Epidemiol. Infect. (2014), **142**, 2433–2446. © Cambridge University Press 2014
doi:10.1017/S0950268813003452

A concept for routine emergency-care data-based syndromic surveillance in Europe

A. ZIEMANN[1]*, N. ROSENKÖTTER[1], L. GARCIA-CASTRILLO RIESGO[2], S. SCHRELL[1], B. KAUHL[1], G. VERGEINER[3], M. FISCHER[4], F. K. LIPPERT[5], A. KRÄMER[6], H. BRAND[1] AND T. KRAFFT[1]

[1] *Department of International Health, CAPHRI School of Public Health and Primary Care, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands*
[2] *Department of Medical Sciences and Surgery, Faculty of Medicine, University of Cantabria, Santander, Spain*
[3] *Dispatch Centre Tyrol (Leitstelle Tirol Gesellschaft mbH), Innsbruck, Austria*
[4] *Department of Anesthesia and Intensive Care, Klinik am Eichert, Goeppingen, Germany*
[5] *Emergency Medical Services, Capital Region of Denmark, Copenhagen, Denmark*
[6] *Department of Public Health Medicine, School of Public Health, University of Bielefeld, Bielefeld, Germany*

Received 22 March 2013; Final revision 21 November 2013; Accepted 13 December 2013; first published online 24 January 2014

SUMMARY

We developed a syndromic surveillance (SyS) concept using emergency dispatch, ambulance and emergency-department data from different European countries. Based on an inventory of sub-national emergency data availability in 12 countries, we propose framework definitions for specific syndromes and a SyS system design. We tested the concept by retrospectively applying cumulative sum and spatio-temporal cluster analyses for the detection of local gastrointestinal outbreaks in four countries and comparing the results with notifiable disease reporting. Routine emergency data was available daily and electronically in 11 regions, following a common structure. We identified two gastrointestinal outbreaks in two countries; one was confirmed as a norovirus outbreak. We detected 1/147 notified outbreaks. Emergency-care data-based SyS can supplement local surveillance with near real-time information on gastrointestinal patients, especially in special circumstances, e.g. foreign tourists. It most likely cannot detect the majority of local gastrointestinal outbreaks with few, mild or dispersed cases.

Key words: Gastrointestinal infections, public health, surveillance system.

INTRODUCTION

Syndromic surveillance (SyS) of pre-diagnostic cases based on signs and symptoms or health-related behaviour is a supplementary approach for timely detection of public health threats and for monitoring events with potential public health impact if information from other surveillance systems are not yet or not at all available [1]. SyS can provide a flexible and cost-effective way to gain timely information about the health impact of known and unknown, communicable and non-communicable, natural and man-made health threats [2, 3].

The European landscape of public health surveillance mainly consists of three parallel schemes. The first scheme comprises the specific communicable

* Author for correspondence: Ms. A. Ziemann, Department of International Health, Faculty of Health, Medicine and Life Sciences, School of Public Health and Primary Care (CAPHRI), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands.
(Email: alexandra.ziemann@maastrichtuniversity.nl)

disease surveillance systems of European Union member states (MS) that provide information on confirmed cases following a common case definition to the European Surveillance System (TESSy) [4]. The second scheme is different reporting systems through which MS give account of communicable or non-communicable events to inform other MS and European institutions, e.g. the Early Warning and Response System (EWRS) [5]. The third scheme comprises unspecific information collated by European networks of different countries, e.g. Influenzanet for self-reported influenza symptoms [6], or EuroMOMO for mortality monitoring [7], and by the Medical Information System (MedISys) that automatically screens online news wires concerning health events [8]. SyS is accomplished in MS at the local, regional, and national levels [9]. A systematic approach towards European SyS could support timely, comparable, cross-border surveillance.

Routinely collected emergency-care data from (i) emergency medical dispatch (EMD) centres, (ii) ambulance or emergency medical services (EMS), and (iii) emergency departments (ED) can be a valuable source for SyS. Across Europe, emergency-care data is available following a common structure [10]. The biggest advantage is the opportunity of real-time reporting of electronic emergency data that can offer timelier and more frequent information compared to established traditional surveillance systems, e.g. based on sentinel doctors [2]. It provides data based on a form of clinical assessment, e.g. working diagnoses from emergency physicians (EP), which have a higher specificity for SyS compared to non-clinical data from, e.g. over-the-counter drug sales [11].

We aimed at developing the first concept for SyS based on three routine emergency data sources that are applicable across Europe. We describe the development of the SyS concept and present results of a case study testing the SyS concept using the example of local gastrointestinal outbreak detection.

METHODS

SyS system concept

Inventory of emergency data availability in Europe

We asked regional (sub-national) emergency service representatives in 12 countries (Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Norway, Spain, Turkey) to assess availability and content of routine datasets collected in EMD, EMS, and ED. Using a semi-standardized survey we asked for the method of data collection, i.e. manual or electronic, the frequency of data availability, e.g. daily, and the available data fields in the routine datasets.

Syndrome definition

Based on the inventory, we defined syndromes of potential public health relevance that could be generated using routine emergency data. Based on a focus group discussion with emergency-care and public-health experts from across Europe and examples from the literature, we developed recommendations for generating syndromes based on the most common diagnostic coding systems used in EMD, EMS, and ED, i.e. Advanced Medical Priority Dispatch System (AMPDS), versions 11.3 and 12.0 (Priority Dispatch Inc., USA), International Classification of Disease (ICD) 9th and 10th revisions, chief complaints based on Canadian Triage and Acuity Scale (CTAS), and the Minimum Dataset for Emergency Physicians (MIND).

SyS system design

Based on a review of the literature and material published on existing SyS systems and a consultation with European emergency-care, public-health and information technology experts, we developed a design concept for an emergency data-based SyS system. We defined a minimum standard dataset as input for the SyS system that is applicable for EMD, EMS and ED, defined the data flow, selected statistical analytical methods for detecting unusual aberrations, and described ways of reporting the output.

Case study on local gastrointestinal outbreak detection

We tested our SyS concept for EMD, EMS and ED, and for different syndromes and purposes, based on retrospective analyses of historical data from regional emergency systems in four countries [12]. In this paper we present the results of a case study on local gastrointestinal outbreak detection.

Datasets

We analysed data from the EMD centre in the state of Tyrol, Austria (EMD-AT dataset), data from EMS staffed by EP in the state of Tyrol, Austria

(EP-AT dataset), the county of Goeppingen, Germany (EP-DE dataset) and the country of Belgium (EP-BE dataset), and data from an ED in a university hospital in the city of Santander, Spain (ED-ES dataset). Table 1 describes the main characteristics of the datasets.

Gastrointestinal syndrome case definition

Table 2 details the definition of gastrointestinal syndrome cases for five common emergency-care coding systems as an example for a syndrome that can be generated based on routine emergency-care data. An emergency case which received any code included in Table 2 was included in the case study.

Temporal aberration detection algorithms

As a first step, three detection algorithms based on cumulative sums were applied for the analysis of aberrations in the time series of gastrointestinal syndrome cases: C1, C2, and C3 based on short-term baselines [13], and two cumulative sum algorithms based on longer baselines, one for normal (CUSUM-N) and one for Poisson-distributed data (CUSUM-P) [14]. If the distribution of the datasets for a specific syndrome was neither normal nor Poisson distributed, as was the case for gastrointestinal syndrome cases, we applied all algorithms in parallel. The CUSUM algorithms were enhanced with the fast initial response (FIR) technique which ensures that large chart values do not inflate following values preventing the production of excessive signals [15]. In the case study the algorithms were applied retrospectively. We analysed periods of six (EP-BE dataset) or 12 (EMD-AT, ED-ES datasets) months and produced a daily CUSUM value. For each analysis period, we calculated baseline means to which the actual values were compared, based on the 6 or 12 months preceding the analysis period (Table 1). For the CUSUM-P analysis, the accepted mean was defined close to the actual mean and the threshold value h was defined by look-up procedure in the table of Lucas [16]. The temporal aberration detection algorithms have been applied using Microsoft Excel 2003 (Microsoft Corp., USA).

Spatio-temporal cluster detection algorithm

In a second step, outbreak periods that were identified based on temporal aberration detection analysis (see definition of outbreaks in the next section) were analysed by a prospective spatio-temporal scan statistic [17]. The scan statistic process can be explained as a cylindrical scanning window that moves flexibly over the study area. The width of the cylinder base represents the geographical area and the height represents the time period which is scanned. The scan statistic evaluates for all possible cylinder locations and sizes if an observed cluster of cases is caused by chance. The scan statistic can be applied to different levels of spatial aggregation of cases. In the case of spatially aggregated datasets, the cases are concentrated on the centroids of an area. In our case study, a prospective spatio-temporal Bernoulli model-based scan statistic was applied to the exact addresses of the emergency sites in the EMD-AT dataset. A prospective spatio-temporal Poisson model was applied to the EP-AT, EP-BE, EP-DE, and ED-ES datasets, based on the centroids of each administrative area (Table 1) [17]. During the scanning process the rates of gastrointestinal cases divided by the total number of emergency cases within the scanning window were compared to the rates outside of the window. The baseline populations were generated using the total number of emergencies in the previous 12 months (EP-AT, EP-DE) and the previous 6 months (EP-BE, ED-ES). The likelihood that a cluster exists by chance was characterized by a P value based on 999 Monte-Carlo simulations [17].

For each syndrome, different parameters have to be defined for detecting relevant clusters. For local gastrointestinal outbreak detection, only clusters with the parameters of 1 day temporal length, enclosing a circular area of up to 1 km radius, and with a significance level of $P<0{\cdot}001$ that the cluster exists by chance were defined as relevant. Pre-tests with different parameters showed that for longer and larger cluster sizes the number of cases that formed a cluster was too low and/or the cases were scattered over too large an area to reflect a true positive outbreak. The analyses were performed using SaTScan™ (v. 9.1.1., M. Kulldorff and Information Management Services Inc., USA). The identified spatio-temporal clusters were visualized using ESRI ArcGIS® v. 10.1. (Environmental Systems Research Institute Inc., USA).

Definition of an outbreak

We followed a decision tree as suggested by Meyer *et al.* [18] and Ansaldi *et al.* [19] to define inclusion criteria for outbreaks based on the signals given by,

Table 1. *Characteristics of emergency datasets and reference data focusing on the example of local gastrointestinal outbreak detection*

Country	Austria	Austria	Belgium	Germany	Spain
Catchment area	State of Tyrol: Innsbruck city, Innsbruck district, Kufstein district	State of Tyrol: Kufstein district	Country of Belgium (national coverage)	County of Goeppingen	City of Santander (reference hospital for the Autonomous Region of Cantabria)
Data source (abbreviation)	EMD (EMD-AT)	EMS staffed with EP (EP-AT)	EMS staffed with EP (EP-BE)	EMS staffed with EP (EP-DE)	ED (ED-ES)
Population served (approximate)	380 000	99 000	10 500 000	256 000	300 000(Santander) 580 000 (Cantabria)
Data provider	Dispatch Centre Tyrol, Innsbruck, Austria	Dispatch Centre Tyrol, Innsbruck, Austria	Ministry of Health, Brussels, Belgium	Emergency physician ambulance service, County of Goeppingen, Germany	University Hospital Marqués de Valdecilla, Santander, Spain
Data analysis period (gastrointestinal syndrome case study	Jan. 2007–Dec. 2009	Jan. 2007–Dec. 2009	Jan. 2009–Dec. 2009	Jan. 2007–Dec. 2008	May 2010–Apr. 2012
Diagnostic coding system	AMPDS v. 11.3	ICD-10	ICD-9	MIND 2, ICD-10	CTAS
Geographical information	Address emergency site	Postal code emergency site	Postal code emergency site	Community code emergency site	Postal code patient residence
Baseline periods CUSUM (gastrointestinal syndrome case study)	Jan. 2006–Dec. 2008 (divided into 12-month periods)	Jan. 2006–Dec. 2008 (divided into 12-month periods)	Jan. 2009–June 2009	Jan. 2006–Dec. 2008 (divided into 12-month periods)	July 2010–Dec. 2011 (divided into 12-month periods; for 2010: 6-month period)
Mean (gastrointestinal syndrome cases in data analysis period)	3·55	0·14	1·73	0·31	5
Standard deviation (gastrointestinal syndrome cases in data analysis period)	1·94	0·54	1·61	0·56	2·69
Reference data (gastrointestinal syndrome case study)	Foodborne diseases (notifiable disease reporting); source: Federal ministry of health and Tyrolean government; reporting period: 2007–2009; content: outbreaks including dates, number of cases, agent, community and district for state of Tyrol		Foodborne diseases (notifiable disease reporting); source: national public health institute; period: 2009; content: number of foodborne outbreaks per month for country of Belgium	Foodborne diseases (notifiable disease reporting); source: national public health institute; period: 2007–2009; content: number of foodborne cases per week for county of Goeppingen	No reference data

EMD, Emergency medical dispatch, EMS, emergency medical services, ED, emergency department; EP, emergency physician; AMPDS, Advanced Medical Priority Dispatch System; ICD, International Classification of Diseases; MIND, Minimum Dataset for Emergency Physicians; CTAS, Canadian Triage and Acuity Scale.

Table 2. *Gastrointestinal syndrome definition for five common emergency care coding systems [21]*

Coding system	Codes included for gastrointestinal syndrome (Boolean operator: OR)
AMPDS v. 11.3, 12.0	A1 Abdominal Pain
ICD-9	001 Cholera
	002 Typhoid and paratyphoid fevers
	003 Other salmonella infections
	004 Shigellosis
	008·5 Bacterial enteritis unspecified
	005 Other food poisoning (bacterial)
	005·9 Food poisoning unspecified
	008·6 Enteritis due to specified virus
	008·69 Enteritis due to other viral enteritis
	009·2 Infectious diarrhoea
	009·3 Diarrhoea of presumed infectious origin
	787·0 Nausea and vomiting
	787·01 Nausea with vomiting
	787·02 Nausea alone
	787·03 Vomiting alone
	787·81 Diarrhoea
	558·9 Other and unspecified non-infectious gastroenteritis and colitis
	535·5 Unspecified gastritis and gastroduodenitis
ICD-10	A00 Cholera
	A01 Typhoid and paratyphoid fevers
	A02 Other salmonella infections
	A03 Shigellosis
	A04 Other bacterial intestinal infections
	A05 Other bacterial foodborne intoxications, not elsewhere classified
	A08 Viral and other specified intestinal infections
	A09 Diarrhoea and gastroenteritis of presumed infectious origin
	R11 Nausea and vomiting
	K52 Other non-infective gastroenteritis and colitis
	K52·9 Non-infective gastroenteritis and colitis, unspecified
	T62·9 Noxious substance eaten as food, unspecified
MIND II	4·1 Abdominal disorders, acute abdomen
	4·2 Abdominal disorders, gastrointestinal bleeding
	4·3 Abdominal disorders, colic
	4·4 Abdominal disorders, other disease abdomen
	6·2 Metabolic disease, dehydrated
CTAS	Abdominal pain adults
	Abdominal pain children
	Diarrhoea
	Vomiting

AMPDS, Advanced Medical Priority Dispatch System; ICD, International Classification of Diseases; MIND, Minimum Dataset for Emergency Physicians; CTAS, Canadian Triage and Acuity Scale.

first, the temporal and, second, the spatio-temporal detection algorithm. For the case of local gastrointestinal outbreak detection, these were (i) at least 2 days of consecutive temporal aberration detection signals, or (ii) days with an exceptionally high aberration in case numbers from the mean [>3 standard deviations (S.D.) from the baseline mean of the previous 6 or 12 months], and (iii) outbreaks identified by the temporal aberration analyses with corresponding spatio-temporal clusters.

Validation of outbreaks

The comparison with reference data from other (traditional) surveillance systems can give additional assurance that a signal could represent a real event. For the case study on local gastrointestinal outbreaks, we compared the detected outbreaks with notifiable surveillance reports of foodborne diseases. This reference data was available for Tyrol (Austria), Belgium, and Goeppingen (Germany) [20] (Table 1).

Table 3. *Availability of selected electronic emergency care information from three sources (EMD, EMS, ED) in regional emergency institutions in 12 countries (status: June 2009)*

Country	Region	Data source	Data availability: date	Data availability: chief complaint/ working diagnosis	Data availability: age	Data availability: sex
Austria	State of Tyrol (City of Innsbruck, District of Innsbruck, District of Kufstein)	EMD	Daily	Daily	Daily	Daily
	District of Kufstein	EMS (EP)	Daily	Daily	Daily	Daily
	District of Kufstein	ED	Daily	n.a.	Daily	Daily
Belgium	Province of Flemish-Brabant	EMD	n.a.	n.a.	n.a.	n.a.
	City of Leuven	EMS (EP)	Daily	Weekly	Daily	Daily
	City of Leuven	ED	Daily	Daily	Daily	Daily
Czech Republic	City of Prague	EMD	Daily	Daily	Weekly	Weekly
	City of Prague	EMS	Daily	Weekly	–	–
	City of Prague	ED	n.a.	n.a.	n.a.	n.a.
Denmark	Capital Region of Denmark	EMD	Daily	Daily	–	–
	Capital Region of Denmark	EMS (EP)	Daily	Daily	Daily	Daily
	Capital Region of Denmark	ED	Daily	–	Daily	Daily
Finland	City of Kuopio	EMD	Daily	Daily	–	Daily
	City of Kuopio	EMS (EP)	Daily	–	Weekly	Weekly
	City of Kuopio	ED	Daily	Daily	Daily	Daily
France	District of Hauts-de-Seine	EMD	n.a.	n.a.	n.a.	n.a.
	District of Hauts-de-Seine	EMS (EP)	Weekly	–	Weekly	Weekly
	n.a.	ED	n.a.	n.a.	n.a.	n.a.
Germany	County of Goeppingen	EMD	Daily	–	–	–
	County of Goeppingen	EMS (EP)	Daily	Daily	Daily	Daily
	County of Goeppingen	ED	Daily	–	Daily	Daily
Hungary	National level	EMD	n.a.	n.a.	n.a.	n.a.
	National level	EMS (EP)	Monthly	–	Monthly	Monthly
	City of Budapest	ED	Daily	–	Daily	Daily
Italy	Province of Genoa	EMD	Daily	Daily	–	–
	Province of Genoa	EMS (EP)	Monthly	–	Monthly	Monthly
	Province of Genoa	ED	n.a.	n.a.	n.a.	n.a.
Norway	City of Bergen	EMD	Daily	Daily	Daily	Daily
	City of Bergen	EMS	Daily	–	–	–
	City of Bergen	ED	n.a.	n.a.	n.a.	n.a.
Spain	Autonomous Region of Cantabria	EMD	Daily	Daily	Daily	Daily
	Autonomous Region of Cantabria	EMS (EP)	Weekly	Weekly	Weekly	Weekly
	Autonomous Region of Cantabria	ED	Daily	Daily	Daily	Daily
Turkey	City of Antalya	EMD	Monthly	n.a.	n.a.	n.a.
	n.a.	EMS	n.a.	n.a.	n.a.	n.a.
	n.a.	ED	n.a.	n.a.	n.a.	n.a.

EMD, Emergency medical dispatch, EMS, emergency medical services, ED, emergency department; EP, emergency physician; n.a., information not available; –, data not available.

RESULTS

SyS system concept

Availability of emergency data

Routine electronic data was available daily in 11 of 12 regions from EMD, EMS and/or ED (Table 3). Information on the patients' chief complaints was available daily and electronically in ten systems, information on age and sex in nine systems (Table 3). Although the datasets comprised common data fields across Europe such as date, age, sex, and diagnostic information, the items were defined differently.

https://doi.org/10.1017/S0950268813003452 Published online by Cambridge University Press

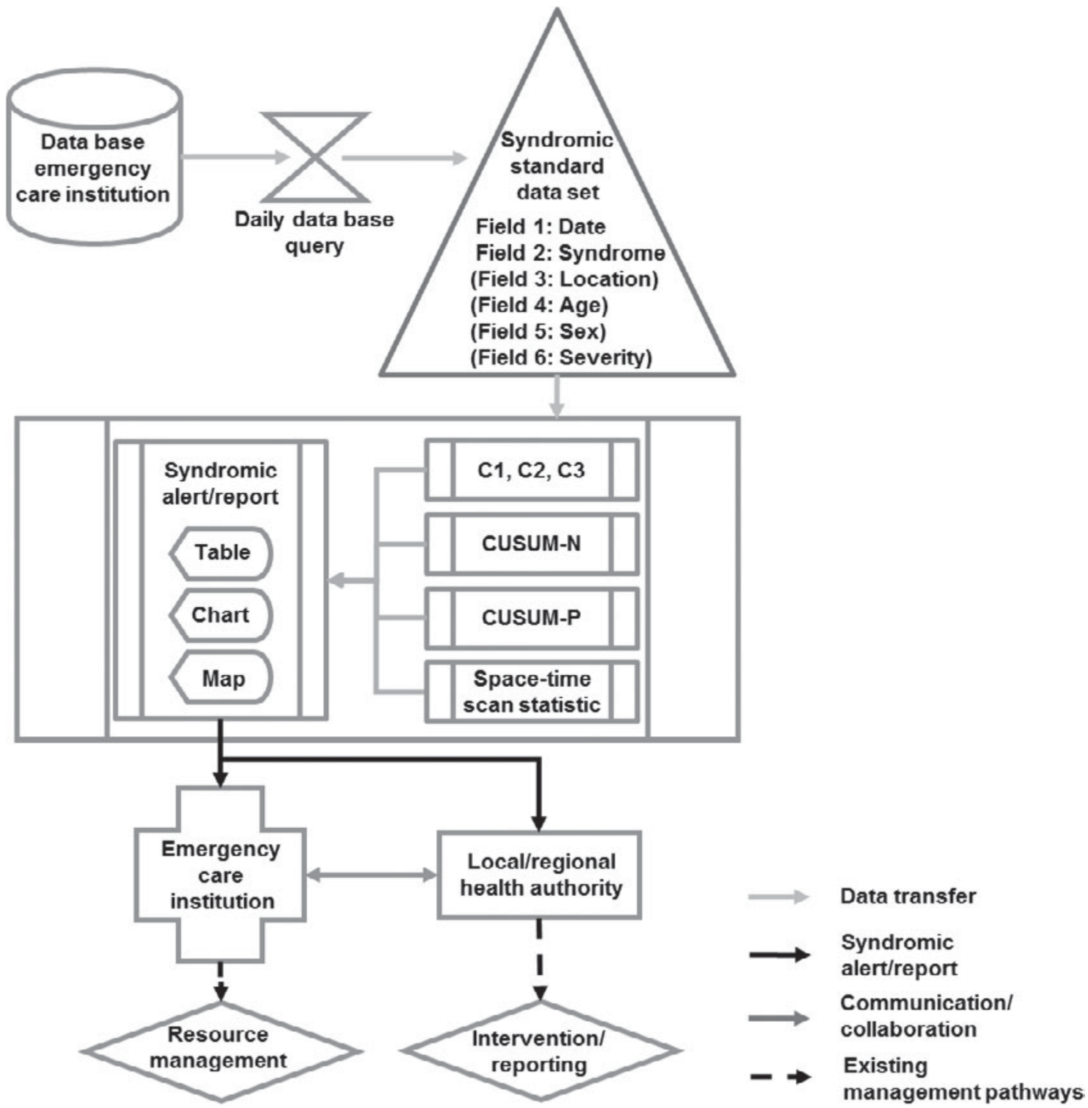

Fig. 1. Concept of an automated emergency data-based SyS system.

In particular, diagnostic information varied. Sometimes international coding systems were used, and sometimes data was collected following regional or national coding systems (Table 1).

SyS system design

We defined a standard dataset for SyS that can be generated based on routine data collected in the majority of EMD, EMS, or ED across Europe: (1) date, (2) syndrome, (3) geographical reference, (4) modifier I: age, (5) modifier II: sex, (6) modifier III: severity.

Figure 1 shows the generic functions and data flow of the automated SyS system. The system can be implemented by emergency institutions using the institution's already established health information technology infrastructure. The emergency institution is supposed to programme a permanent, daily translation between the emergency database and the surveillance system following the standard SyS dataset, e.g. an extract transform load (ETL) process. Afterwards, the syndromic data should automatically be analysed by applying temporal and spatio-temporal aberration detection algorithms in parallel. The proposed algorithms can be operationalized using open source software such as R [21] and SaTScan, or can be programmed directly in other, already applied data analysis software. The parameters of the algorithms have to be calculated once for each monitored syndrome and each emergency dataset, based on historical emergency data. During regular operation of the SyS system, these parameters

https://doi.org/10.1017/S0950268813003452 Published online by Cambridge University Press

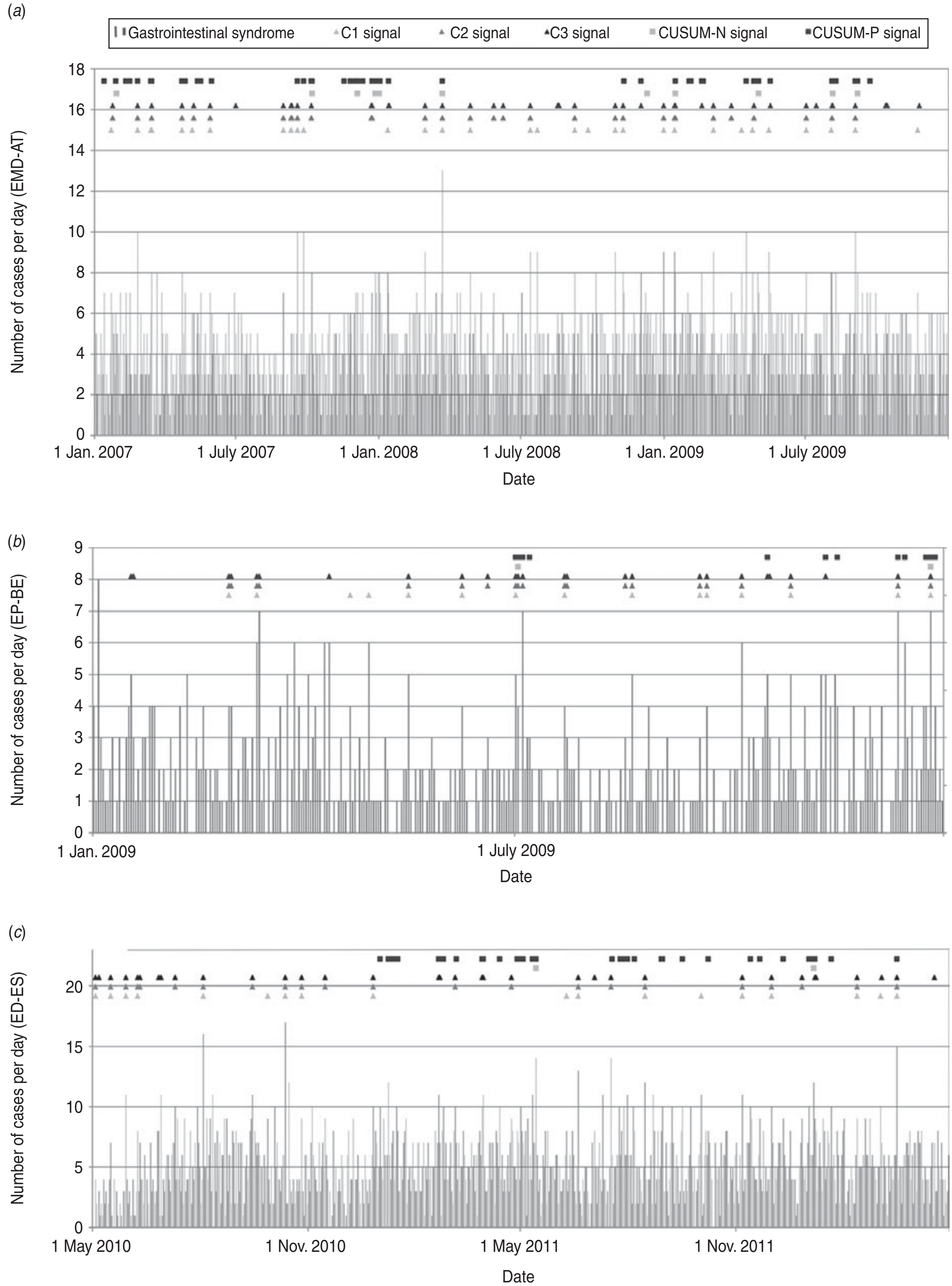

Fig. 2. Time series and temporal aberration detection algorithm signals for local gastrointestinal cases in three regions: (*a*) state of Tyrol, Austria, based on emergency medical dispatch data (EMD-AT); (*b*) Belgium (national coverage) based on data from emergency medical services (EMS) staffed with emergency physicians (EP-BE); (*c*) city of Santander, Spain, based on emergency department data (ED-ES).

https://doi.org/10.1017/S0950268813003452 Published online by Cambridge University Press

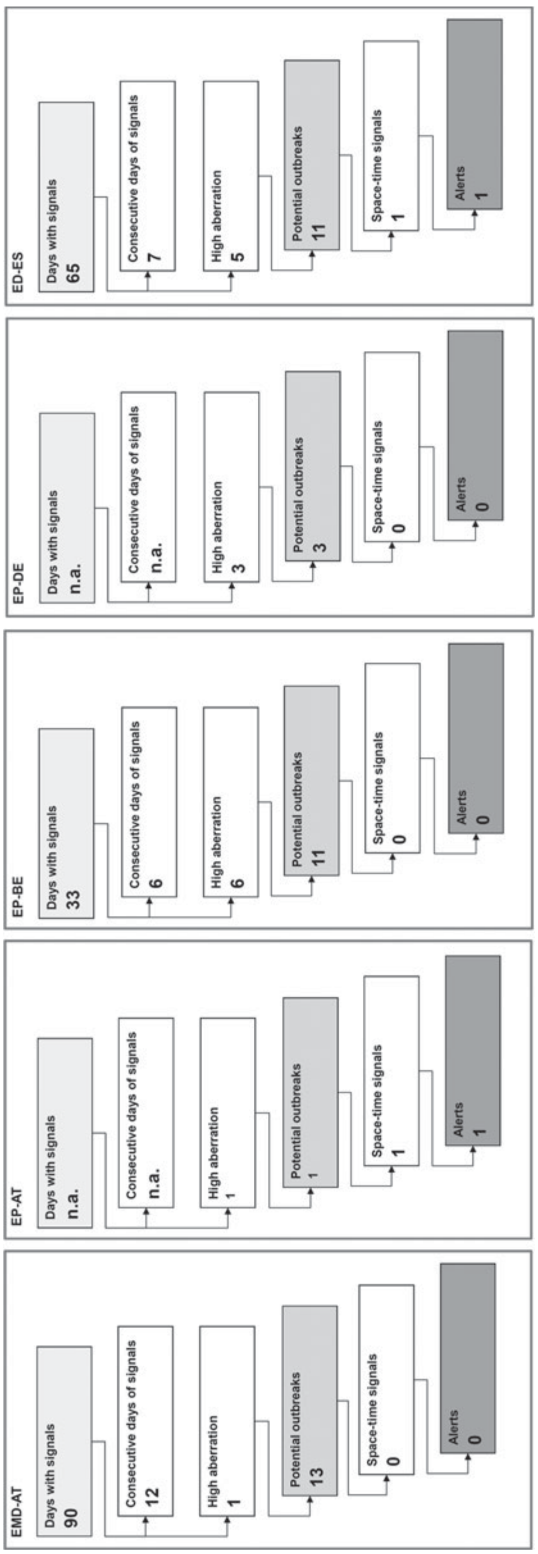

Fig. 3. Decision tree from signal to alert in four regions for the detection of local gastrointestinal outbreaks.

should be updated regularly and after changes in the data collection procedure. The outputs of the SyS analyses are statistical signals that can be displayed in tables, charts and maps, which can be disseminated within the emergency-care institution and to the local/regional public health authority. Reporting can be accomplished by establishing a regular automatic email message, by incorporating the results in already established reports, or by allowing stakeholders to access a virtual dashboard online that is automatically updated on a regular basis. The public health authority and/or emergency institution decide if the signals could represent a real event following a pre-defined decision tree for each syndrome. The public health authority can incorporate SyS alerts into existing surveillance systems and response procedures. The emergency institution can use the information for resource planning. At the time of writing this paper, two institutions have implemented an automatic SyS system following this concept, the EMD centre of the State of Tyrol, Austria, and the ED of the University Hospital in Santander, Spain.

Case study on local gastrointestinal outbreak detection

The case study showed that the case numbers in the datasets based on data from EMS staffed with EP in the Austrian (EP-AT dataset) and the German (EP-DE dataset) regions, with an average of 0·14 and 0·31 cases per day, respectively, were too low for providing valid results based on the temporal aberration detection analysis. Figure 2 shows the time series of the number of gastrointestinal syndrome cases and the signals of the temporal aberration detection analyses for the EMD-AT, EP-BE and ED-ES datasets.

The temporal aberration detection analyses resulted in many signals. When applying the decision tree to identify outbreaks, there were many events with high aberration from the mean and with signals on at least 2 consecutive days. When applying the spatio-temporal analysis during these outbreak periods, we were able to further narrow down the number of relevant outbreaks. Figure 3 provides an overview on the number of signals and the application of the decision tree for each dataset. One outbreak was located in Tyrol, Austria (EP-AT) (14 February 2007, 12 cases within a circle of 0 km radius, $P < 0·0001$), and one in Santander, Spain (3 August 2010, seven cases within a circle of 0·68 km radius or

https://doi.org/10.1017/S0950268813003452 Published online by Cambridge University Press

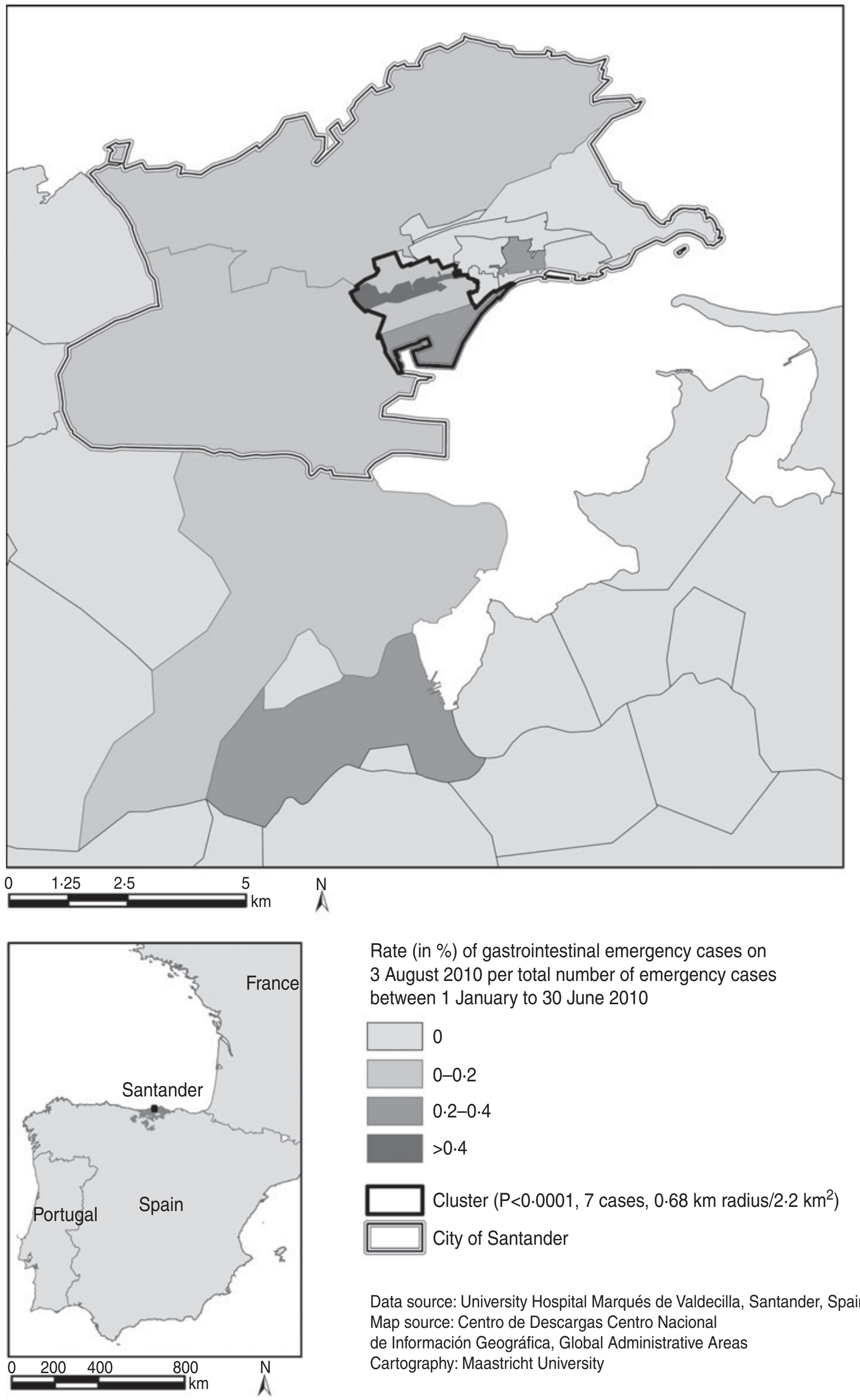

Fig. 4. Exemplary spatio-temporal cluster of gastrointestinal syndrome cases in Santander, Spain, on 3 August 2010 and rate of gastrointestinal syndrome cases on 3 August 2010 per total number of ED cases between 1 January and 30 June 2010 per postal code.

distribution across postal code areas of 2·2 km^2, $P<0{\cdot}0001$). Figure 4 exemplifies the cluster in Santander, Spain.

The comparison with notifiable disease reporting data confirmed the alert on 14 February 2007 as a norovirus outbreak in a group of foreign students

https://doi.org/10.1017/S0950268813003452 Published online by Cambridge University Press

who stayed in one hotel in the city of Kufstein ($n = 26$ cases). The alert was not confirmed by the EMD-AT dataset which refers to the same region. Two subsequent norovirus outbreaks in the following days in two foreign tourist groups in the same hotel ($n = 10$ and $n = 53$ cases) were not identified in the syndromic datasets. No other notified foodborne outbreak in Tyrol, Austria ($n = 42$), and Belgium ($n = 105$) could ultimately be linked to signals in the syndromic datasets. The reference data from Goeppingen, Germany did not provide the number of outbreaks.

DISCUSSION

SyS system concept

We developed the first concept for a SyS system based on routinely collected emergency medical care data from EMD, EMS and ED for different countries in Europe.

Routine emergency data was available in many regions in Europe in electronic form and on a daily basis. It provided relevant information for SyS, such as date and geographical information and the patients' chief complaints. We defined recommendations for syndrome coding, based on the most common coding systems in emergency care, and designed a concept for an emergency data-based SyS system able to be implemented at the local/regional level in Europe. Two regional emergency institutions in Austria and Spain have initially implemented an automatic SyS system following our concept.

As the emergency data inventory revealed differences in data coding and availability across Europe, we conceptualized the system to be implemented at single emergency institutions or in one jurisdiction. This allows for raw data to be analysed in the emergency institution, respecting data privacy. This flexibility of the concept supports a relatively rapid set-up of a SyS system as no agreements or technical connections outside of the emergency institution have to be established. The syndrome definitions based on the most common emergency-care coding systems ease the implementation and support the portability of the SyS concept across Europe. Next to the gastrointestinal syndrome, the expert consortium defined syndromes for respiratory and influenza-like illness, for heat-related illness and unspecific syndrome (= volume of medical cases without specification) [12, 22]. The results of case studies analysing these syndromes are discussed elsewhere [23, 24].

Case study on local gastrointestinal outbreak detection

Our SyS concept was tested for the detection of local outbreaks of gastrointestinal illness in four regions in Europe. In this case study, we identified two potentially relevant outbreaks. The outbreak identified in Spain could not be confirmed due to missing reference data. The alert in Austria was confirmed as a norovirus outbreak in a group of foreign students. No other notified outbreak was identified by the SyS analyses. This low validity shows that our SyS concept cannot replace traditional surveillance of gastrointestinal diseases.

Gastrointestinal diseases are often the focus of SyS applications [25], pursuing three major purposes: (i) early information on the onset of expected seasonal outbreaks such as winter vomiting disease [26], (ii) situational awareness during potentially health-threatening events such as disasters or mass gatherings [18], and (iii) detection of local gastrointestinal illness clusters [27]. Earlier studies suggested that comparatively large outbreaks at the local or regional levels were successfully detected by SyS systems [28]. Rather small outbreaks, however, appear to be difficult to detect as Xing *et al.* [29], Balter *et al.* [30] and Heffernan *et al.* [31] found based on ED data. Moreover, in our study most notified outbreaks in the study regions, which mainly consisted of few cases, were not detected by our SyS analyses. Emergency-care data, similar to other health services-based data sources for surveillance, are unlikely to reflect outbreaks with few or dispersed cases such as foodborne outbreaks comprised of visitors to a restaurant who later develop symptoms when they are in different areas [30].

Another explanation for the low validity is the fact that emergency-care data sources are not anticipated to catch all gastrointestinal outbreaks as most gastrointestinal illness patients with mild symptoms would self-treat their symptoms or utilize primary-care services. This assumption would suggest additional analysis of other data sources for SyS with a bigger coverage of mild gastrointestinal illness cases. Andersson and colleagues [32] compared three syndromic data sources able to cover people affected by gastrointestinal illness who were not seeking care in Sweden: telephone helplines, web queries and over-the-counter drug sales. This study also confirmed the finding that only larger outbreaks were detected by SyS. From nine point-source outbreaks only the four largest were detected with case numbers between 369 and 27 000. Five smaller outbreaks with case numbers between 100 and 185 were not detected. We could not test our concept on large outbreaks as no outbreaks with more than 53 cases occurred during the study period. The reference data in Belgium and Germany did not provide the number of cases per outbreak.

Emergency care especially comes into contact with gastrointestinal illness in case of severe illness, e.g. during the Shiga toxin-producing *Escherichia coli* outbreak in Germany in 2011 during which ED reported on bloody diarrhoea cases [33]. Further, emergency services are approached by gastrointestinal patients during crisis situations such as the 2003 blackout in the USA [34]. In ED in the USA, seasonal increases of gastrointestinal cases are seen during winter suggesting that gastrointestinal patients visit emergency services not only for severe illness but most likely because other health facilities are not accessible, e.g. during Christmas holidays [30]. In addition, emergency services cover patients with special characteristics, e.g. as in our case of Austria foreign tourists that might have decided to use emergency care as the easiest point of access to care. Hence, compared to other SyS data sources, emergency-care data-based SyS can have an added value for gastrointestinal surveillance if patients with severe symptoms or in special circumstances are using emergency care instead of other health services.

In the case study, we received many temporal signals for aberrations consisting of small case numbers which could not be confirmed by data from notifiable disease surveillance, which was also the case in other studies [35, 36]. This could be due to the choice or calibration of the statistical methods applied for temporal aberration detection analysis [37]. The application of other detection algorithms such as regression analysis or moving averages could yield more valid results. However, we saw the greatest potential to increase validity by additionally applying spatio-temporal detection algorithms which are expected to add information to solely temporal analyses of local gastrointestinal outbreaks as many cases tend to cluster in relatively small areas [38].

Other studies applying spatio-temporal scan statistics detected rather large or severe outbreaks [36, 39]. In order to enhance the validity of detecting small clusters, adjustment of the analysis parameters was suggested [38]. Our case study showed promising results for identifying smaller outbreaks and reducing the number of potential false alerts when applying relatively restrictive parameters to the analysis. This limited our analysis to only detect point-source outbreaks although it increased the probability of receiving alerts for true positive outbreaks. We also tested less restrictive parameters to scan for clusters up to 1 week and up to 5 km radius but found only insignificant results.

The aggregation of cases to a larger geographical area yields the problem of lower validity of the identified clusters [40]. In our case study, the Spanish study area contained both urban and rural areas with very large zip-code areas. If a cluster had been detected comprised of such a large postal code, the risk of it being a false alert is much higher compared to a cluster comprised of only small urban postal code areas. Another limitation in the applied scan statistic is the fixed circular form of the scanning window which cannot identify clusters of another shape. Flexible shapes have been tested but are not commonly used [17]. Due to high computing time we applied the prospective spatio-temporal analysis to shorter, previously defined outbreak periods based on the temporal analysis for the whole study area, which might have led to missing outbreaks that cluster in space and time, but are not visible in the purely temporal analysis. This problem would be diminished if the analyses ran automatically.

We are the first to have used run-sheet data from EMS staffed with EP for SyS. Although in two areas the case numbers were too low to perform a valid temporal aberration detection analysis, the data source appears to be promising for SyS. The true positive norovirus outbreak in Tyrol, Austria, was only captured by the data from the EP run sheets and not by the EMD data covering the same area. This indicates a higher specificity of EP-staffed EMS compared to EMD data. It also indicates that SyS based on data sources with such low case numbers tend to detect point-source outbreaks with a high number of cases rather than continuous or propagated source outbreaks with low case numbers or cases dispersed over space and time. We encourage further research using ambulance data for SyS to confirm our findings.

The case study was performed retrospectively and was not based on results from active automated SyS systems. The performance of the two currently implemented automated systems needs to be evaluated prospectively in the future to further confirm the usefulness of our concept.

CONCLUSIONS

We have provided a practical concept for implementing SyS in Europe based on routine emergency-care data from EMD, EMS and ED that can be used as supplementary and timely surveillance information source at the local/regional level. Emergency-care data-based SyS can supplement local surveillance with near real-time information on gastrointestinal patients, especially in special circumstances or with special treatment-seeking behaviour, e.g. foreign tourists. It should be able to detect large outbreaks and outbreaks comprised of patients with severe symptoms. It is not very likely to detect the majority of local gastrointestinal outbreaks with few, mild or dispersed cases. We recommend using a combination of temporal and spatial outbreak detection algorithms in parallel and to apply a decision tree for initiating public health action based on statistical signals, in order to increase the validity of SyS.

ACKNOWLEDGEMENTS

This paper describes results from the SIDARTHa project, which is an initiative of the European Emergency Data (EED) Research Network. We thank all SIDARTHa Consortium partners for their valuable contributions for conceptualizing and testing the SIDARTHa concept. Furthermore, we thank the Belgian Scientific Institute of Public Health and the Tyrolean Government for providing reference data for the case study of gastrointestinal outbreak detection and the Belgian Ministry of Health for providing national ambulance service data for Belgium. We also thank the anonymous reviewer for providing valuable comments to earlier versions of this paper.

This paper arose from the SIDARTHa project which has received funding from the European Union, within the framework of the Public Health Programme (grant agreement no. 2007208).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Triple S Project.** Assessment of syndromic surveillance in Europe. *Lancet* 2011; **378**: 1833–1834.
2. **Buehler JW, *et al.*** Situational uses of syndromic surveillance. *Biosecurity and Bioterrorism* 2009; **7**: 165–177.
3. **Ziemann A, *et al.*** Syndromic surveillance. Enhancing public health responsiveness to global change – a European perspective. *IHDP Update* 2011; **1**: 12–18.
4. **The European Surveillance System (TESSy)** (http://ecdc.europa.eu/en/activities/surveillance/tessy/pages/tessy.aspx). Accessed 18 June 2013.
5. **European Communities.** The Commission Health Emergency Operations Facility: for a coordinated management of public health emergency at EU level. Luxembourg, 2007.
6. **Influenzanet** (http://influenzanet.eu). Accessed 16 November 2013.
7. **EuroMOMO** (http://euromomo.eu). Accessed 16 November 2013.
8. **Medical Information System** (http://medusa.jrc.it/medisys). Accessed 18 June 2013.
9. **Fouillet A, *et al.*** Inventory of syndromic surveillance systems in Europe by the Triple-S project. *Emerging Health Threats Journal* 2011; **4**: 58.
10. **Krafft T, *et al.*** European Emergency Data Project (EED Project): EMS data-based health surveillance system. *European Journal of Public Health* 2003; **13**: 85–90.
11. **Das D, *et al.*** Monitoring over-the-counter medication sales for early detection of disease outbreaks – New York City. *Morbidity and Mortality Weekly Report* 2005; **54** (Suppl.): 41–46.
12. **Rosenkötter N, *et al.*** Retrospective data analysis and simulation study as basis for an automated syndromic surveillance system. Results from the SIDARTHa project. Bad Honnef: SIDARTHa scientific/technical coordination office, 2010 (http://www.sidartha.eu/docs/2007208_SIDARTHa_D6_WP6-1%20Report%20data%20analysis%20incl%20WP7%20case%20studies.pdf). Accessed 12 March 2013.
13. **Hutwagner L, *et al.*** The bioterrorism preparedness and response Early Aberration Reporting System (EARS). *Journal of Urban Health* 2003; **80**: i89–96.
14. **Burkom H.** Alerting algorithms for biosurveillance. In: Lombardo JS, Buckeridge DL, eds. *Disease Surveillance: A Public Health Informatics Approach*. Hoboken: John Wiley & Sons Inc., 2007, pp. 159–163.
15. **Lucas JM, Crosier RB.** Fast initial response for CUSUM quality-control schemes: Give your CUSUM a head start. *Technometrics* 2000; **42**: 102–107.
16. **Lucas JM.** Counted data CUSUMs. *Technometrics* 1985; **27**: 129–144.
17. **Kulldorff M.** Scan statistics for geographical disease surveillance: an overview. In: Lawson AB, Kleinman K, eds. *Spatial and Syndromic Surveillance*. Chichester: Wiley, 2005, pp. 115–131.
18. **Meyer N, *et al.*** A multi-data source surveillance system to detect a bioterrorism attack during the G8 Summit in Scotland. *Epidemiology and Infection* 2008; **136**: 876–885.
19. **Ansaldi F, *et al.*** Emergency department syndromic surveillance system for early detection of 5 syndromes: a pilot project in a reference teaching hospital in Genoa, Italy. *Journal of Preventive Medicine and Hygiene* 2008; **49**: 131–135.

20. **SurvStat.** (http://www3.rki.de/SurvStat/). Accessed 1 March 2013.
21. **R Development Core Team.** R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008.
22. **Garcia-Castrillo Riesgo L, *et al.*** The SIDARTHa coding manual. How to generate syndromes based on routinely collected emergency care data for the European syndromic surveillance system SIDARTHa. Bad Honnef: SIDARTHa Scientific/Technical Co-ordination Office, 2009 (http://www.sidartha.eu/docs/2007208_SIDARTHa_D5_WP5-2%20Report_Coding%20Manual.pdf). Accessed 12 March 2013.
23. **Schrell S, *et al.*** Local implementation of a syndromic influenza surveillance system using emergency department data in Santander, Spain. *Journal of Public Health (Oxford)* 2013; **35**: 397–403.
24. **Rosenkötter N, *et al.*** Validity and timeliness of syndromic influenza surveillance during the autumn/winter wave of A(H1N1) influenza 2009. Results of emergency medical dispatch, ambulance and emergency department data from three European regions. *BMC Public Health*. Published online: 1 October 2013. doi:10.1186/1471-2458-13-905.
25. **Buckeridge DL.** Outbreak detection through automated surveillance: a review of the determinants of detection. *Journal of Biomedical Informatics* 2007; **40**: 370–379.
26. **Loveridge P, *et al.*** Vomiting calls to NHS Direct provide an early warning of norovirus outbreaks in hospitals. *Journal of Hospital Infection* 2010; **74**: 385–393.
27. **Edge VL, *et al.*** Syndromic surveillance of gastrointestinal illness using pharmacy over-the-counter sales. A retrospective study of waterborne outbreaks in Saskatchewan and Ontario. *Canadian Journal of Public Health* 2004; **95**: 446–450.
28. **Moore KM, Edgar BL, McGuinness D.** Implementation of an automated, real-time public health surveillance system linking emergency departments and health units: rationale and methodology. *Canadian Journal of Emergency Medicine* 2008; **10**: 114–119.
29. **Xing J, Burkom H, Tokars J.** Method selection and adaptation for distributed monitoring of infectious diseases for syndromic surveillance. *Journal of Biomedical Informatics* 2011; **44**: 1093–1101.
30. **Balter S, *et al.*** Three years of emergency department gastrointestinal syndromic surveillance in New York City: what have we found? *Morbidity and Mortality Weekly Report* 2005; **54** (Suppl): 175–180.
31. **Heffernan R, *et al.*** Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases* 2004; **10**: 858–864.
32. **Andersson T, *et al.*** Syndromic surveillance for local outbreak detection and awareness: evaluating outbreak signals of acute gastroenteritis in telephone triage, web-based queries and over-the-counter pharmacy sales. *Epidemiology and Infection*. Published online: 15 May 2013. doi:10.1017/S0950268813001088.
33. **Wadl M, *et al.*** Enhanced surveillance during a large outbreak of bloody diarrhoea and haemolytic uraemic syndrome caused by Shiga toxin/verotoxin-producing Escherichia coli in Germany, May to June 2011. *Eurosurveillance* 2011; **16**(24): pii=19893.
34. **Marx MA, *et al.*** Diarrheal illness detected through syndromic surveillance after a massive power outage: New York City, August 2003. *American Journal of Public Health* 2006; **96**: 547–553.
35. **Steiner-Sichel L, *et al.*** Field investigations of emergency department syndromic surveillance signals – New York City. *Morbidity and Mortality Weekly Report* 2004; **53** (Suppl.): 184–189.
36. **Yih WK, *et al.*** Evaluating real-time syndromic surveillance signals from ambulatory care data in four states. *Public Health Reports* 2010; **125**: 111–120.
37. **Hadler JL, Siniscalchi A, Dembek Z.** Hospital admissions syndromic surveillance – Connecticut, October 2001-June 2004. *Morbidity and Mortality Weekly Report* 2005; **54** (Suppl.): 169–173.
38. **Horst MA, Coco AS.** Observing the spread of common illnesses through a community: using Geographic Information Systems (GIS) for surveillance. *Journal of the American Board of Family Medicine* 2010; **23**: 32–41.
39. **Greene SK, *et al.*** Gastrointestinal disease outbreak detection using multiple data streams from electronic medical records. *Foodborne Pathogens and Disease* 2012; **9**: 431–441.
40. **Chen D, *et al.*** Spatial and temporal aberration detection methods for disease outbreaks in syndromic surveillance systems. *Annals of GIS* 2011; **17**: 211–220.