Spatio-temporal Analysis of Flu-related Drugs Uptake in an Online Cohort in England.

Daniela Perrotta
ISI Foundation
Turin, Italy
daniela.perrotta@isi.it

Dario Delle Vedove
University of Turin
Turin, Italy
dario.dellevedove@edu.unito.it

Chinelo Obi
Public Health England
London, United Kingdom
Chinelo.Obi@phe.gov.uk

Richard Pebody
Public Health England
London, United Kingdom
Richard.Pebody@phe.gov.uk

Rossano Schifanella
University of Turin
ISI Foundation
Turin, Italy
rossano.schifanella@unito.it

Daniela Paolotti
ISI Foundation
Turin, Italy
daniela.paolotti@isi.it

ABSTRACT

The effective monitoring and control of disease outbreaks and epidemics rely on accurate and timely data, including the number of disease cases as well as the amount of medicines needed to alleviate the burden of the disease in the general population. Official public health sources of information, despite being reliable and accurate, often fail to be delivered in a timely manner. On the other hand, participatory Web-based monitoring systems, which rely on the participation of self-selected volunteers, can help complement traditional public health practices and overcome these issues.

In this study, we investigate the spatio-temporal patterns of flu-related drugs uptake in England, as measured by the Flusurvey platform, which is the largest crowd-sourced Web platform for the monitoring of influenza-like illness activity in United Kingdom. Flu-related drugs prescriptions reported by the National Institute of Health in England represent our ground truth. We retrospectively evaluate the performance of self-reported data collected by Flusurvey over the course of four influenza seasons, from 2014-2015 to 2017-2018. Our results show a high temporal correlation (ranging from 0.60 to 0.96) between the prescriptions data and the Flusurvey time series for antibiotics and cough medications. The spatial correlation between the two datasets is instead not statistically significant. In conclusion, Web-based monitoring systems such as Flusurvey, can capture the temporal patterns of flu-related drugs consumption in the general population and help deliver this information to public health authorities in a more timely fashion than traditional systems.

CCS CONCEPTS

• **Applied computing** → *Health informatics.*

DPH '19, November 20–23, 2019, Marseille, France

ACM ISBN 978-1-4503-7208-4/19/11…$15.00
https://doi.org/10.1145/3357729.3357739

KEYWORDS

Drugs uptake, Prescriptions, Web-based surveillance, Participatory surveillance, Influenza-like illness, Flu

ACM Reference Format:
Daniela Perrotta, Dario Delle Vedove, Chinelo Obi, Richard Pebody, Rossano Schifanella, and Daniela Paolotti. 2019. Spatio-temporal Analysis of Flu-related Drugs Uptake in an Online Cohort in England.. In *Proceedings of the 9th International Digital Public Health Conference (2019), Nov. 20–23, 2019, Marseille, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3357729.3357739

1 INTRODUCTION

Seasonal influenza epidemics occur annually during winter seasons in temperate regions with an estimated annual attack rate of 3 to 5 million cases of severe illness and around 250 to 500 thousand deaths worldwide each year [2]. Timely and effective surveillance systems are therefore crucial to monitor the circulation of influenza in the population and promptly allocate public health resources to implement adequate prevention strategies and mitigate the impact of particularly severe outbreaks of influenza. Next to national public health surveillance infrastructures, in recent years participatory Web-based systems have been used more and more to provide an additional layer of surveillance of influenza activity by collecting information directly from a cohort of individuals who self-report their health status through Internet-based surveys [28]. Similar systems have been implemented independently in various parts of the globe, such as Influenzanet in Europe [22], Flu Near You in United States [27] and Flutracking in Australia [12].

Here we focus on Influenzanet [3], a network of participatory surveillance systems established in 2009 with the aim of monitoring influenza-like illness (ILI) epidemics in Europe. Previous studies have extensively shown the uniqueness of Influenzanet as an innovative digital tool to complement traditional data, with a wide variety of applications in public health, including the monitoring of ILI incidence rates [22, 25, 31], the estimation of risk factors for ILI [4, 14], age-specific influenza attack rates [23, 25], and influenza vaccine effectiveness [13, 16], the assessment of health care seeking behavior [24, 30], and the unsupervised inference of ILI syndromes [15]. In addition to the monitoring task, previous studies have also shown the potential impact of employing Influenzanet data prospectively, thus further highlighting the added value provided by such digital tool to support the existing traditional practices in public health [7, 26, 32].

Indeed, such studies have been allowed thanks to an adequate data collection that has been properly designed to draw a high-level picture of the participants and the determinants of health. Firstly, after registering to the system, participants are prompted with an intake questionnaire, which ask for some demographic factors, such as age, gender, location (first part of postcode), level of education, household composition, influenza vaccine status, and pre-existing health conditions. Subsequently, through a weekly e-mail newsletter, participants are invited to complete a symptoms survey in which they are asked to report their health status, i.e. whether they are in good health or have experienced an episode of illness. In the latter case, a set of follow-up questions is asked, including onset and duration of symptoms, health-seeking behavior, changes in daily routine and drugs uptake.

In this study, we leverage this latter self-reported information to investigate the performance of participatory surveillance in capturing the spatio-temporal usage of influenza-related drugs as compared to the official data of drugs prescription reported by public health authorities. Remarkably, participatory surveillance systems monitor the ILI activity from symptomatic individuals who do not necessarily seek medical attention for their symptoms, thus representing the ideal tool to assess drugs consumption in the general population as a further approach to monitor and control disease epidemics. In particular, here we focus on the Flusurvey platform [1], launched in 2009 in the United Kingdom as part of the Influenzanet network, and we compare the self-reported data collected by Flusurvey against the reference data reported by the National Health Service (NHS) [21]. This system is responsible for collecting, processing and publishing data and information from all health and social care systems across England. Notably, this type of study is feasible limited to the country of England for which prescriptions data are publicly available at the practice level. Prescriptions data have been adopted in several studies to explore long-term temporal trends and have been proved to be a valuable source to observe changes in practice, to provide feedback, to ensure there are no unexpected or undesirable changes, and to facilitate tracking and forecasting of costs [10]. A wide body of literature has focused on several distinct categories of pharmaceuticals, such as opioids [9, 18], antidepressants [29], antibiotics [11, 17], and antipsychotics [5].

In this paper, we identify four categories of drugs related to influenza-like illness, namely antibiotics, antivirals, cough medications, and painkillers, and we retrospectively evaluate their consumption as reported by Flusurvey participants in England over the course of four influenza seasons, from 2014-2015 to 2017-2018. Below we provide a detailed description of the two datasets we have used in this study, from raw data to processed data in the form of monthly drugs rates. Then we present the results of our analysis in terms of accuracy of the self-reported data collected by Flusurvey in capturing the spatio-temporal trend of flu-related drugs consumption in England, both nationally and in the 9 regions of England. We conclude by discussing our findings, also in relation to the advantages and limitations of both sources of data, and proposing future work to help advance research in this field.

2 DATA

In this section, we present the two sources of data we have used to study the spatio-temporal trend of flu-related drugs uptake in England. First we describe the self-reported data collected by the Flusurvey platform and then the prescriptions data reported by general practitioners and collected by the NHS system.

2.1 Flusurvey

The Flusurvey platform [1] was launched in the United Kingdom in 2009, under the umbrella of the Influenzanet network [3], to help monitor the influenza-like illness activity with the aid of self-selected volunteers from the general population. Generally, the data collection runs approximately from November to May in order to cover the period of highest activity of influenza and it is usually advertised at the beginning of each influenza season in order to recruit new participants. Participation is voluntary and anonymous, and open to all individuals living in the country.

Upon registration, participants are invited to complete an intake questionnaire and provide some general information on their demographic and medical background, including age, gender, geographical location, educational and socio-economic status, employment, household composition, use of public transport, vaccination status and pre-existing health-conditions. This intake questionnaire can be then updated at any time, if necessary, for example to update the vaccination status. Participants can also create accounts on behalf of other members of their household, such as children or elderly people, and record data on their behalf.

Once enrolled in the system, participants are invited weekly, via an e-mail newsletter, to complete a symptoms survey in which they are asked whether they experienced any general, respiratory or gastrointestinal symptoms since the last time they visited the platform. If symptoms are reported, further questions are asked in order to assess both the syndrome and the behavior of the participants, detailing the onset of symptoms, body temperature, health-seeking behavior (i.e. if they have visited or contacted a primary care because of their symptoms), changes in daily routine (i.e. if they stayed at home from school or work) and medicine uptake.

Data collected by the Flusurvey platform are gathered in accordance with the European privacy legislation which establishes that only aggregated and anonymized data can be published and shared. Raw data are available upon request from third parties wishing to conduct scientific research and upon discussion with other members of the Influenzanet Scientific Committee.

2.2 Prescriptions

Prescriptions are made publicly available since August 2010 by the National Health Service (NHS) [21] in the form of a list of all medicines, dressings and appliances prescribed each month by general practitioners (GPs) in England.

For each medicine, dressing and appliance, we collect information on the GP where the prescription was issued, the presentation name and a code, called BNF code, which uniquely identifies the item as a dressing or an appliance in the UK health system. It is worth noting that a prescription item refers to a single supply of a medicine, dressing or appliance listed on a prescription form. It is a different concept from the quantity of units, e.g. pills, ampules,

Table 1: Description of the Prescriptions data.

Attribute	Description
BNF code	15 characters code identifying a medicine, dressing or appliance and describing its hierarchical information (see paragraph *Medicine taxonomy*).
BNF name	Individual preparation name, which may be proprietary, i.e. a name associate to a brand, or generic, i.e. not related to any specific firm. A medicine is followed by format and formulation (active substance and its amount).
Practice	6 characters (1 letter + 5 numbers) code identifying a practice dispensing the prescription.
Items	A prescription item refers to a single supply of a medicine, dressing or appliance prescribed on a prescription form. If a prescription form includes N pharmaceuticals, it is counted as N distinct prescription items.
Date	Year and month in which the prescription was written, reported as '*yyyymm*'. For example, June 2012 would be 201206.

or tablets, relative to a single prescription and it does not provide details on the intensity of the active ingredients.

The monthly consumption in units of product and the reference time frame are also provided. Additionally, we collected information on the general practices and, in particular, the zip code that enables to geolocate a GP in the spatial unit of interest. Other information on the cost and the health districts of origin are also available, but will not be used in this analysis.

Refer to Table 1 for a summary of the fields used in this study. All the cited datasets are available at https://digital.nhs.uk/.

Medicines taxonomy. To organize drugs in functional groups we refer to the British National Formulary (BNF) [19], a pharmaceutical reference book providing key information on the selection, prescribing, dispensing and administration of medicines. This categorization is used for all medicines, dressings and appliances dispensed in the UK and it is helpful to distinguish distinct features characterizing each pharmaceutical. BNF codes are organized in chapters, sections, paragraphs and sub-paragraphs that are categories of increasing specificity.

In this study, we focus on the following groups:

- *Cough Preparations* (BNF section 3.9);
- *Painkillers*, i.e. non-opioid analgesics (BNF section 4.7.1);
- *Antivirals* for influenza and respiratory syncytial virus (BNF sections 5.3.4 and 5.3.5, respectively);
- *Antibiotics* as cephalosporins and other beta-lactams, macrolides, sulfonamides and trimethoprim, quinolones (BNF sections 5.1.2, 5.1.5, 5.1.8, and 5.1.12, respectively);

Patients provenance. To account for the provenance of the patients, we refer to another publicly available dataset [20] from NHS that lists for each GP the number of registered patients living in a particular Lower Layer Super Output Area (LSOA). LSOAs are geographic areas designed by Census for statistical aims, with a mean population figure of 1,500. A gender split is also available. Provenance data covers the period from 2014 to 2018 and is provided quarterly. We notice an inconsistency in the reference spatial units for the 2014 provenance data that resulted in the exclusion from the analysis of 810 LSOAs with not matching codes (3.6% of the population). Refer to Table 2 for a summary.

Table 2: Description of the data on patients provenance.

Attribute	Description
Practice	6 characters (1 letter + 5 numbers) code identifying a practice to which patients are registered during a specific quarter.
Total	Total number of patients registered to the practice.
Males	Total number of males registered to the practice.
Females	Total number of females registered to the practice.
Date	Year and month of activity, reported as '*yyyymm*'. Patients provenance is reported every quarter, particularly in January, April, July and October. Thus, *mm* can be equal only to 01, 04, 07, 10.

3 METHODS

In this section, we describe the methodology adopted to calculate the monthly drugs rate for both sources of data, i.e. Prescriptions data and Flusurvey data. The resulting time series are then compared in terms of temporal trend and spatial distribution. We report results both nationally and for the 9 regions of England.

3.1 Flusurvey

In this study, we use the Flusurvey data for four influenza seasons, from 2014-2015 to 2017-2018, and we include the data reported in the period from November to May of each season, except for the 2014-2015 season for which the data collection was interrupted earlier, in March. For each influenza season we include in the analysis only those participants whose home postcode resides in England and who have completed at least one background survey since their registration to the platform and at least one symptoms survey, excluding the first symptoms survey submitted at the time of first participation to the data collection. This is to avoid sporadic participation and to guarantee a basic level of confidence with the questions proposed in the questionnaire. Moreover, previous studies showed a potential correlation between symptoms presence and willingness to join the platform [4, 6].

For each influenza season, we retrospectively compute the time series of drugs rate from symptoms surveys in which participants have reported the presence of at least one symptom and the uptake of at least one medicine among the ones proposed in the questionnaire. Specifically, participants can choose from the following list of possible answers and tick all that applies in their case:

- No medication
- Painkillers (e.g. paracetamol, lemsip, ibuprofen, aspirin, calpol)
- Cough medication (e.g. expectorants)
- Antivirals (Tamiflu, Relenza)

- Antibiotics
- Other
- I don't know/can't remember

Thus, here we identify four categories of drugs, namely Antibiotics, Antivirals, Cough medications, and Painkillers. We refer to the symptoms onset date to retrieve when participants took medicines for their symptoms. Only records with symptoms onset dates occurring between November 1st and May 31st are retained. If the symptom onset date is not reported, we use the survey submission date. If the symptoms onset date is marked as incorrect (e.g. future date or date before the start of data collection), the survey is removed from the analysis. Duplicates of symptoms surveys submitted within the same day are removed from the analysis in order to avoid multiple counting of the same episode.

In addition, participants are asked to report whether they have sought health care treatment for their symptoms in the form of a face-to-face visit, including general practitioners, hospitals, emergency rooms, and any other medical services, or in the form of telephone/internet contact, including for example having talked to receptionist, doctor or nurse, or contacted the NHS. Here, we consider only face-to-face visits due to the underlying assumption that drugs prescriptions do not occur via telephone or internet contact.

Finally, the time series of drugs rate are obtained as the ratio between the monthly drugs counts and the number of participants involved in the data collection each season. Hereafter, we will refer to them as Flusurvey time series. Furthermore, we also compute the specific time series of drugs uptake associated to the seeking of medical services.

3.2 Prescriptions

As described in Section 2.2, the BNF codes associated with each pharmaceutical allows to effectively select different functional groups. Similarly to the Flusurvey case, we focus on the *Cough Preparations*, *Painkillers*, *Antibiotics*, and *Antivirals* categories. First, we compute the monthly consumption per category and general practice across the period of interest. Note that in this first step the unit of study is the general practice; however, we are interested in characterizing drugs consumption at a finer spatial granularity. To this extent, we develop a redistribution mechanism that combines the prescriptions and the patients provenance datasets to compute consumption rates at the level of LSOAs.

To this extent, we define $i(g)$ as the number of prescriptions dispensed by the practice g and $p(u,g)$ as the number of patients registered at the practice g *and* living in the spatial unit u. These quantities are directly available from the data. We hypothesize that the amount of prescriptions dispensed by a GP g in a specific geographical area u is proportional to the number of patients registered at g and living in u. Consequently, we estimate the number of items prescribed in each spatial unit u as:

$$i(u) = \sum_{g \in G_u} i(u,g)$$

where G_u is the set of GPs with at least a patient living in u and $i(u,g)$ is the estimated amount of prescription items dispensed in a spatial unit u by the GP g. To evaluate $i(u,g)$ we compute:

$$i(u,g) = i(g)\frac{p(u,g)}{p(g)}$$

where $p(g)$ is the total number of patients enrolled in the GP g, which can be determined as

$$p(g) = \sum_{u \in U_g} p(u,g)$$

with U_g as the collection of the LSOAs where patients of GP g live. Having assumed uniformity in the prescribing distribution, we point out that $i(u,g)$ results to be the proportion of prescription items dispensed by GP g in the LSOA u in accordance with the amount of patients living in that area.

The number of patients living in a certain spatial unit u and registered at any GP can be easily computed as:

$$p(u) = \sum_{g \in G_u} p(u,g).$$

Observe that $p(u)$ could be potentially split by gender that allows the computation of the male $p_{male}(u)$ and female $p_{female}(u)$ rates respectively.

The *prescription rate* in the spatial unit u can be now computed as the ratio between items and patients:

$$r(u) = \frac{i(u)}{p(u)}.$$

We compute $r(u)$ monthly during the period 2014-2018 for each of the drugs classes of interest, ending up with 48 observational points. Amongst those, only the 26 observations that matches with the Flusurvey data are preserved for the analysis. Hereafter, we will refer to them as Prescriptions time series.

4 RESULTS

In Table 3, we report for each season the number of participants, the total number of symptoms surveys submitted to the system as well as the average number of monthly surveys considered in our analysis (i.e. symptoms surveys reporting both symptoms and medicines). A total of 8,825 unique individuals participated to the Flusurvey data collection during one or more of the four influenza seasons under study. The majority of participants live in the region of Greater London (i.e. UKI) with an average of about 20% of participants and in the neighboring region of South East (i.e. UKJ) with an average of about 21% of participants.

The number of symptoms surveys in which participants reported having symptoms and taking medicines ranges approximately from 14% to 17% per season compared to the total number of symptoms surveys submitted each season. The total drugs count per season refers mainly to painkillers with an average of about 76%, followed by cough medications with 18%, antibiotics with 6% and antivirals with 0.3%. On average, the proportion of participants who reported having visited or contacted a medical service for their symptoms is approximately 90% for antibiotics, 78% for antivirals, 18% for cough medications and 14% for painkillers.

In Figure 1, we report the distributions of the monthly drugs rate of the Prescriptions data and the Flusurvey data for the four influenza seasons under study. In particular, Prescriptions data show a decreasing trend over time of the consumption of antibiotics,

Table 3: Participation to Flusurvey during the four influenza seasons under study.

season	no. participants	total no. surveys	average no. surveys per month
2014-2015	4,279	44,779	1,502
2015-2016	4,216	57,313	1,305
2016-2017	3,323	46,595	1,000
2017-2018	5,219	68,362	1,382

Table 4: Spearman's correlation and Hit Rate for antibiotics and cough medications for the four influenza seasons under study.

season	Antibiotics		Cough medications	
	Spearman's correlation	Hit Rate (%)	Spearman's correlation	Hit Rate (%)
2014-2015	0.60	75.0	0.70	75.0
2015-2016	0.82*	66.7	0.96*	100.0
2016-2017	0.89*	66.7	0.96*	83.3
2017-2018	0.86*	83.3	0.96*	83.3

*p-value<0.05

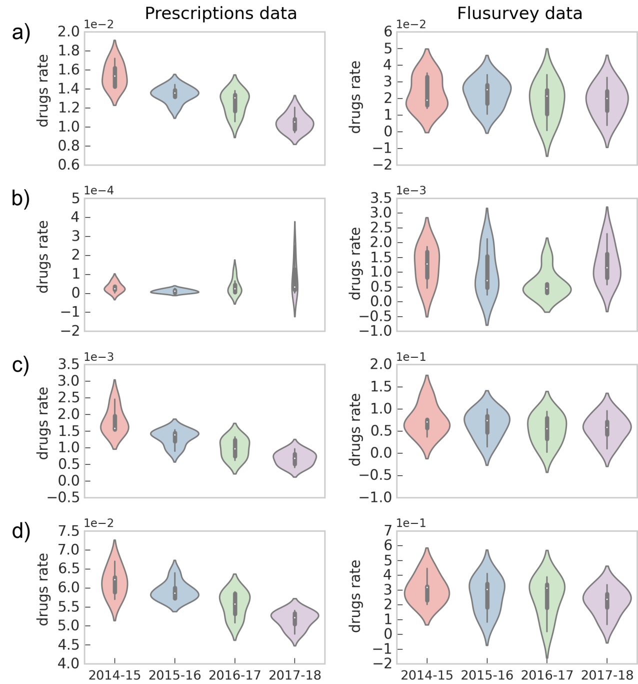

Figure 1: Distributions of the monthly drugs rate of the Prescriptions data (left) and Flusurvey data (right). Rows refer to: a) Antibiotics; b) Antivirals; c) Cough medications; d) Painkillers.

cough medications and painkillers, whereas the temporal pattern of the monthly consumption rate as detected by Flusurvey is quite similar across the various influenza seasons. Antivirals have the lowest rates of consumption in both datasets. It is worth noting that drugs rate cannot be compared in amplitude as the two data sets refer to different measures.

In Figure 2, we show the resulting Flusurvey time series of drugs uptake as compared to the Prescriptions time series. For the sake of visual comparison the time series are rescaled according to the highest peak of each season. The time series of antivirals lack some data points due to the low rate of prescription and consumption. The Spearman's correlations between the time series of prescription data and Flusurvey data correspond to 0.54 for antibiotics (p-value<0.01), 0.47 for antivirals (p-value<0.05), 0.58 (p-value<0.01) for cough medications and 0.38 (p-value=0.06) for painkillers. The values of the Hit Rate instead correspond to 68.0 for antibiotics, antivirals and painkillers, while 84.0 for cough medications. In Table 4, we further report the values of the Spearman's correlations and Hit Rate per season for antibiotics and cough medications.

To assess the spatial distribution between the Flusurvey and Prescriptions time series at the level of the 9 regions of England, we performed a Kendall's tau correlation. Specifically, we found weak positive correlations for antibiotics (τ=0.19, p-value<0.01), cough medications (τ=0.30, p-value<0.01), and painkillers (τ=0.09, p-value<0.05), while no statistically significant correlation for antivirals (τ=0.06, p-value=0.53). Looking at the spatial relationship between the cumulative drugs rates, instead, we found no statistically significant correlations, ranging from -0.03 for antibiotics to 0.24 for cough medications (all p-values>0.05), as also shown in the maps in Figure 3. Moreover, in Figure 4 we show the correlations between the Flusurvey and Prescriptions time series at the level of the 9 regions of England, sorted by the number of participants involved in Flusurvey.

5 DISCUSSION AND CONCLUSIONS

In this study, we investigated the performance of a digital participatory surveillance system called Flusurvey in capturing the spatio-temporal patterns of flu-related drugs consumption in the general population in England. Specifically, we analyzed the monthly consumption of four categories of flu-related drugs, namely antibiotics, antivirals, cough medications and painkillers. Self-reported drugs uptake collected by the Flusurvey platform is compared against the prescription data released by public health officials in England, here considered our ground truth.

As already shown in previous studies, in participatory surveillance systems only a small fraction of participants refers to seek health care treatment for their symptoms [25, 31]. Consequently, self-reported drugs consumption in the Flusurvey cohort refers mainly to those drugs that do not require a prescription from a public health official, mainly painkillers which represent about 76% of the total drugs count, on average. In particular, among the participants who reported having sought medical treatment, the majority also reported having taken antibiotics (about 90%) and antivirals (about 78%), while only a smaller portion of participants reported having sought medical treatment in conjunction with the consumption of cough medications (about 18%) and painkillers (about 14%). Interestingly, our results show significant temporal

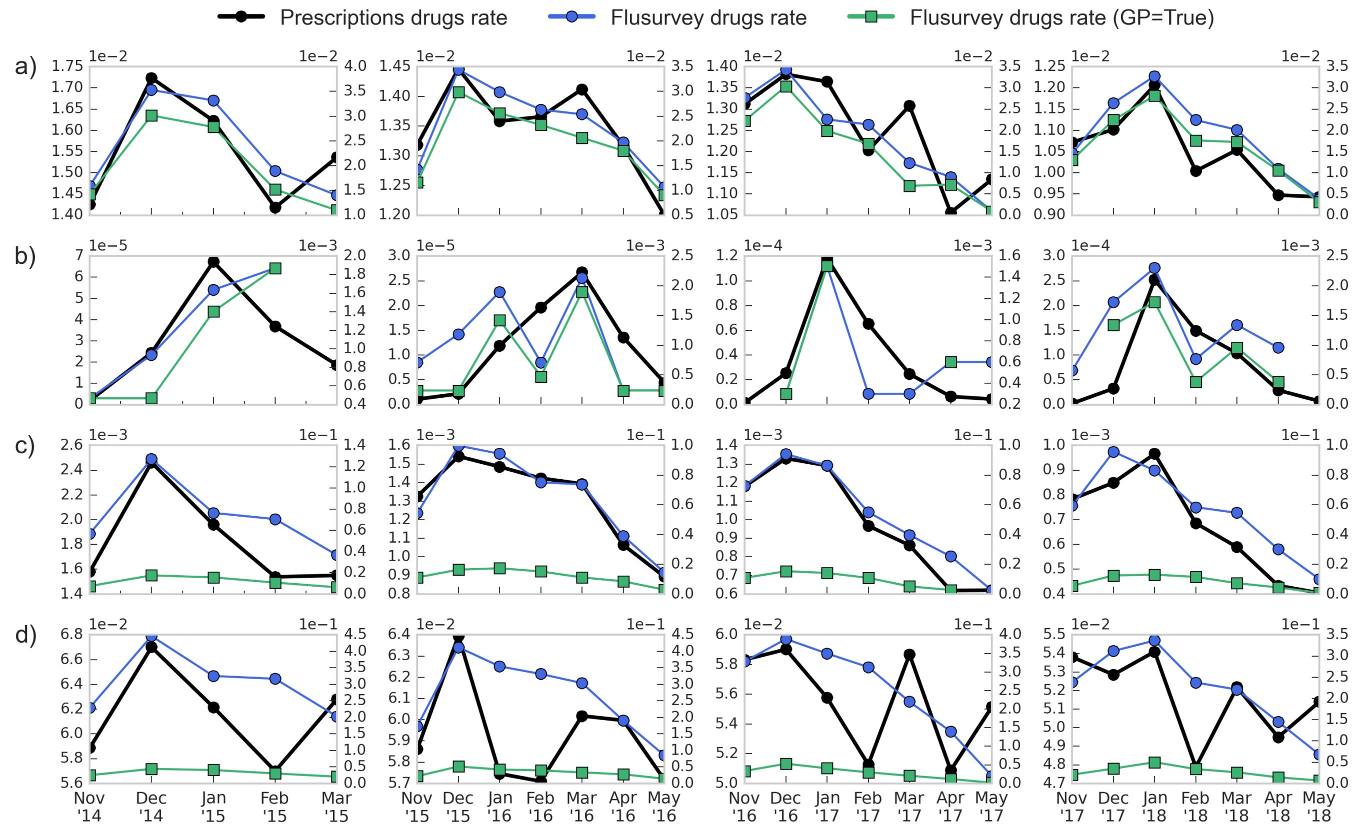

Figure 2: Time series of drugs rate as obtained by Prescriptions data (left y-axis) and Flusurvey data (right y-axis). Rows refer to: a) Antibiotics; b) Antivirals; c) Cough medications; d) Painkillers. The time series for the 2014-2015 flu season have data points only until March because the Flusurvey data collection ceased earlier that year (see Section 3.1 for more details).

correlations between the self-reported data collected by Flusurvey and the ground truth at the national level for antibiotics, antivirals and cough medications, while no significant temporal correlation instead for painkillers. It is also worth noting that in the prescriptions dataset, painkillers show no seasonality, whereas the consumption of antibiotics, antivirals and cough medications shows well-defined peaks during the winter months, thus representing the closest pattern to the purpose of participatory surveillance system like Flusurvey for capturing the influenza-like illness activity during the winter season.

On the other hand, the spatial distribution of the Flusurvey self-reported drugs consumption is not statically correlated with the spatial distribution of prescriptions drugs consumption. This might be mainly due to the spatial distribution of the Flusurvey participants, which is not representative of the spatial distribution of the general population in the United Kingdom and is strongly biased towards urban areas [8].

Prescriptions data can instead praise a higher spatial resolution up to the level of LSOAs. However, prescriptions data are released typically with a lag of 2 months due to the time to collect and aggregate data reported by the general practitioners. Moreover, prescriptions data are aggregated monthly, while Flusurvey can praise a higher temporal resolution since the weekly reports are generally submitted on a weekly basis.

The main assumption in adopting prescriptions to characterize drugs consumption in certain areal units is the homogeneity in their spatial distribution. For example, we assume that if a general practice prescribes *n* antibiotics items in a month they are uniformly spatially distributed amongst registered patients without taking into account demographic or socio-economic determinants of their provenance areas. Even if it looks a reasonable assumption for drugs related to widespread medical conditions like influenza, a quantitative validation is missing and planned for future work. A further issue that affects the redistribution methodology is the inability to derive accurate estimations in the case of drugs with low consumption rates. In fact, for certain categories, the prescriptions volume is comparable in magnitude to the number of spatial units, i.e. about 33,000 in the case of LSOAs, that makes the allocation procedure not accurate for high-resolution spatial granularities. In these cases, a solution would be to aggregate the prescriptions data to coarser geographical regions, as described in several previous works.

In conclusion, in this pilot study we have shown the feasibility of using self-reported information collected by Flusurvey to measure

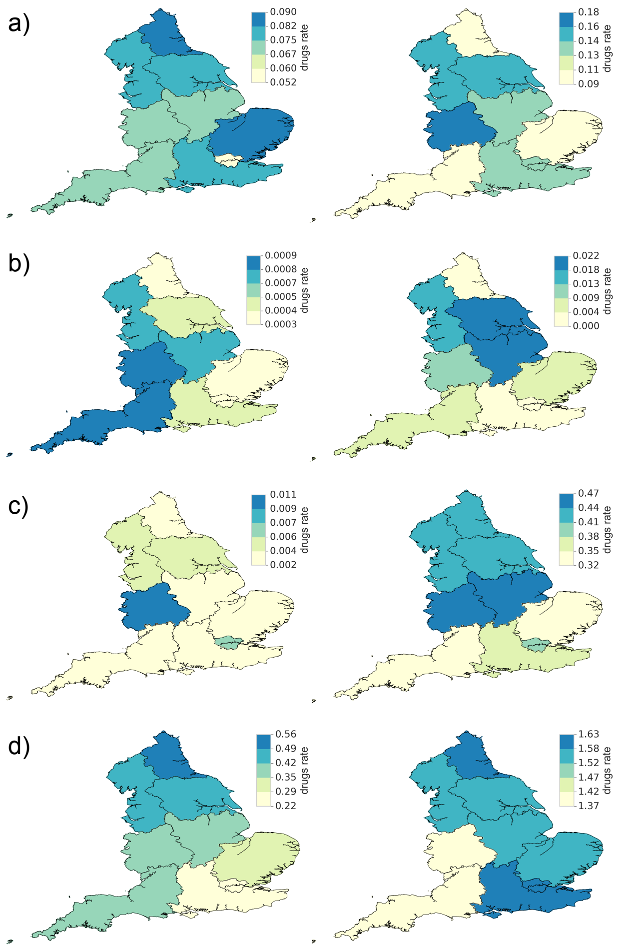

Figure 3: Cumulative drugs rate of Prescriptions data (left) and Flusurvey data (right) at the resolution of the 9 regions of England for the 2017-2018 influenza season. Rows refer to: a) Antibiotics; b) Antivirals; c) Cough medications; d) Painkillers.

the flu-related drugs consumption in the general population. This represents a further step towards the integration of novel digital approaches into existing traditional practices in public health and will allow the further development of a public health tool capable of describing the spatio-temporal patterns of disease activity and drugs consumption in the general population. Since both data sources we have explored in this work are affected by biases and limitations, one of the immediate applications of our study could be to combine and integrate the two datasets in order to overcome the lack of spatial and temporal granularity in the Flurvey platform and in the Prescriptions data, respectively. In particular, the timeliness of the information collected in near real-time by the Flusurvey platform could well complement the prescription data collected and released monthly by the NHS system. This highlights the added value of participatory surveillance data in the application of public health policies, such as the preparedness for particularly severe influenza seasons for which a higher drugs consumption is expected in the general population. On the other hand, the Flusurvey platform would particularly benefit from specific enrollment campaigns in order to reach out to additional participants and ensure more accurate geographical coverage.

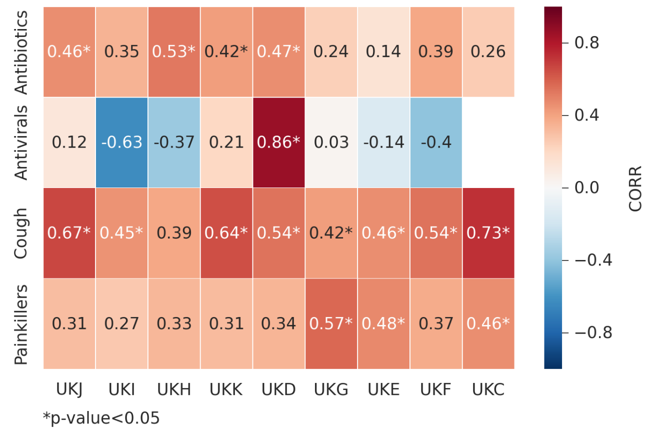

Figure 4: Correlations between the Prescriptions and the Flusurvey time series at the resolution of the 9 regions of England.

Further future applications of this study include the use of predictive models to estimate the expected pattern of drugs consumption in the general population and help inform and guide the public health decision making. Moreover, drugs consumption rates can also be integrated into forecasting models for influenza-like illnesses as an additional layer of information to better capture the unfolding of seasonal influenza epidemics in a timely manner.

Finally, this work serves as the first validation of the reliability of the information on drugs consumption as detected by a participatory surveillance platform. In countries where official information on drugs prescriptions is not timely collected or publicly released, drug consumption data provided by participatory surveillance platforms could provide relevant public health insight that would not otherwise be available.

ACKNOWLEDGMENTS

We thank all the Flusurvey users who contributed to this study. Daniela Paolotti and Daniela Perrotta acknowledge support from the Lagrange Project of the Institute for Scientific Interchange Foundation (ISI Foundation) funded by Fondazione Cassa di Risparmio di Torino (Fondazione CRT). Dario Delle Vedove has been partially supported by the project "Countering Online hate speech through Effective on-line Monitoring" funded by Compagnia di San Paolo.

REFERENCES

[1] Flusurvey (United Kingdom). Available at: https://flusurvey.org.uk/ (accessed on May, 2019).
[2] Influenza (Seasonal), Fact Sheet Number 211. Available at: http://www.who.int/mediacentre/factsheets/fs211/en/index.html.
[3] Influenzanet website. Available at: http://influenzanet.info.
[4] A. J. Adler, K. T. Eames, S. Funk, and W. J. Edmunds. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey. BMC

infectious diseases, 14(1):232, 2014.
[5] A. Brimicombe and P. Mungroo. Geographical variation in GP drug prescribing for schizophrenia and similar psychosis in England. Journal of Public Mental Health, 17(4):218–223, 2018.
[6] E. Brooks-Pollock, N. Tilston, W. J. Edmunds, and K. T. Eames. Using an online survey of healthcare-seeking behaviour to estimate the magnitude and severity of the 2009 H1N1v influenza epidemic in England. BMC infectious diseases, 11(1):1, 2011.
[7] J. S. Brownstein, S. Chu, A. Marathe, M. V. Marathe, A. T. Nguyen, D. Paolotti, N. Perra, D. Perrotta, M. Santillana, S. Swarup, et al. Combining participatory influenza surveillance with modeling and forecasting: Three alternative approaches. JMIR Public Health and Surveillance, 3(4):e83, 2017.
[8] P. Cantarelli, M. Debin, C. Turbelin, C. Poletto, T. Blanchon, A. Falchi, T. Hanslik, I. Bonmarin, D. Levy-Bruhl, A. Micheletti, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. BMC public health, 14(1):1, 2014.
[9] H. J. Curtis, R. Croker, A. J. Walker, G. C. Richards, J. Quinlan, and B. Goldacre. Opioid prescribing trends and geographical variation in England, 1998-2018: a retrospective database study. The Lancet Psychiatry, 6(2):140 – 150, 2019.
[10] H. J. Curtis and B. Goldacre. OpenPrescribing: normalised data and software tool to research trends in English NHS primary care prescribing 1998-2016. BMJ Open, 8(2):e019921, 2018.
[11] H. J. Curtis, A. J. Walker, K. R. Mahtani, and B. Goldacre. Time trends and geographical variation in prescribing of antibiotics in England 1998âĂŞ2017. Journal of Antimicrobial Chemotherapy, 74(1):242–250, 09 2018.
[12] C. Dalton, D. Durrheim, J. Fejsa, L. Francis, S. Carlson, E. T. d'Espaignet, F. Tuyl, et al. FluTracking: a weekly Australian community online survey of influenza-like illness in 2006, 2007 and 2008. Communicable diseases intelligence quarterly report, 33(3):316, 2009.
[13] W. J. Edmunds and S. Funk. Using the Internet to estimate influenza vaccine effectiveness. Expert review of vaccines, 11(9):1027–1029, 2012.
[14] C. Guerrisi, C. Turbelin, T. Blanchon, T. Hanslik, I. Bonmarin, D. Levy-Bruhl, D. Perrotta, D. Paolotti, R. Smallenburg, C. Koppeschaar, et al. Participatory syndromic surveillance of influenza in Europe. Journal of Infectious Diseases, 214(suppl 4):S386–S392, 2016.
[15] K. Kalimeri, M. Delfino, C. Cattuto, D. Perrotta, V. Colizza, C. Guerrisi, C. Turbelin, J. Duggan, J. Edmunds, C. Obi, et al. Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms. PLoS computational biology, 15(4):e1006173, 2019.
[16] C. E. Koppeschaar, V. Colizza, C. Guerrisi, C. Turbelin, J. Duggan, W. J. Edmunds, C. Kjelsø, R. Mexia, Y. Moreno, S. Meloni, et al. Influenzanet: citizens among 10 countries collaborating to monitor influenza in Europe. JMIR public health and surveillance, 3(3), 2017.
[17] E. M, A. A, A.-O. D, and H. S. Trends and patterns in antibiotic prescribing among out-of-hours primary care providers in England, 2010-14. J Antimicrob Chemother., 72(12):3490–3495, 2017.
[18] L. Mordecai, C. Reynolds, L. J. Donaldson, and A. C. de C Williams. Patterns of regional variation of opioid prescribing in primary care in England: a retrospective observational study. British Journal of General Practice, 68(668):e225–e233, 2018.
[19] NHS. Bnf booklet 76. march-september 2018, 2018.
[20] NHS. Patients registered at a GP practice, 2018.
[21] NHS. Practice level prescribing data, 2018.
[22] D. Paolotti, A. Carnahan, V. Colizza, K. Eames, J. Edmunds, G. Gomes, C. Koppeschaar, M. Rehn, R. Smallenburg, C. Turbelin, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clinical Microbiology and Infection, 20(1):17–21, 2014.
[23] O. Patterson-Lomba, S. Van Noort, B. J. Cowling, J. Wallinga, M. G. M. Gomes, M. Lipsitch, and E. Goldstein. Utilizing syndromic surveillance data for estimating levels of influenza circulation. American journal of epidemiology, 179(11):1394–1401, 2014.
[24] M. Peppa, W. J. Edmunds, and S. Funk. Disease severity determines health-seeking behaviour amongst individuals with influenza-like illness in an internet-based cohort. BMC infectious diseases, 17(1):238, 2017.
[25] D. Perrotta, A. Bella, C. Rizzo, and D. Paolotti. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in Italy. PloS one, 12(1):e0169801, 2017.
[26] D. Perrotta, M. Tizzoni, and D. Paolotti. Using participatory Web-based surveillance data to improve seasonal influenza forecasting in Italy. In Proceedings of the 26th International Conference on World Wide Web, pages 303–310. International World Wide Web Conferences Steering Committee, 2017.
[27] M. S. Smolinski, A. W. Crawley, K. Baltrusaitis, R. Chunara, J. M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, and J. S. Brownstein. Flu Near You: crowdsourced symptom reporting spanning two influenza seasons. American journal of public health, 105(10):2124–2130, 2015.
[28] M. S. Smolinski, A. W. Crawley, J. M. Olsen, T. Jayaraman, and M. Libel. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR public health and surveillance, 3(4):e62, 2017.
[29] R. Spence, A. Roberts, C. Ariti, and B. M. Focus on: Antidepressant prescribing. 2014.
[30] N. L. Tilston, K. T. Eames, D. Paolotti, T. Ealden, and W. J. Edmunds. Internet-based surveillance of influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic. BMC Public Health, 10(1):650, 2010.
[31] S. P. van Noort, C. T. Codeço, C. E. Koppeschaar, M. Van Ranst, D. Paolotti, and M. G. M. Gomes. Ten-year performance of influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. Epidemics, 13:28–36, 2015.
[32] Q. Zhang, C. Gioannini, D. Paolotti, N. Perra, D. Perrotta, M. Quaggiotto, M. Tizzoni, and A. Vespignani. Social data mining and seasonal influenza forecasts: The FluOutlook platform. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 237–240. Springer, 2015.