, 1012 (2009)



nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive

Archive > Volume 494 > Issue 7436 > News > Article

NATURE | NEWS

عربي

When Google got flu wrong

US outbreak foxes a leading web-based method for tracking seasonal flu.

Declan Butler

Science 14 March 2014:
Vol. 343 no. 6176 pp. 1203–1205
DOI: 10.1126/science.1248506

‹ Prev | Table
Read Full T

POLICY FORUM

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

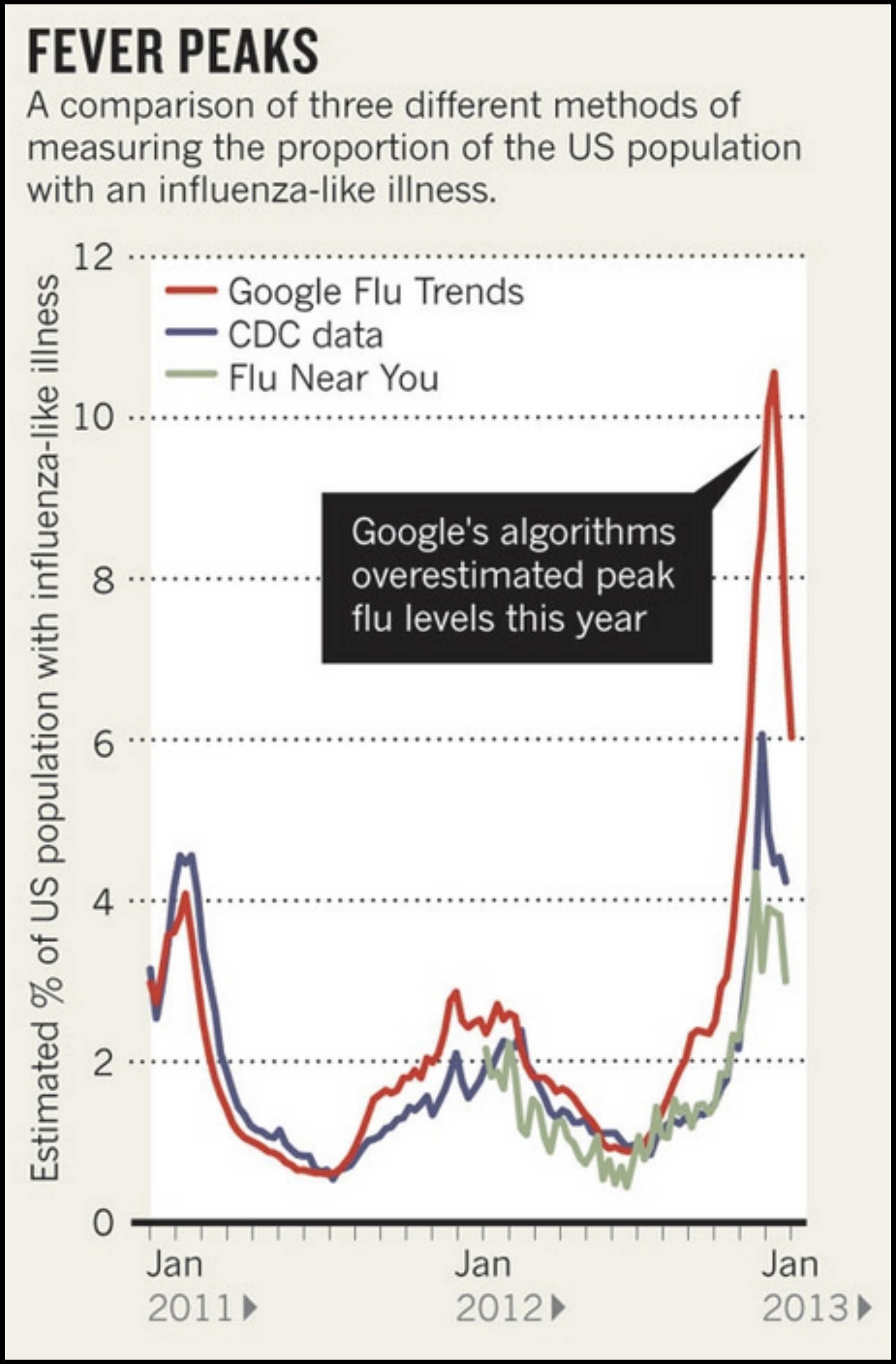
David Lazer[1,2,*], Ryan Kennedy[1,3,4], Gary King[3], Alessandro Vespignani[5,6,3]

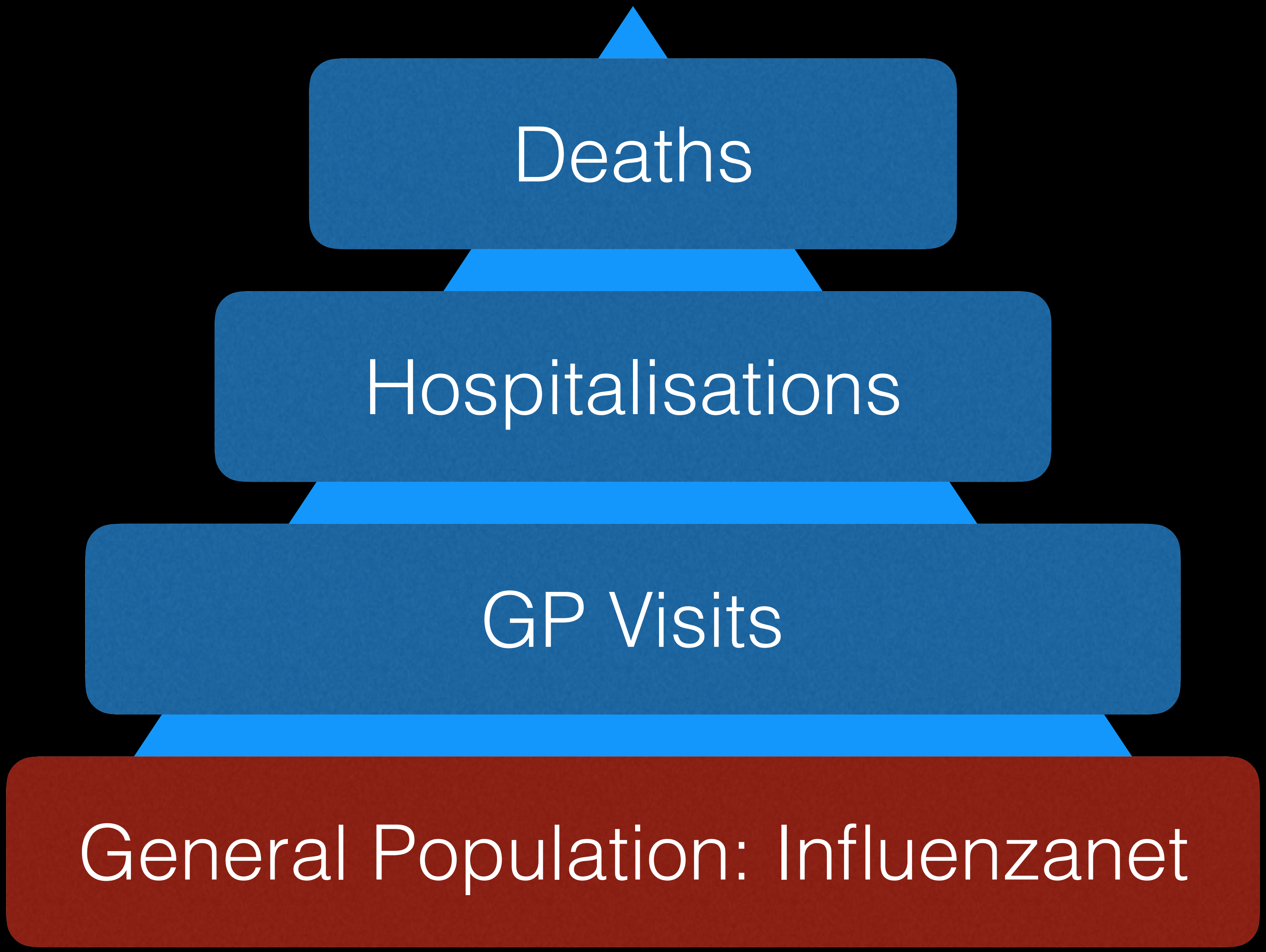
+ Author Affiliations

*Corresponding author. E-mail: d.lazer@neu.edu.

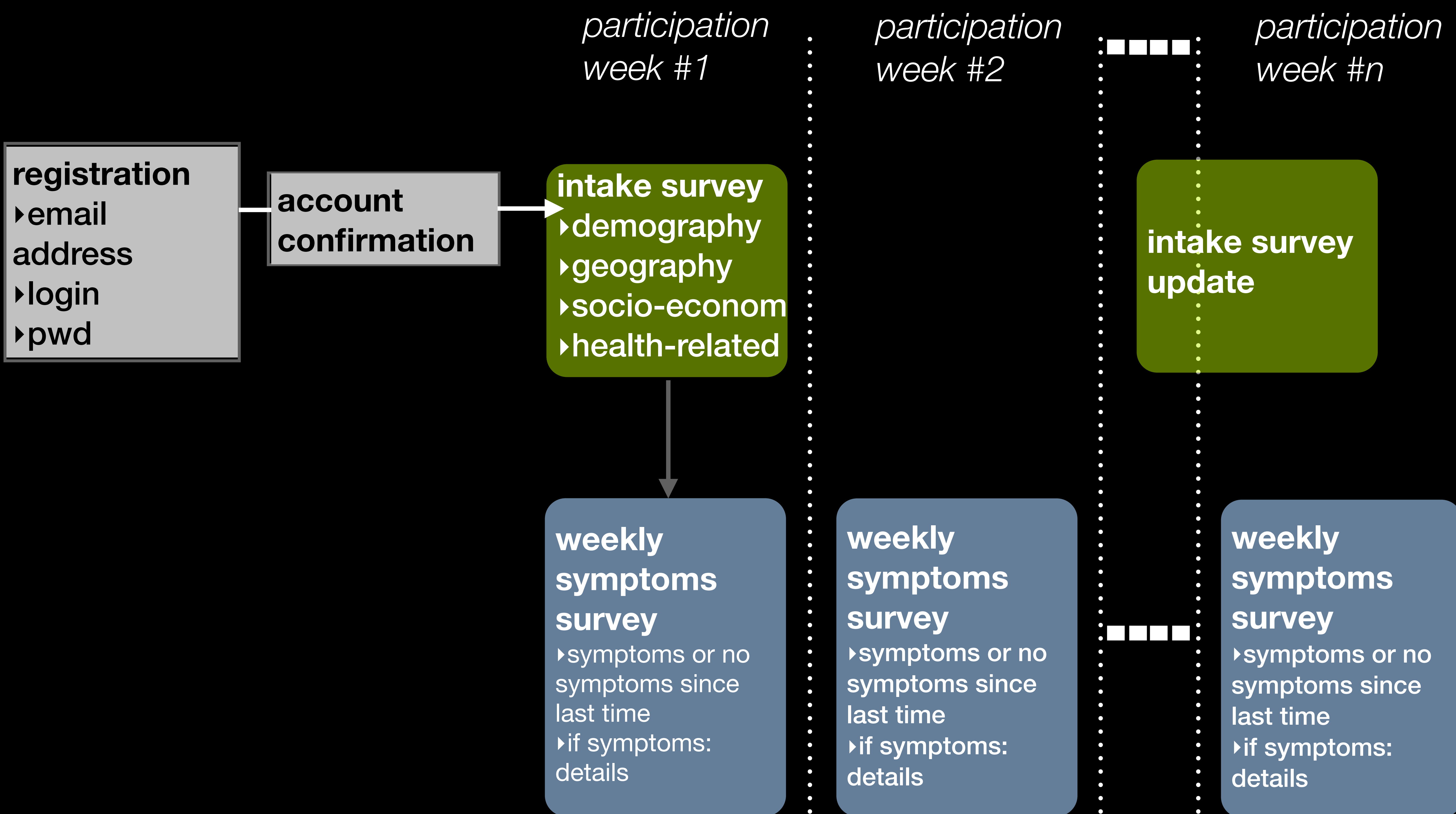
In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that G
executives or the creators of the flu tracking system would have hoped. *Nature* report
predicting more than double the proportion of doctor visits for influenza-like illness (
Centers for Disease Control and Prevention (CDC), which bases its estimates on survei
from laboratories across the United States (*1*, *2*). This happened despite the fact that
predict CDC reports. Given that GFT is often held up as an exemplary use of big data
lessons can we draw from this error?

- Passive data sources don’t describe who is well
- Low specificity



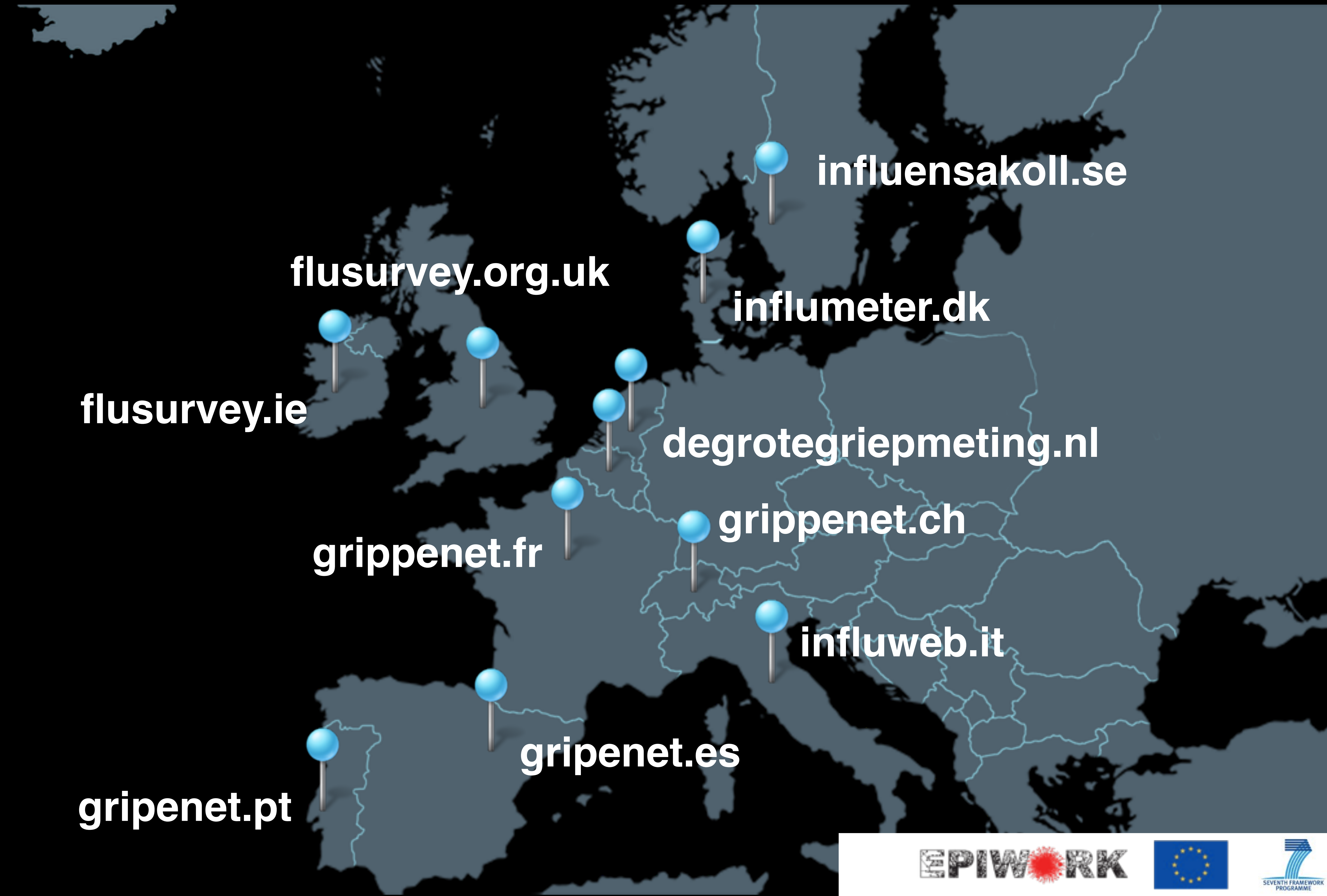

Deaths
Hospitalisations
GP Visits
General Population: Influenzanet

INFLUENZANET study design

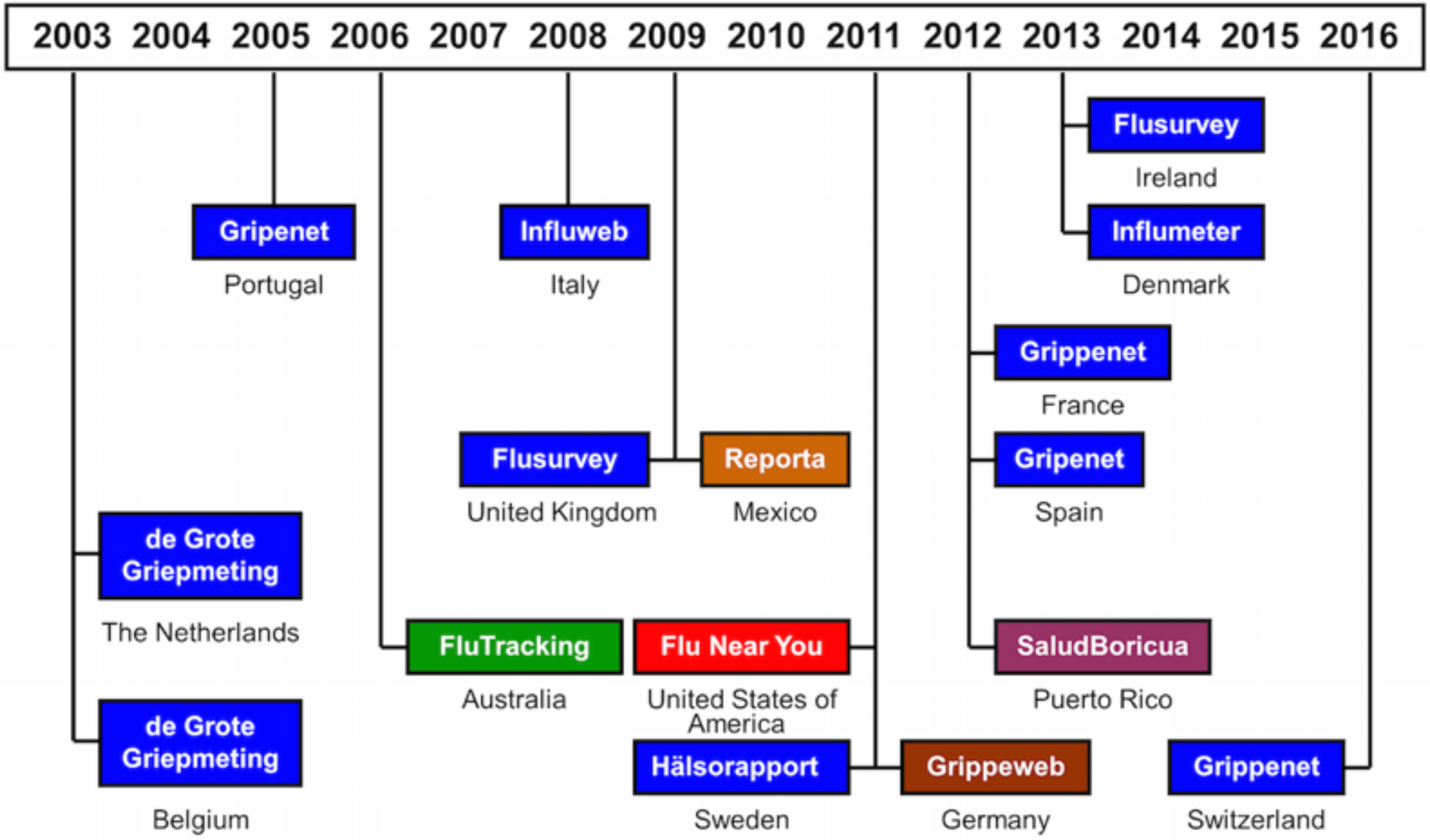


<u>uniform data collection across countries</u>

INFLUENZANET.INFO



INFLUENZANET: a timeline

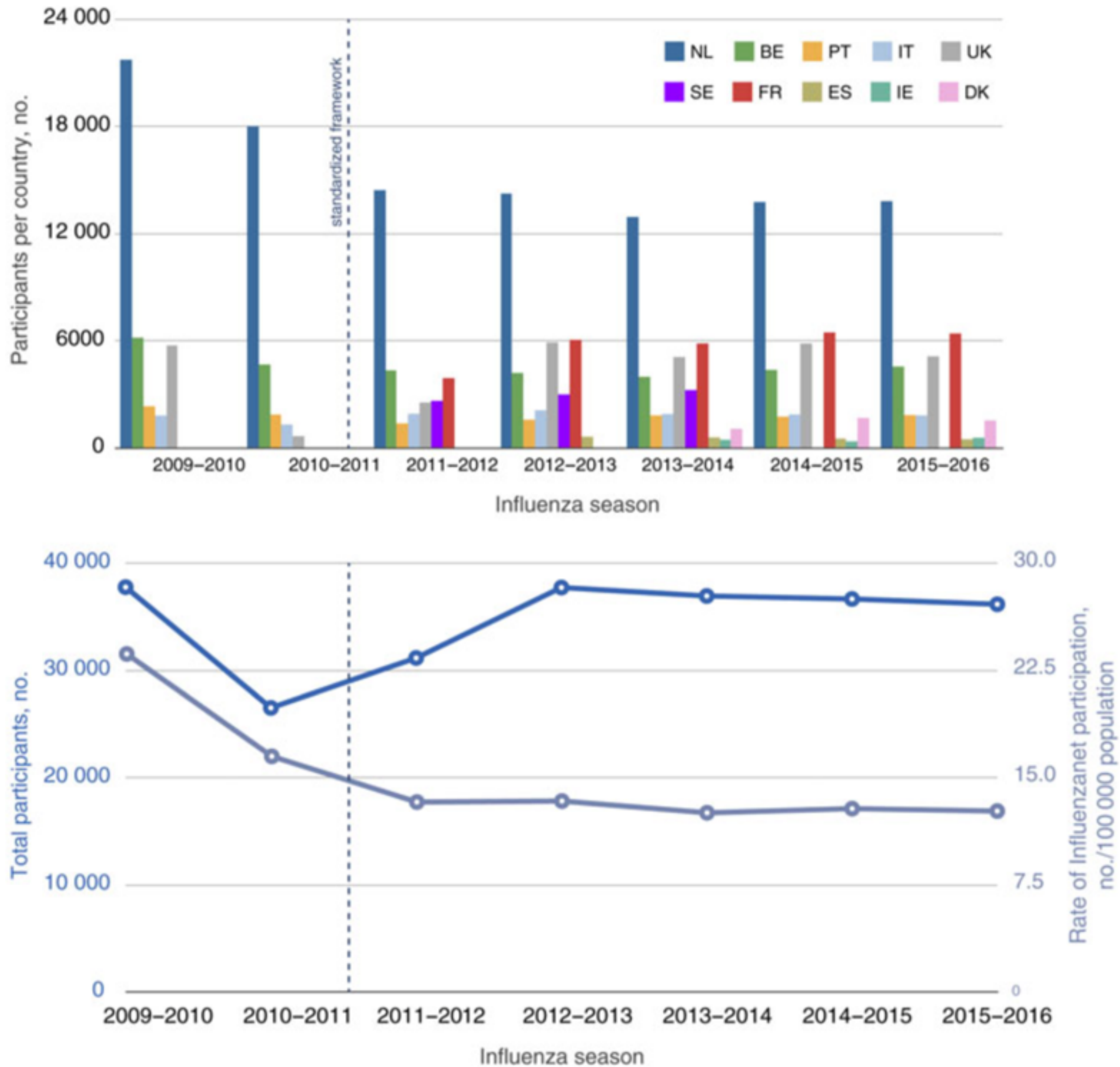


Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, Kjelsø C, Mexia R, Moreno Y, Meloni S, Paolotti D, Perrotta D, van Straten E, Franco AO
Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe
JMIR Public Health Surveill 2017;3(3):e66

study design: key questions

- ☑ weekly individual symptoms (or lack thereof)
- ☑ vaccine coverage
- ☑ healthcare seeking
- ☑ drugs uptake (including antibiotics)
- ☑ perceived severity
- ☑ duration of the symptoms
- ☑ absenteeism

C. Guerrisi et al, *Participatory Syndromic Surveillance of Influenza in Europe* Journal of Infectious Diseases (2016) 214 (suppl 4): S386-S392.



INFLUENZANET.INFO

influweb.it - ISI Foundation & Istituto Superiore di Sanità, Italy

grippenet.fr - INSERM, France

gripenet.pt - INSA, Portugal

gripenet.es - University of Saragoza, Spain

flusurvey.net - Public Health England

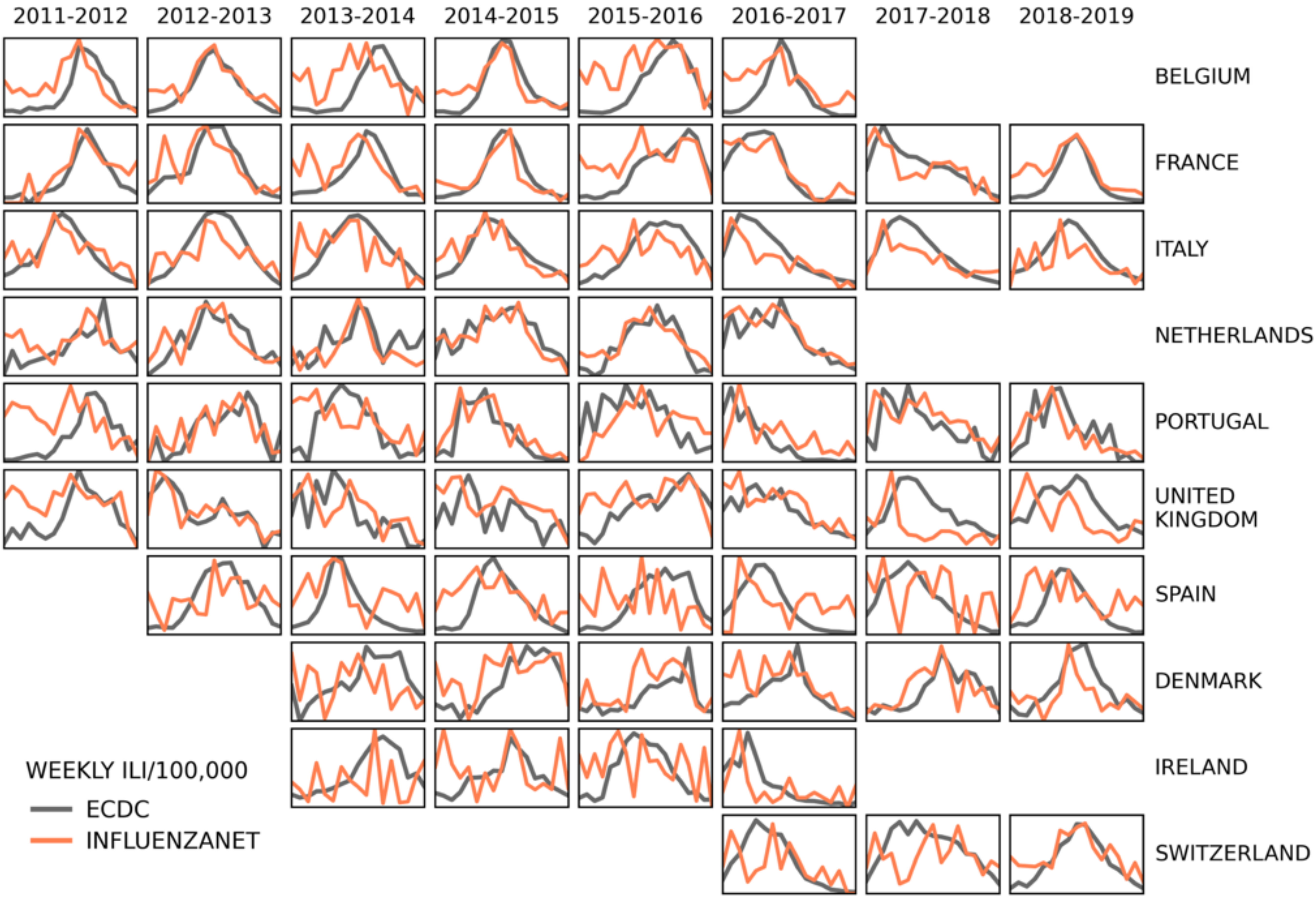
influensakoll.se - Public Health Agency of Sweden

influmeter.dk - Staten Serum Institute, Denmark

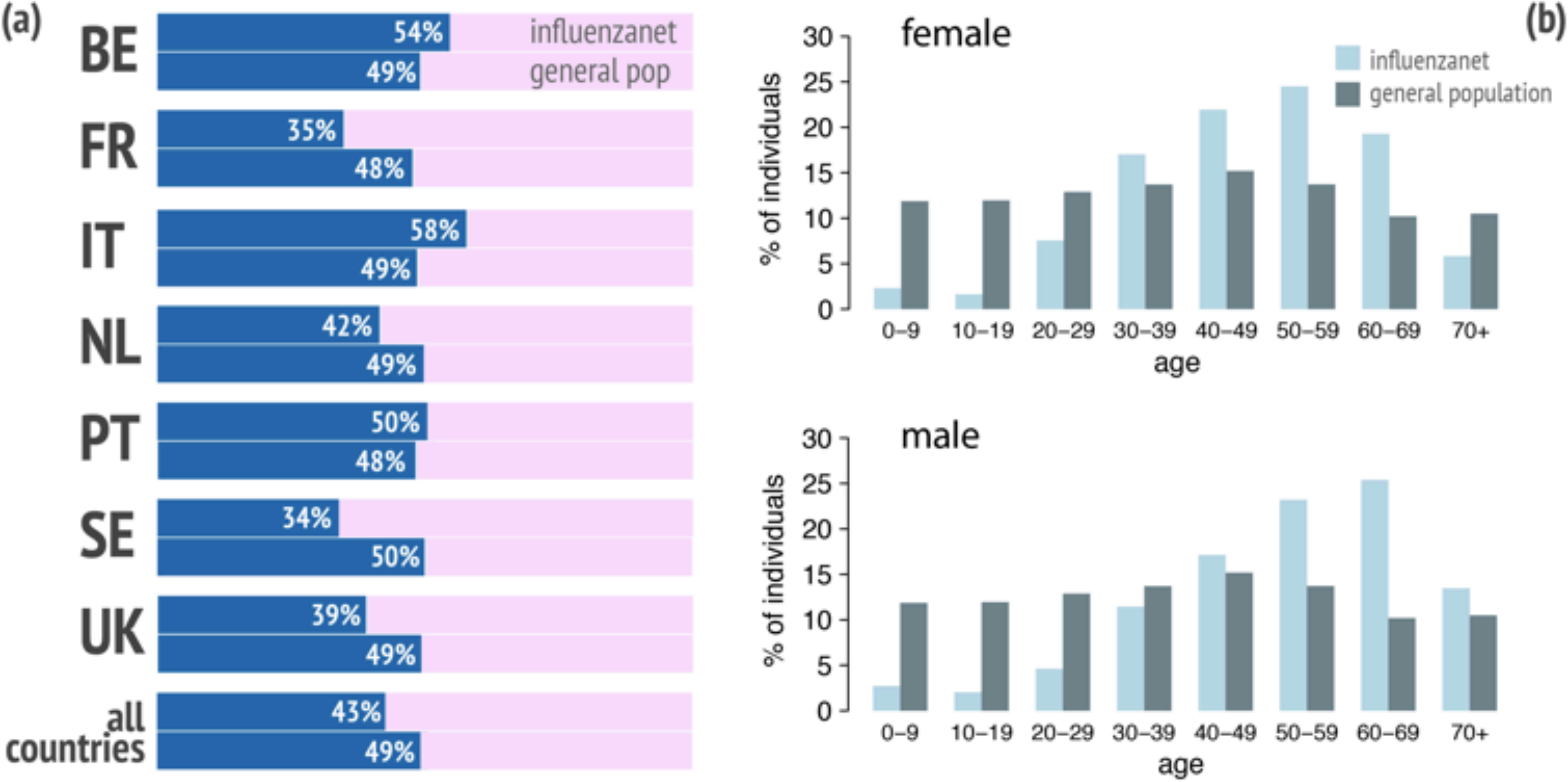
fr.grippenet.ch, de.grippenet.ch - Global Health Institute, Geneva

flusurvey.ie - University of Galway

grippeweb.rki.de - Robert Koch Institute, Germany



D. Paolotti et al., Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience, Clinical Microbiology and Infection, 20(1), 17 (2014)



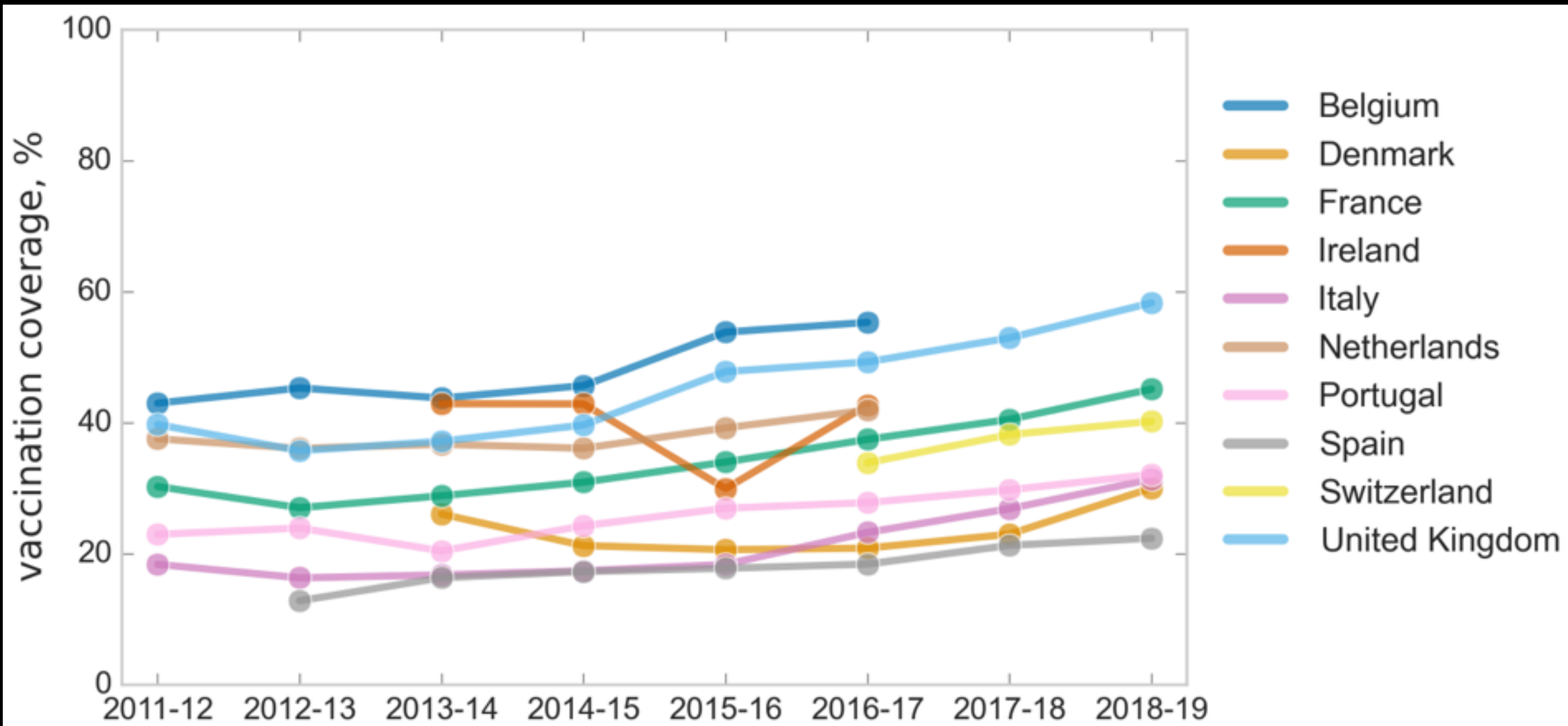
AGE / GENDER

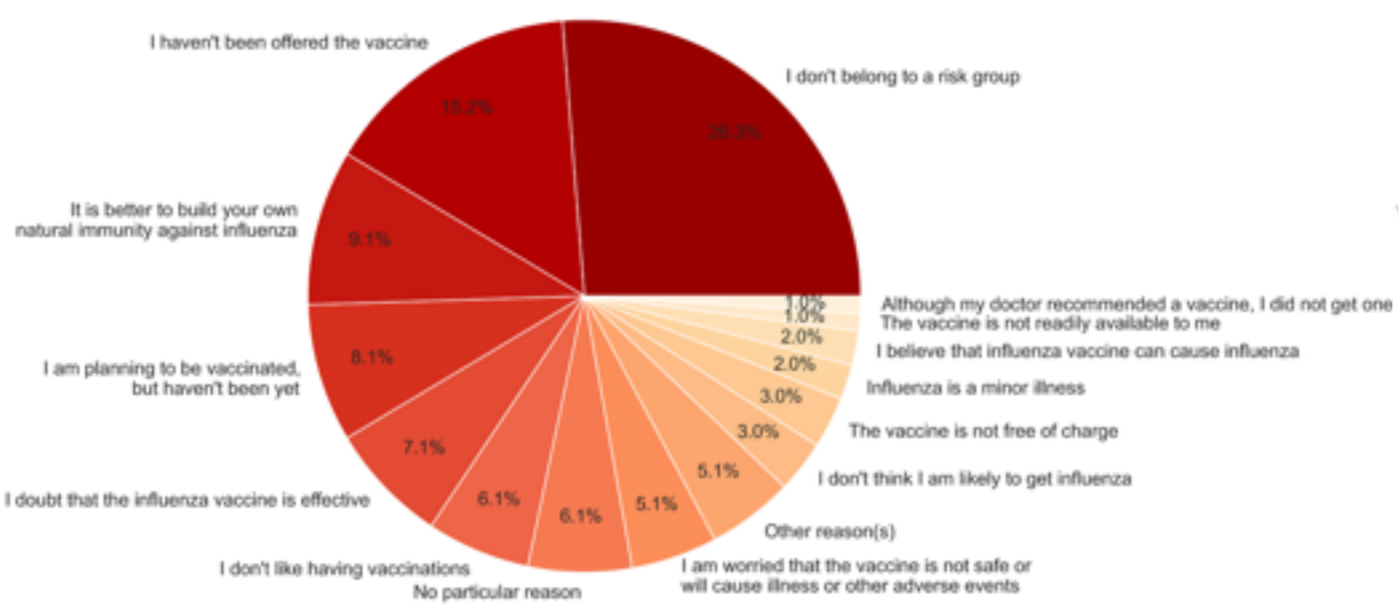
P. Cantarelli et al.,

The representativeness of a European multi-center network for influenza-like-illness participatory surveillance

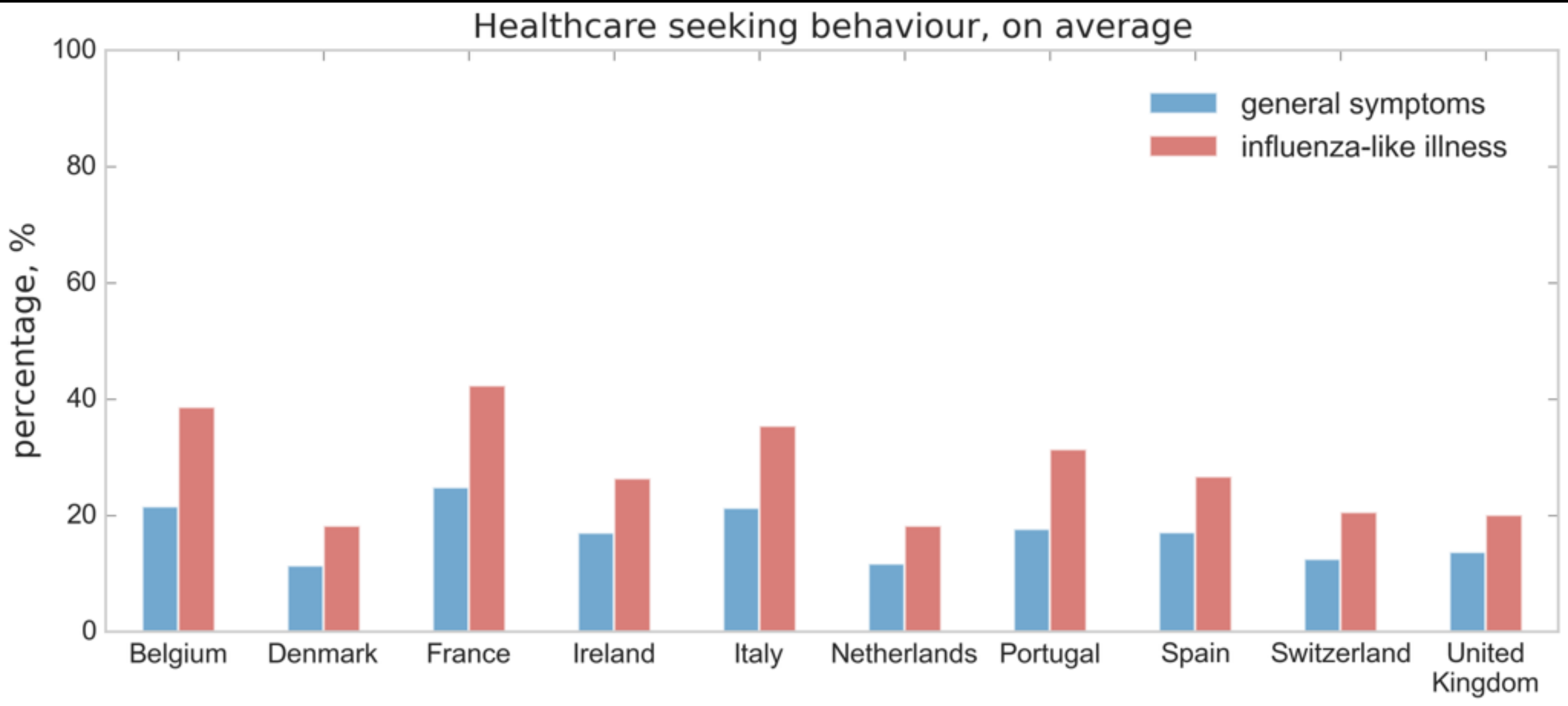
BMC Public Health 2014, 14:984

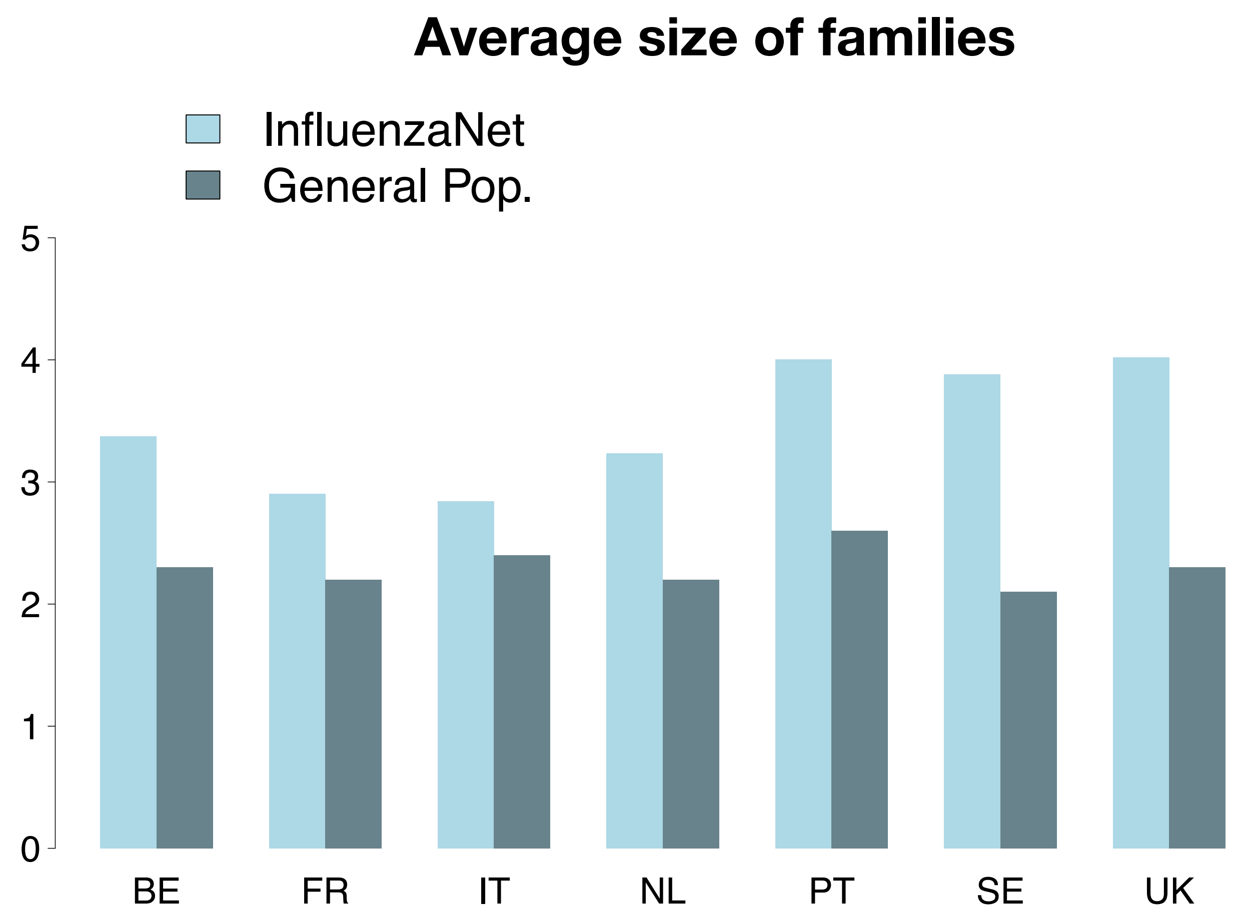
VACCINATION COVERAGE

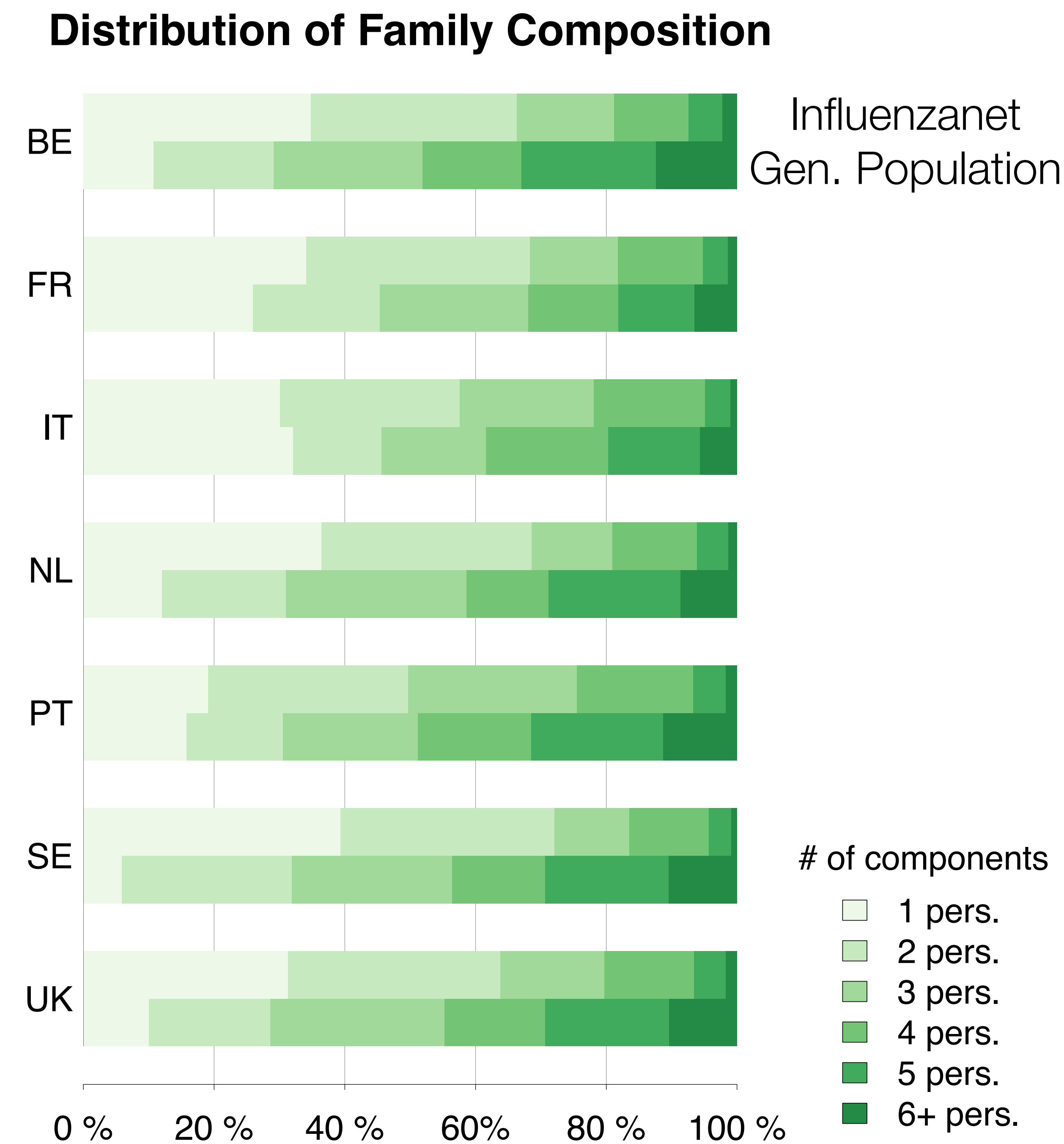




HEALTHCARE SEEKING







SOCIO-ECONOMIC INDICATORS

P. Cantarelli et al.,

The representativeness of a European multi-center network for influenza-like-illness participatory surveillance

BMC Public Health 2014, 14:984

	OR	CI	P
Gender	**1.52**	**1.33, 1.73**	**0.03**
Groups of children	**1.47**	**1.22, 1.77**	**0.03**
Unvaccinated	**1.94**	**1.66, 2.27**	**<0.001**
18-24	0.95	0.64, 1.42	0.80
25-34	1.12	0.81, 1.55	0.50
35-44	**1.42**	**1.04, 1.94**	**0.03**
45-64	**1.51**	**1.12, 2.03**	**<0.01**
65+	0.95	0.65, 1.38	0.79
Living w child	1.10	0.95, 1.28	0.31
Smoking	**1.32**	**1.07, 1.64**	**0.01**
Risk factor	**1.53**	**1.27, 1.83**	**<0.001**
Public transport	0.91	0.82, 1.15	0.73

RISK FACTORS FOR ILI (UK)

Adler A et al.,
Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey
BMC Infect Dis. 2014, 14:232

Countries	NL, BE, PT, IT[a]	UK[b]	FR[b]
Study period	2003–2013	2014–2015	2014–2015
Study participants per season, no., mean	24 666[c]	2629	4475
Female sex	1.22 (1.17–1.28)	1.25 (1.14–1.36)	1.12 (1.02–1.23)
Vaccinated	0.80 (.71–.91)[d]	S[e]	0.87 (.78–.97)
Age, y			
NL, BE, PT, IT			
<18	1.59 (1.46–1.74)	. . .	. . .
18–49	Reference	. . .	. . .
50–64	0.82 (.78–.86)	. . .	. . .
≥65	0.46 (.41–.51)	. . .	. . .
UK, FR			
0–14	. . .	1.27 (1.03–1.50)	0.95 (.76–1.16)
15–44	. . .	Reference	Reference
45–64	. . .	0.99 (.89–1.11)	0.87 (.76–0.98)
≥65	. . .	0.82 (.71–.94)	0.68 (.58–.79)
Children in household (vs living alone)	1.31 (1.22–1.40)	S[e]	S[e]
Contact with groups[f]	ND	1.11 (1.01–1.21)	1.12 (1.01–1.23)
Smoker	1.16 (1.10–1.22)	S[e]	S[e]
Underlying health condition			
Asthma[g]	1.58 (1.47–1.69)	. . .	. . .
Diabetes	1.27 (1.15–1.41)	. . .	. . .
Heart	1.29 (1.13–1.47)	. . .	. . .
Kidney	1.23 (.80–1.90)	. . .	. . .
Immune	1.23 (1.02–1.49)	. . .	. . .
Any	. . .	S[e]	1.17 (1.05–1.30)
Having respiratory allergies	ND	1.14 (1.05–1.24)	1.19 (1.07–1.29)
Declaring often having ILI	ND	ND	1.31 (1.17–1.45)
Sports participation for >1 h/wk	0.95 (.90–1.00)	ND	ND
Having pets			
Dogs	1.15 (1.09–1.22)	. . .	. . .
Cats	1.07 (1.02–1.12)	. . .	. . .
Any	. . .	ND	1.17 (1.08–1.28)
Daily transportation method			
Bike/foot	0.95 (.90–1.00)	NS[h]	S[e]
Car	Reference	. . .	. . .
Public	0.97 (.89–1.05)	. . .	. . .

C. Guerrisi et al, *Participatory Syndromic Surveillance of Influenza in Europe* Journal of Infectious Diseases (2016) 214 (suppl 4): S386-S392.

RISK FACTORS FOR ILI (multicountry)

grippenet.fr

Devenez acteur de la surveillance de la grippe

G grippenet

Objectifs

Estimer la fréquence de la grippe chez les femmes enceintes et le nombre de femmes vaccinées contre la grippe pendant leur grossesse en France métropolitaine.

G-GrippeNet est un projet de recherche fonctionnant sur le même principe que GrippeNet.fr, qui s'adresse spécifiquement aux **femmes enceintes**.

G-GrippeNet permet de participer activement à la surveillance de la grippe en France, de façon volontaire, anonyme et bénévole. Les données recueillies permettront de faire avancer la recherche sur la grippe afin de **mieux protéger les femmes enceintes et leurs bébés**.

Toutes les femmes enceintes résidant en France métropolitaine peuvent participer à G-GrippeNet, pour la durée qu'elles souhaitent. La participation à l'étude ne prend pas plus de 5 minutes par semaine. Le processus d'inscription est le même que pour Grippenet.fr, toutes les informations que vous trouverez sur ce site sont valables pour l'étude G-GrippeNet sauf mention contraire.

Influenzanet

Influenzanet is a system to monitor the activity of influenza-like-illness (ILI) with the aid of volunteers via the internet

flusurvey

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

i-sense

EPSRC IRC in Early Warning Sensing Systems for Infectious Diseases

Log in

Register

Forgot login details?

Virological Swabbing

This year at Flusurvey we're pleased to annouce that we are piloting taking swab samples from a subset of the participants to ascertain whether, once experiencing symptoms, they actually do have the flu virus or not.

Self-Swabbing for Virological Confirmation of Influenza-Like Illness Among an Internet-Based Cohort in the UK During the 2014-2015 Flu Season: Pilot Study

C. Wenham, E. Gray, C. Keane, M. Donati, D. Paolotti, R. Pebody, E. Fragaszy, R. McKendry, J. Edmunds

J Med Internet Res 20,3 (2018)

Why have I been chosen to take part in the swabbing sample?

You have been invited to take part in an additional flusurvey sub-study. This year, in addition to collecting information provided by you online relating to your symptoms during the flu season, we are asking a small sample of people to undertake a swabbing

What is the definition of Influenza-like illness?

Sudden onset of symptoms

AND

at least one of the following four systemic symptoms:
Fever or feverishness, Malaise, Headache, Myalgia

AND

at least one of the following three respiratory symptoms:
Cough, Sore throat, Shortness of breath

ECDC case definition

“What if the observed symptoms are the result of a superposition of latent syndromes characterised by an unknown incidence and and unknown composition in terms of symptoms?”

Weekly Symptoms Survey

1. Fever
2. Chills
3. Runny/blocked nose
4. Sneezing
5. Sore throat
6. Cough
7. Shortness of breath
8. Headache
9. Muscle/joint pain
10. Chest pain
11. Feeling tired (malaise)
12. Loss of appetite
13. Phlegm
14. Watery, bloodshot eyes
15. Nausea
16. Vomiting
17. Diarrhoea
18. Stomachache
19. Sudden Onset

boolean variables

time series of daily symptoms counts

$$\mathbf{X} = [x_{ij}]$$

matrix whose elements contains the occurrence of symptom j on day i

Latent Syndromes detection

- it is reasonable to expect that a specific combination of symptoms reported by a user is the symptomatic expression of one or more illnesses, i.e. syndromes, experienced by the user.

- In accordance with this consideration, we postulate that the time series x_{ij} of observed symptoms counts are the result of a linear mixing process driven by K unknown sources, corresponding to the latent syndromes we want to detect.

$$x_{ij} = \sum_{k \in \{1,..,K\}} w_{ik} \, h_{kj} + e_{ij}.$$

Latent Syndromes detection

The mixing equations can be expressed in matrix notation:

$$\mathbf{X} = \mathbf{WH} + \mathbf{E}$$

$$\mathbf{W} = [w_{ik}]\,,\ \mathbf{H} = [h_{kj}]\,,\ \mathbf{E} = [e_{ij}]$$

In this notation, the problem of detecting the unknown K latent sources can then be formulated as a matrix decomposition problem

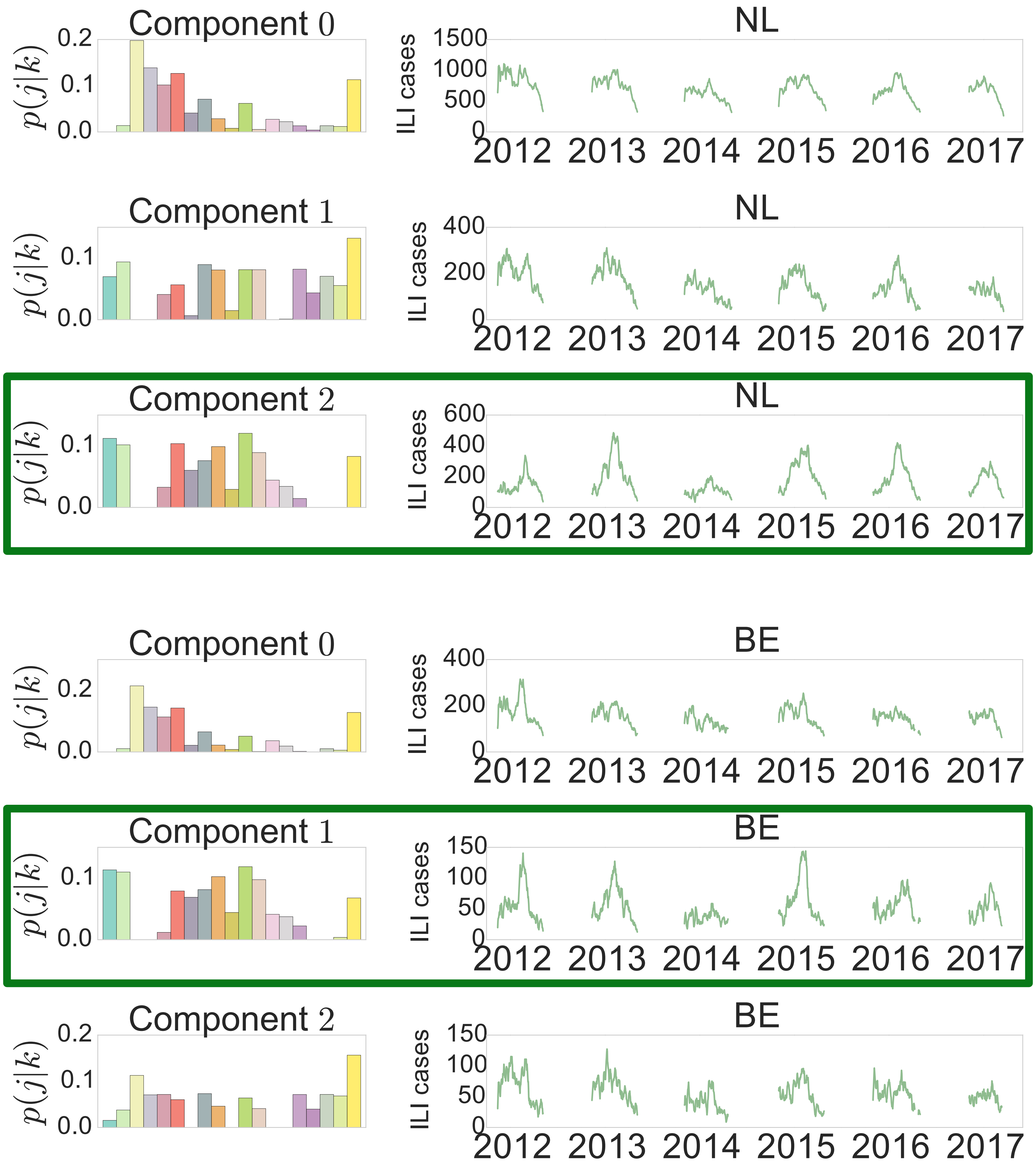
Non-negative matrix factorization

As a result of the probabilistic formulation, the total number of symptoms counts N will be proportionally split among K latent components according to p(k). p(j|k) describes each component in terms of the expected proportion of symptoms. According to this formulation, the total number of counts associated to a latent component k in day i will be given by:

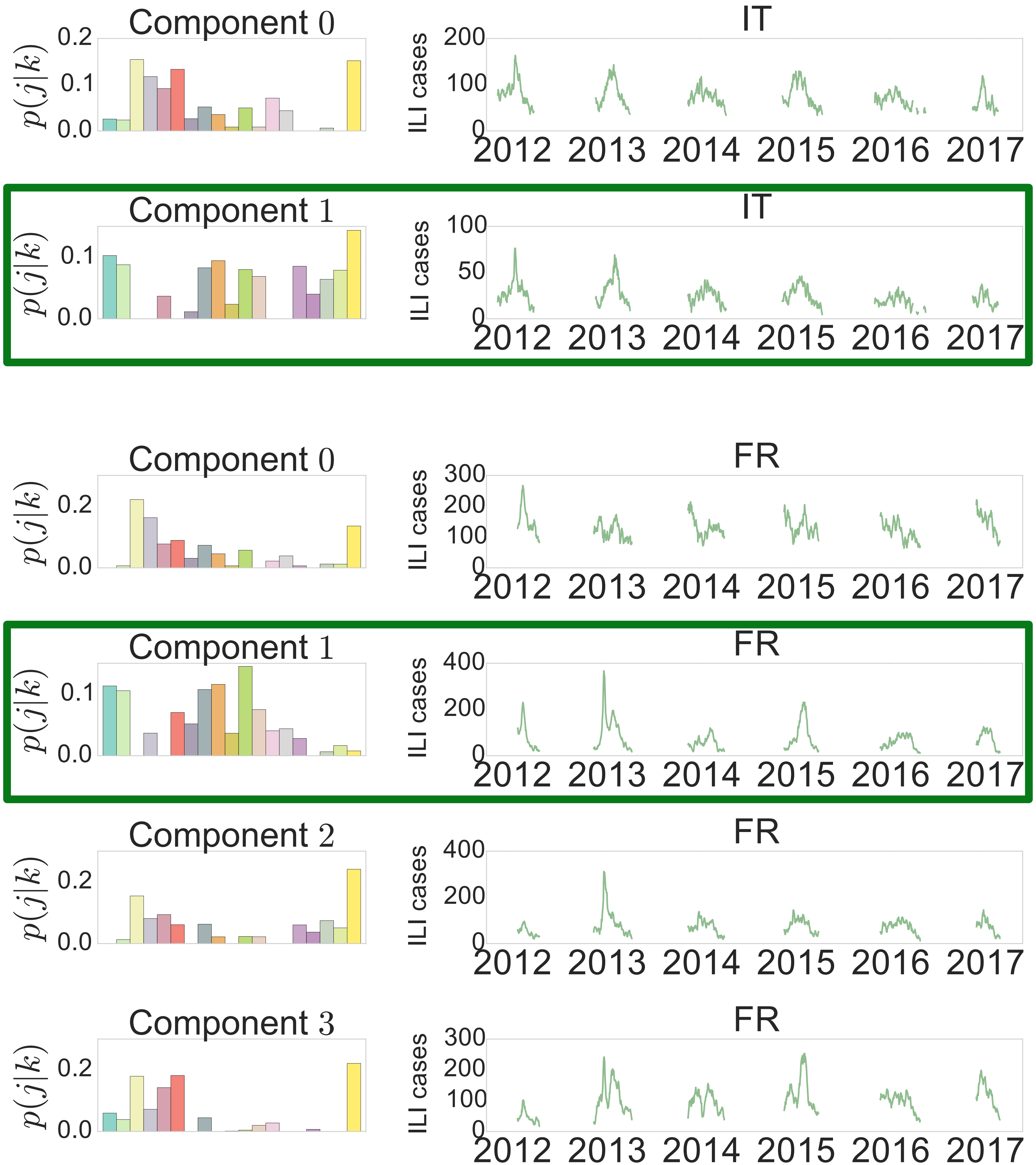
$$y_{ik} = N\,p(i,k) = N\,p(k)\,p(i|k)$$

Back to data

	NL	BE	IT	FR	UK	ES	PT	DK	IE
Number of seasons									
	6	6	6	6	6	5	6	4	4
Average participants per season									
	13, 450	4, 209	1, 830	5, 757	4, 676	526	1, 663	1, 391	406
Average # surveys per season									
	206, 987	67, 420	17, 807	68, 567	45, 543	5, 894	17, 852	22, 782	3, 220
Average % of surveys with symptoms									
	20%	16%	19%	20%	29%	22%	17%	18%	25%
Average of surveys per participant per season									
	15	16	9	12	9	11	10	16	8

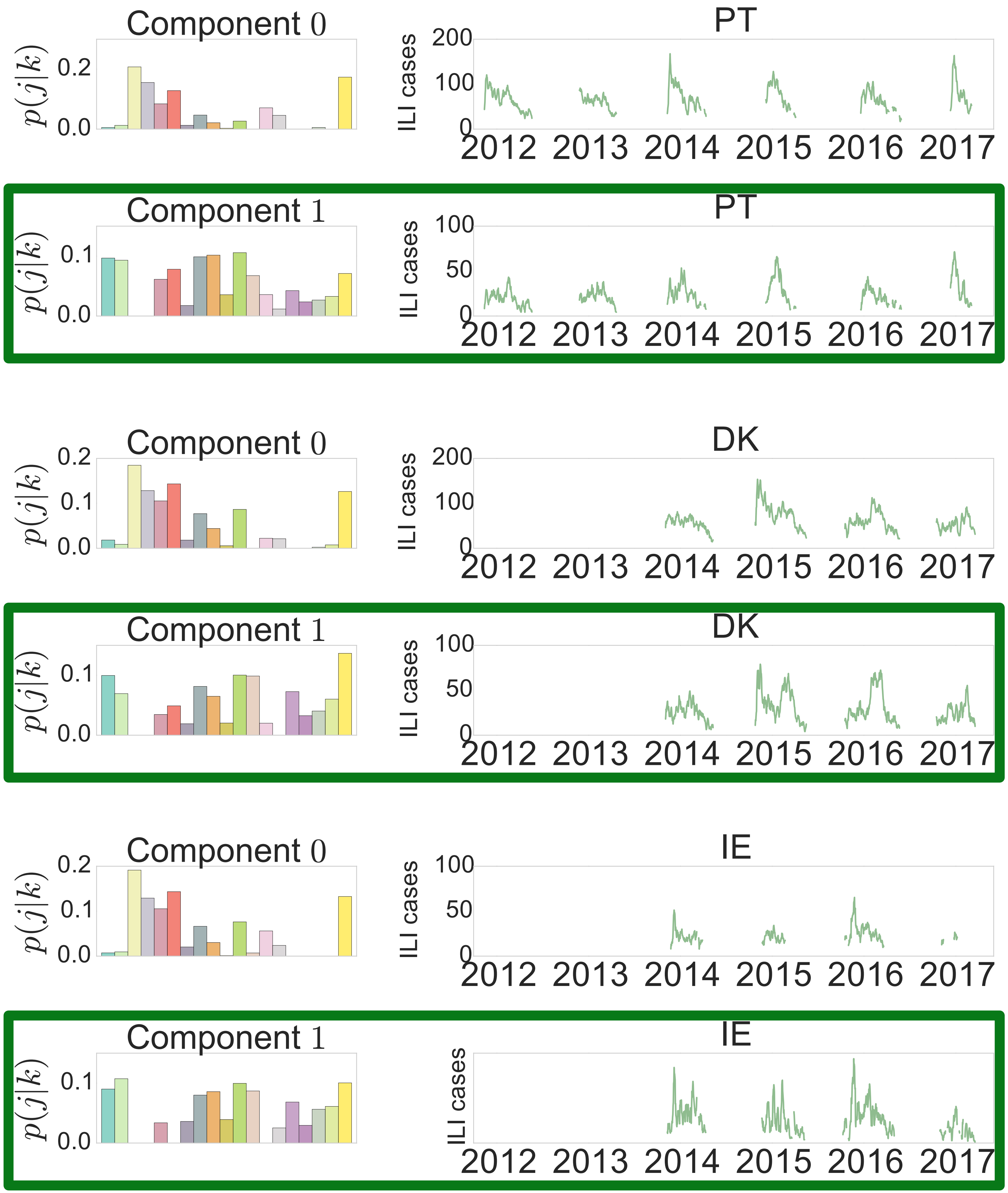


Component 0
$p(j|k)$
NL
ILI cases
2012 2013 2014 2015 2016 2017
Component 1
Component 2
BE

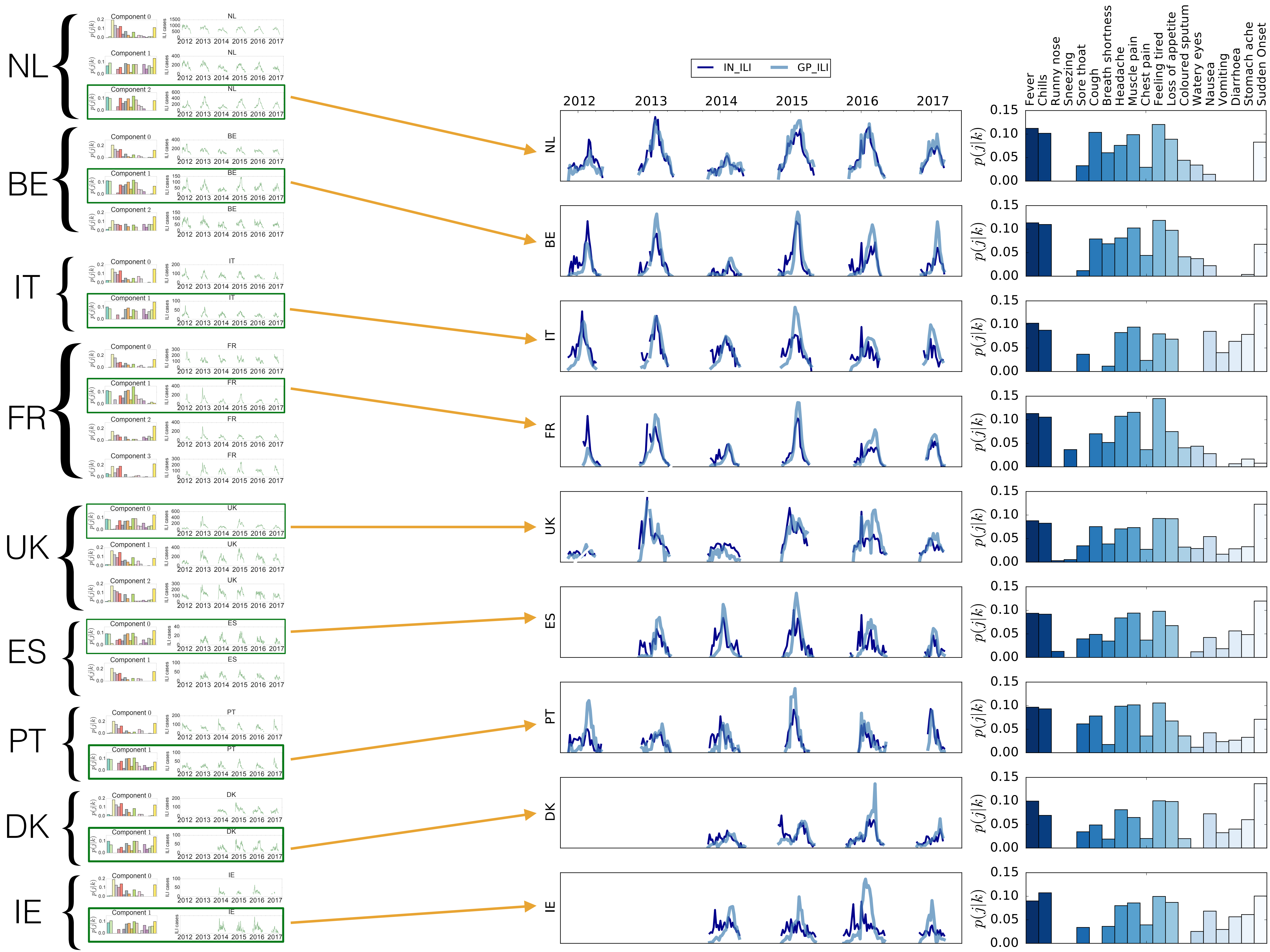


Component 0
$p(j|k)$
0.2
0.1
0.0
IT
ILI cases
200
100
0
2012
2013
2014
2015
2016
2017
Component 1
$p(j|k)$
0.1
0.0
IT
ILI cases
100
50
0
2012
2013
2014
2015
2016
2017
Component 0
$p(j|k)$
0.2
0.0
FR
ILI cases
300
200
100
0
2012
2013
2014
2015
2016
2017
Component 1
$p(j|k)$
0.1
0.0
FR
ILI cases
400
200
0
2012
2013
2014
2015
2016
2017
Component 2
$p(j|k)$
0.2
0.0
FR
ILI cases
400
200
0
2012
2013
2014
2015
2016
2017
Component 3
$p(j|k)$
0.2
0.0
FR
ILI cases
300
200
100
0
2012
2013
2014
2015
2016
2017


Component 0
UK
p(j|k)
0.1
0.0
ILI cases
600
400
200
0
2012 2013 2014 2015 2016 2017
Component 1
UK
0.2
0.1
0.0
400
200
0
Component 2
UK
0.2
0.1
0.0
300
200
100
0
Component 0
ES
0.1
0.0
40
20
0
Component 1
ES
0.2
0.0
100
50
0



Component 0
Component 1
$p(j|k)$
ILI cases
PT
DK
IE
0.0
0.1
0.2
0
50
100
200
2012
2013
2014
2015
2016
2017



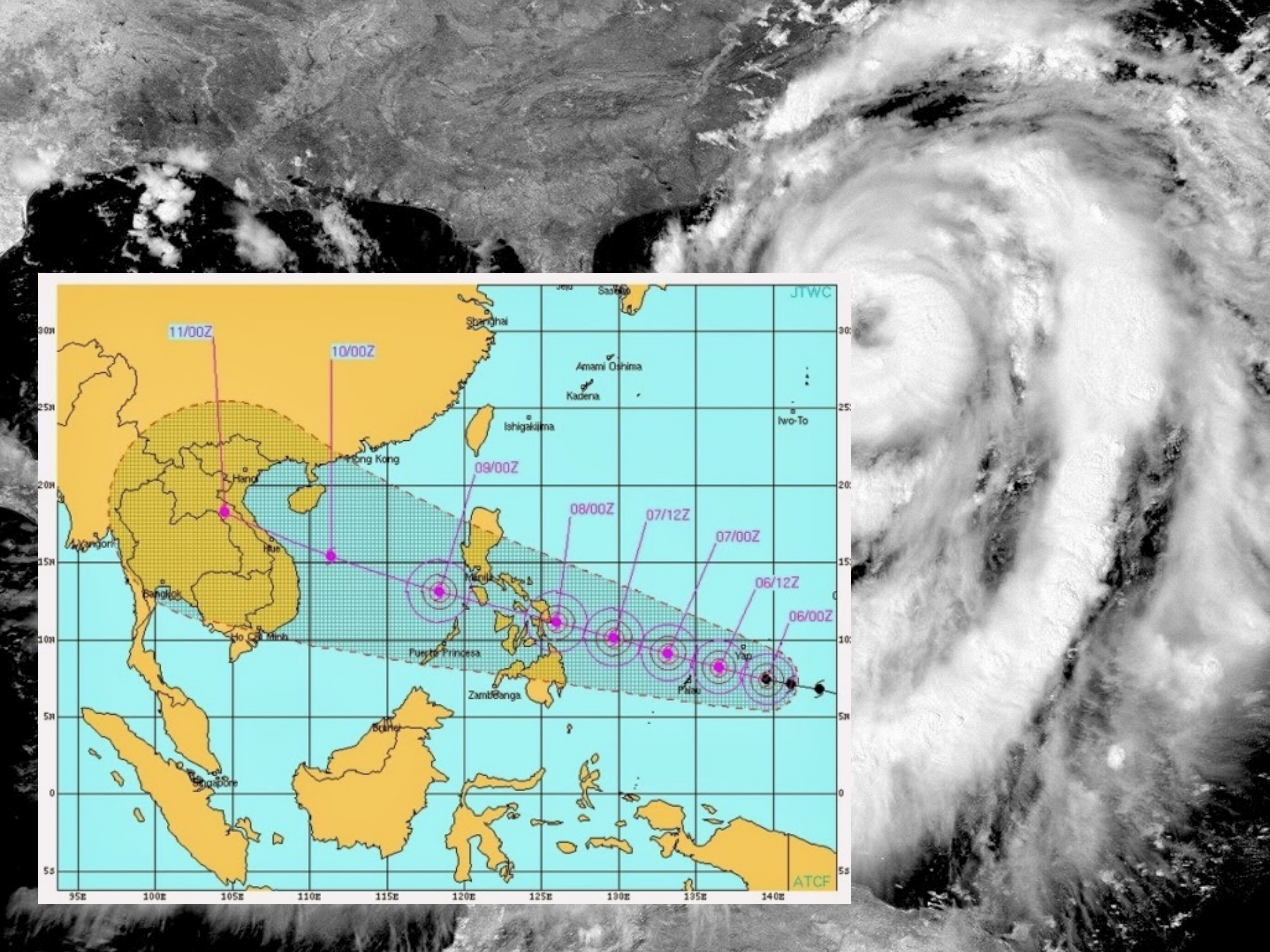
NL
BE
IT
FR
UK
ES
PT
DK
IE
Component 0
Component 1
Component 2
Component 3
p(j|k)
ILI cases
IN_ILI
GP_ILI
2012
2013
2014
2015
2016
2017
Fever
Chills
Runny nose
Sneezing
Sore thoat
Cough
Breath shortness
Headache
Muscle pain
Chest pain
Feeling tired
Loss of appetite
Coloured sputum
Watery eyes
Nausea
Vomiting
Diarrhoea
Stomach ache
Sudden Onset
0.00
0.05
0.10
0.15

	NL	BE	IT	FR	UK	ES	PT	DK	IE
(i) Correlation between IN_ILI_ECDC and IN_ILI for the seasons 2011-2017									
	0.91	0.92	0.86	0.83	0.92	0.86	0.84	0.90	0.82
(ii) Correlation between IN_ILI and GP_ILI for the seasons 2011-2017									
	0.88	0.80	0.69	0.79	0.74	0.65	0.66	0.71	0.38
(iii) Correlation between IN_ILI_ECDC and GP_ILI for the seasons 2011-2017									
	0.79	0.72	0.80	0.86	0.75	0.67	0.63	0.68	0.23
(iv) Correlation between IN_ILI prediction and GP_ILI for the season 2016-2017									
	0.85	0.82	0.69	0.80	0.60	0.84	0.80	0.76	0.60
(v) Correlation between between IN_ILI_ECDC and IN_ILI for the season 2016-2017									
	0.85	0.82	0.86	0.93	0.67	0.59	0.88	0.80	0.71

Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms, K. Kalimeri, M. Delfino, C. Cattuto, D. Perrotta, V. Colizza, C. Guerrisi, C. Turbelin, J. Duggan, J. Edmunds, C. Obi, R. Pebody, A. O Franco, Y. Moreno, S. Meloni, C. Koppeschaar, C. Kjelsø, R. Mexia, D. Paolotti, PLOS Computational Biology 15 / 4 (2019)

What’s next - I?

- Virological confirmation is needed to estimate more accurately the scaling factor
- extension of the method to other countries and syndromes
- Assess the validity of the methods for detection of new emerging diseases


JTWC
ATCF
11/00Z
10/00Z
09/00Z
08/00Z
07/12Z
07/00Z
06/12Z
06/00Z
Shanghai
Amami Oshima
Kadena
Ishigakijima
Iwo-To
Hong Kong
Hue
Manila
Puerto Princesa
Zamboanga
Yap
Palau
Singapore
95E
100E
105E
110E
115E
120E
125E
130E
135E
140E

Forecasting seasonal influenza

Baseline Model: forecasting influenza through sentinel doctors reports

$$y_{w+k} = \alpha_1 \widetilde{y}_{w-1} + \alpha_2 \widetilde{y}_{w-2} + \alpha_3 \widetilde{y}_{w-3}$$

y_w denotes the ILI incidence value reported by the sentinel doctors at week w (k=0 nowcasting, k>1 forecasting)

Participatory Surveillance + sentinel doctors reports

Linear Autoregressive exogenous model: ARX1w, ARX4w

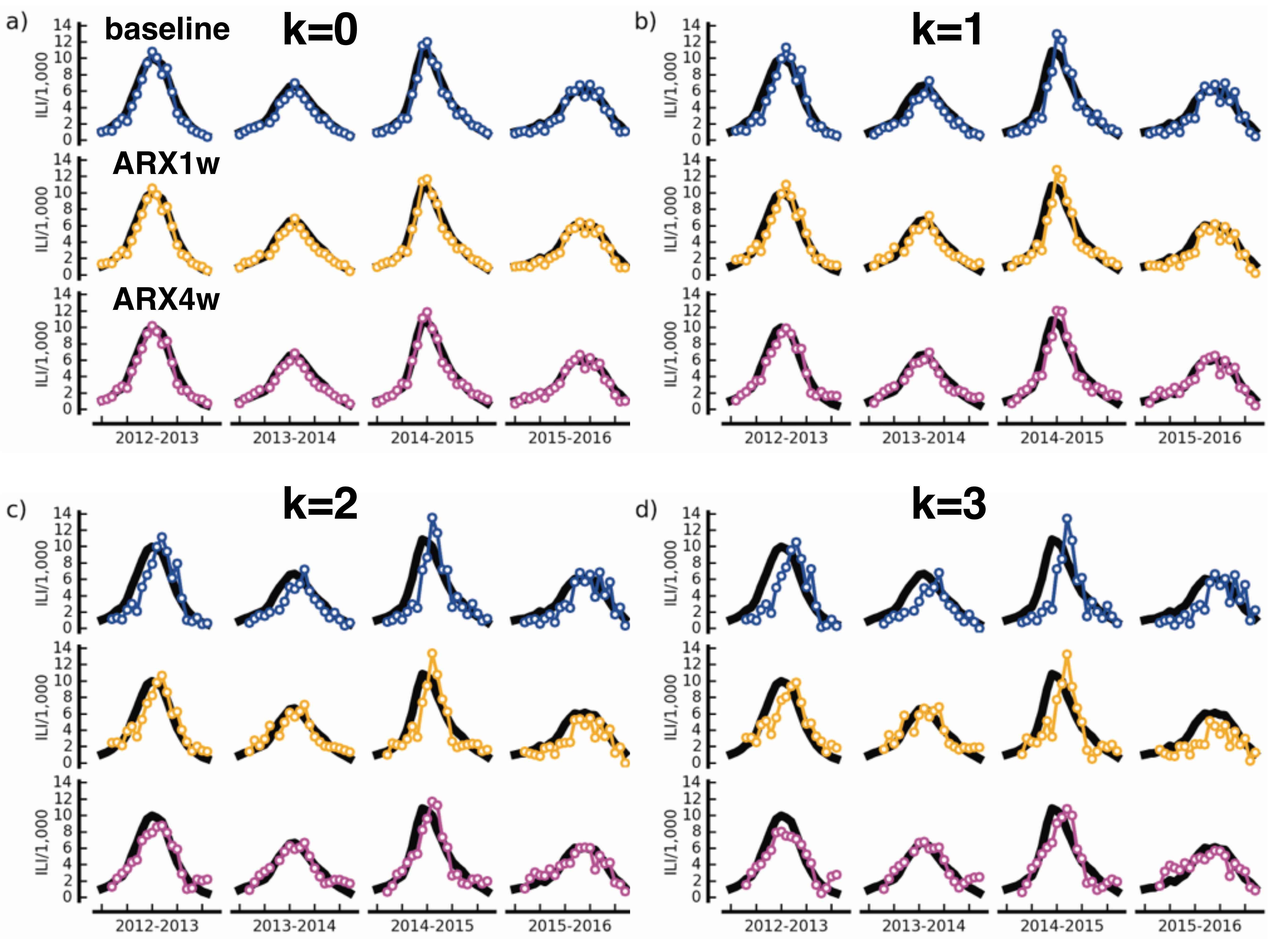
$$y_{w+k} = \sum_{i=1}^{3} \alpha_i \widetilde{y}_{w-i} + \gamma_0 \widetilde{z}_w$$

$$y_{w+k} = \sum_{i=1}^{3} \alpha_i \widetilde{y}_{w-i} + \sum_{j=0}^{3} \gamma_j \widetilde{z}_{w-j}$$

INFLUENZANET in Italy: influweb.it

Participation to Influweb during the five influenza seasons under study.

Season	Number of registered users	Average number of weekly active users	Total number of symptoms surveys
2011-2012	2,270	1,085	14,681
2012-2013	3,057	1,165	15,789
2013-2014	3,513	1,244	18,471
2014-2015	3,753	1,290	20,224
2015-2016	4,053	1,224	20,823


a)
baseline
k=0
ARX1w
ARX4w
b)
k=1
c)
k=2
d)
k=3
ILI/1,000
2012-2013
2013-2014
2014-2015
2015-2016

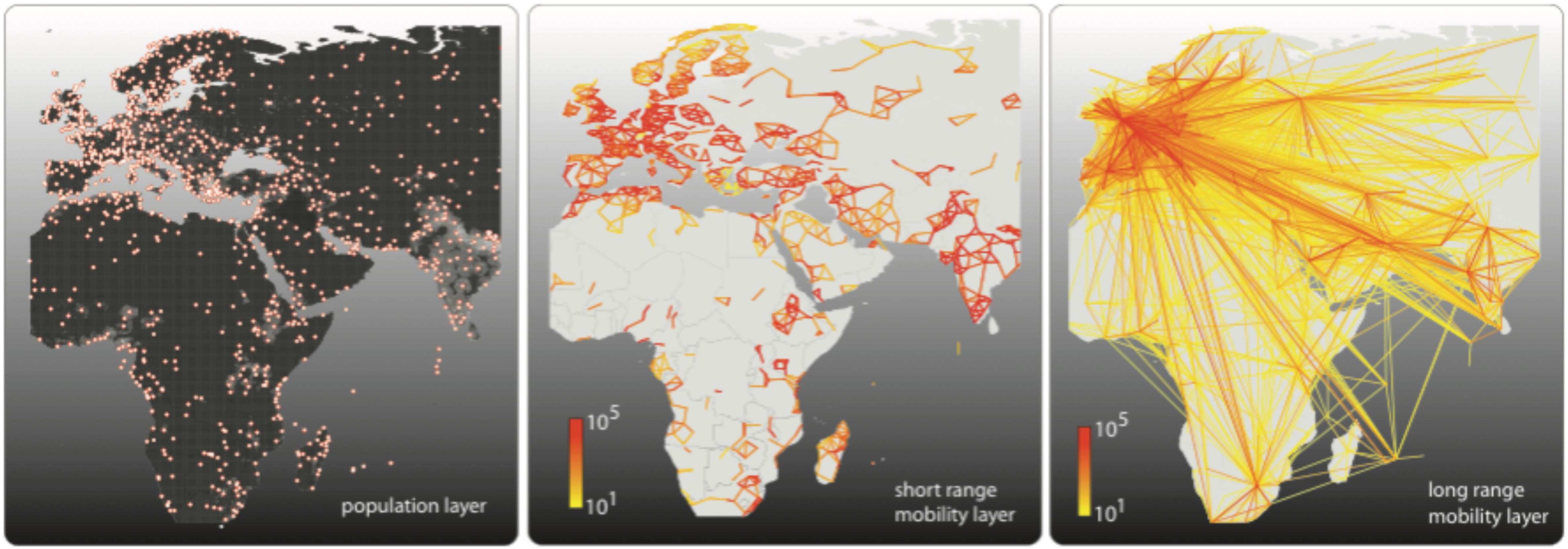
Similarity metrics and peak analysis of the forecasting models with respect to the ground truth. The best performing model per metric is bold faced.

	Model	CORR	MAE (ILI/1,000)	RMSE (ILI/1,000)	MAPE (%)	weeks lag [Min, Max]	PE (%) Mean [Min, Max]
		Similarity Metrics				Peak Analysis	
$k = 0$	baseline	0.981	0.45	0.5955	**12.93**	[0, 1]	8.6 [4.8, 10.9]
	ARX_{1w}	0.983	0.42	0.5419	12.95	[0, 1]	5.5 [3.5, 7.7]
	ARX_{4w}	**0.984**	**0.39**	**0.5052**	13.13	[0, 1]	5.6 [1.9, 9.0]
$k = 1$	baseline	0.934	0.83	1.1174	**22.39**	[1, 2]	13.7 [8.9, 19.3]
	ARX_{1w}	0.942	0.73	0.9635	24.31	[1, 2]	9.3 [0.8, 18.1]
	ARX_{4w}	**0.957**	**0.63**	**0.7952**	25.63	[0, 1]	5.5 [0.7, 10.6]
$k = 2$	baseline	0.835	1.29	1.7497	33.79	[2, 3]	13.8 [7.7, 24.3]
	ARX_{1w}	0.864	1.05	1.4488	**32.72**	[2, 3]	11.7 [6.5, 23.0]
	ARX_{4w}	**0.918**	**0.84**	**1.0885**	34.62	[0, 2]	5.2 [0.3, 12.0]
$k = 3$	baseline	0.677	1.80	2.4287	45.23	[0, 3]	9.7 [2.3, 23.2]
	ARX_{1w}	0.751	1.42	1.9404	43.23	[0, 3]	10.7 [1.7, 22.1]
	ARX_{4w}	**0.879**	**1.02**	**1.3058**	**42.93**	[0, 3]	7.2 [0.6, 19.7]

MAE = Mean Absolute Error
RMSE = Root Mean Square Error
MAPE = Mean Absolute Percent Error

proceedings of WWW2017: D. Perrotta, M. Tizzoni, D. Paolotti, “Using Participatory Web-based Surveillance Data to Improve Seasonal Influenza Forecasting in Italy”

THE MODEL: GLEAM



Commuting

Air travel

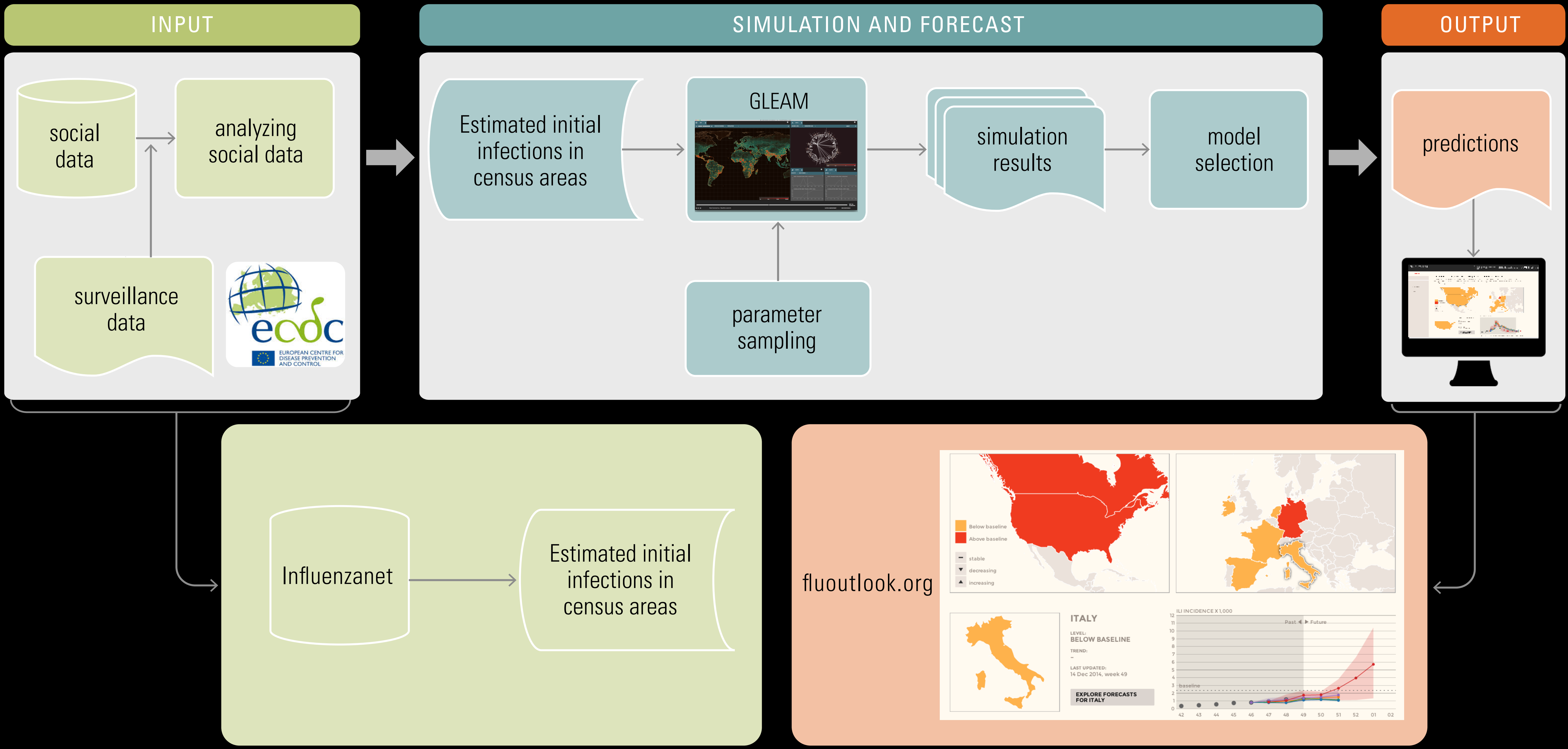
Population layer

Human mobility layers

Geographic scale

Balcan et al. PNAS 2009

Forecasting seasonal influenza: fluoutlook.org



Social Data Mining and Seasonal Influenza Forecasts: The FluOutlook Platform
Q. Zhang, C. Gioannini, D. Paolotti, N. Perra, D. Perrotta, M. Quaggiotto, M. Tizzoni, A. Vespignani
Lecture Notes in Computer Science 9286 (2015) p.237-240

Fluoutlook is a web platform for the exploration of influenza forecasts

It provides a visual interface to numerical forecasts of the current influenza season in North America and Europe through maps and charts. Activity data and forecasts are updated weekly, based on the reports of the official influenza surveillance systems in each country. **Tell me more.**

NORTHERN EMISPHERE | SOUTHERN EMISPHERE

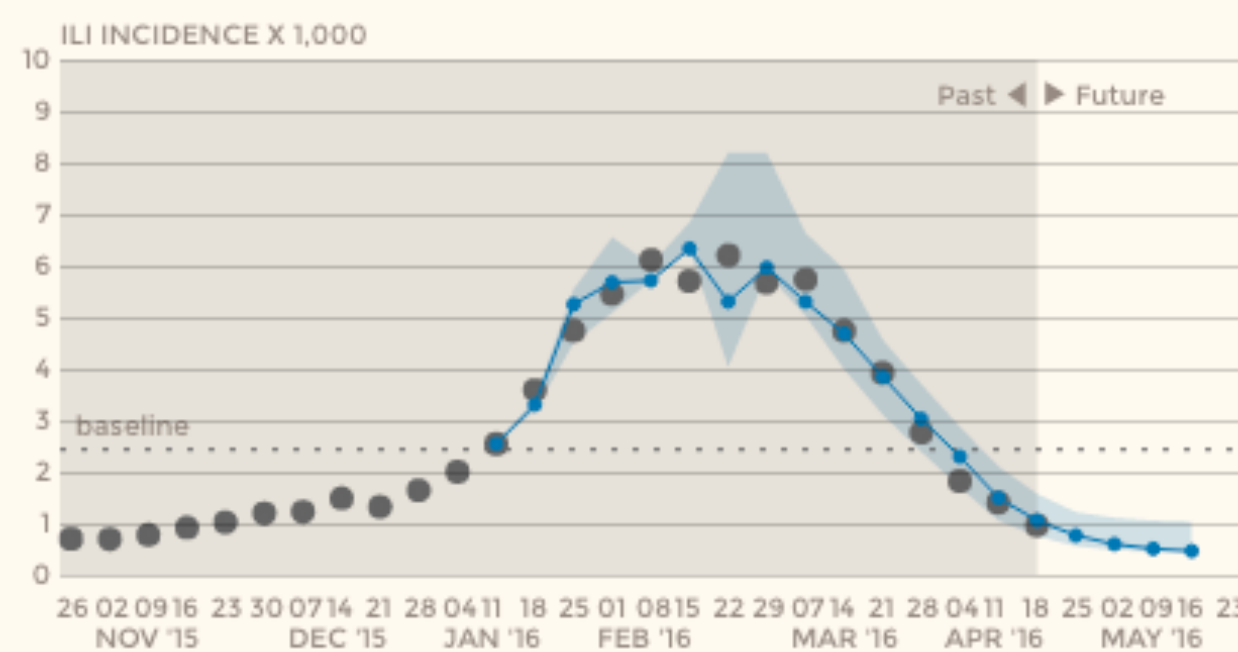
ITALY

LEVEL:
BELOW BASELINE

TRENDː
–

LAST UPDATED:
18 Apr 2016, week 16

EXPLORE FORECASTS FOR ITALY



Conclusions - 1

In summary, Influenzanet is a well-established standardized participatory surveillance system for ILI in Europe, covering more than one third of the 28 EU member states. Its strength lies in:

- (1) the **standardized technological and epidemiological framework** for a coherent surveillance across countries and the centralized fashion that characterize the technological deployment (except for France and Sweden, all the platforms have been deployed and are now managed by ISI Foundation in Turin, Italy).

- (2) the ability to timely **monitor ILI in the general population**, including individuals who do not seek medical assistance;

- (3) its **sensitivity** in detecting substantial changes in population health earlier than GP sentinel networks;

- (4) its potential **scalability** to large numbers with rather **limited costs**;

- (5) its **flexibility** in exploring different ILI definitions;

- (6) the detailed profile data allowing **individual-level epidemiological** analyses generally not possible in standard systems;

- (7) its potential extension to **other diseases**.

What's next - II?

- Participatory surveillance in developing countries
- Behavioural changes
- Other health-related issues: pregnant women and vaccination, chronic conditions, vaccine hesitancy
- Validation of passive web data: linking behavioural data from participatory systems with on-line data

ECDC/WHO bulletin: indicators

- Weekly number of active users
- Weekly incidence (with various ILI case definition)
- % of ILI cases in contact with healthcare
- cumulative % of vaccinated by key target group

Conclusions - 2

- Its limitations are mainly due to the self-selected sample, potential misreporting, and lack of validation by a physician or by virological testing. However, the agreement found with GP incidence trends suggests that these limitations have little effect once results are adjusted for lack of representativeness.

- Main challenges remain the baseline maintenance resources to sustain the system in the long run and the recruitment and retention of participants. While the identification of sociocultural determinants for participation will provide additional insights, the strong willingness for engagement found in most countries' participants confirms the feasibility of the approach.

- Moreover, the platform represents a crucial channel for communication with the public, to inform and increase awareness, an increasingly important aspect after the 2009 pandemic.

INFLUENZANET.INFO

Daniela Paolotti (IT)

Vittoria Colizza, Clement Turbelin, et al (FR)

Yamir Moreno (ES)

Ricardo Mexia (PT)

Richard Pebody, Chinelo Obi (UK)

Edward Van Straten (SE)

Charlotte Kjelsø (DK)

Jim Duggan (IE)

Antoine Flahaut (CH)

Udo Bucholz (DE)

influweb

grippenet.fr

gripenet

flusurvey

www.influensakoll.se
Kartlägger förekomst och spridning av influensa i Sverige

influmeter.dk
Kortlægger forekomst og spredning af influenza i Danmark

flusurvey.ie

grippenet

GrippeWeb

- *C. Gioannini*
- *D. Perrotta*
- *M.Quaggiotto*
- *L. Rossi*
- *M. Tizzoni*
- *M. Delfino*
- *K. Kalimeri*
- *C. Cattuto*

ISI
ISI Foundation
& ISI Global Science Foundation

- *N. Perra*
- *N. Samay*
- *Q. Zhang*
- *A. Vespignani*

MOBS LAB
LABORATORY FOR THE MODELING OF BIOLOGICAL AND SOCIO-TECHNICAL SYSTEMS

Northeastern

Thank you!

Daniela Paolotti
@danielapaolotti