

Automated use of WHONET and SaTScan to detect outbreaks of *Shigella* spp. using antimicrobial resistance phenotypes

J. STELLING¹*, W. K. YIH², M. GALAS⁴, M. KULLDORFF², M. PICHEL³,
R. TERRAGNO³, E. TUDURI³, S. ESPETXE⁴, N. BINSZTEIN³, T. F. O'BRIEN¹,
R. PLATT² AND Collaborative Group WHONET-Argentina

¹ Department of Medicine, Brigham and Women's Hospital, WHO Collaborating Centre for Surveillance of Antimicrobial Resistance, Boston, MA, USA

² Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

³ Departamento Bacteriología, Instituto Nacional de Enfermedades Infecciosas ANLIS 'Dr C. Malbrán', Buenos Aires, Argentina

⁴ Dirección Epidemiología, Ministerio de Salud, Argentina

(Accepted 25 August 2009; first published online 2 October 2009)

SUMMARY

Antimicrobial resistance is a priority emerging public health threat, and the ability to detect promptly outbreaks caused by resistant pathogens is critical for resistance containment and disease control efforts. We describe and evaluate the use of an electronic laboratory data system (WHONET) and a space–time permutation scan statistic for semi-automated disease outbreak detection. In collaboration with WHONET-Argentina, the national network for surveillance of antimicrobial resistance, we applied the system to the detection of local and regional outbreaks of *Shigella* spp. We searched for clusters on the basis of genus, species, and resistance phenotype and identified 19 statistical 'events' in a 12-month period. Of the six known outbreaks reported to the Ministry of Health, four had good or suggestive agreement with SaTScan-detected events. The most discriminating analyses were those involving resistance phenotypes. Electronic laboratory-based disease surveillance incorporating statistical cluster detection methods can enhance infectious disease outbreak detection and response.

Key words: Antibiotic resistance, medical informatics, outbreaks, *Shigella*, surveillance.

INTRODUCTION

Clinical and public health microbiology laboratories worldwide constitute a rich, underutilized resource in monitoring the changing epidemiology of microbial populations and in identifying emerging public health

threats. Through the routine capture and interpretation of diagnostic test results such as organism identification and antimicrobial resistance phenotypes, electronic laboratory-based surveillance can improve the diagnostic specificity and response time of traditional infectious disease surveillance [1–3]. The ability to promptly detect outbreaks caused by highly resistant pathogens is of particular importance.

To realize the full potential of microbiology laboratory data for disease outbreak detection on a national

* Author for correspondence: Dr J. Stelling, Microbiology Laboratory, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA.
(Email: jstelling@whonet.org)

scale, new tools and strategies for data management, analysis, interpretation, and communication are required. Visual monitoring of weekly case counts for pre-defined geographic units is inadequate for tracking hundreds of microbial species or thousands of phenotypically distinct pathogenic strains which do not respect artificial geographic boundaries. Time-trend analysis and other temporal statistics have been applied to laboratory data for the detection of public health outbreaks in the community [4, 5] and hospital setting [6, 7], but these do not address the geographic component of pathogen emergence and spread. Incorporating both time and geography, the prospective space-time permutation scan statistic [8], which applies minimal assumptions concerning the time, length, and geographical size of potential outbreaks, is increasingly used for the early detection of disease outbreaks of public health importance [9–11].

The adoption of a uniform microbiology database by laboratories across Argentina through the Collaborative Group WHONET-Argentina provides a unique opportunity to apply such outbreak detection methods to laboratory data on a national scale and evaluate their performance. We selected *Shigella* sp. as a pathogen for an initial evaluation of the approach because of the public health importance of shigellosis outbreaks in Argentina and the particular challenge of detecting outbreaks in an organism which exhibits strong seasonal dependency. The early detection of outbreaks caused by foodborne and waterborne pathogens, such as *Shigella*, is key to tracing sources of contamination and implementing control measures. Because of the increased morbidity and mortality associated with multidrug-resistant *Shigella* infections [12, 13], recent reports documenting the first appearance of highly resistant *Shigella* clones in Argentina are a cause for concern [14–16].

Our goal was to develop and assess the performance of WHONET and SaTScan for the detection of outbreaks of antimicrobial-resistant *Shigella* in Argentina as a first evaluation of the potential utility of such a system for other pathogens and locations. We were also interested in assessing the role of antimicrobial resistance phenotypes as a practical and specific marker for early detection of outbreak clones. This work was conducted under the National Institutes of Health Models of Infectious Disease Agent Study (MIDAS), a consortium whose goal is to develop and apply modelling methods for understanding the spread and detection of infectious agents.

METHODS

WHONET-Argentina laboratory data

The Collaborative Group WHONET-Argentina was established in 1986 under the coordination of the Antimicrobials Service of the Instituto Nacional de Enfermedades Infecciosas ANLIS ‘Dr C. Malbrán’ (INEI) to support laboratory training, epidemiological and molecular research, and public health surveillance of antimicrobial resistance [17]. The network comprises 70 microbiology laboratories representing the 23 provinces of Argentina as well as the federal capital of Buenos Aires. Isolate-level data with available patient demographics and sample details are entered into WHONET, a free software developed by our group at the WHO Collaborating Centre for Surveillance of Antimicrobial Resistance and distributed through the World Health Organization to support antimicrobial resistance surveillance in more than 90 countries involving over 1000 clinical, public health, veterinary, and food microbiology laboratories [18, 19].

Participation in WHONET-Argentina is contingent upon ongoing acceptable performance in semi-annual proficiency tests of species identification and antimicrobial susceptibility testing. A total of 38 of the 70 WHONET-Argentina laboratories (Fig. 1) were included in this study based on compliance with national protocols for susceptibility testing, capacity for identification of *Shigella* isolates, and completeness of data entry for the period July 2005–June 2007. Data from 29 laboratories were available for the whole period, and data from an additional nine were available from 1 January 2006 onward. The primary reasons for excluding laboratories were insufficient data because of incomplete data entry (14 laboratories) or enrolment in the WHONET-Argentina network after January 2006 (15 laboratories). Three laboratories were excluded because *Shigella* isolates were usually identified only to genus level.

Antimicrobial susceptibility testing was performed by disk diffusion or broth microdilution in accordance with guidelines of the Clinical and Laboratory Standards Institute (CLSI) [20]. CLSI interpretive criteria were used for all antimicrobials with the exception of fosfomycin, for which recommendations of the French Society of Microbiology were used [21]. Susceptibility to six antimicrobials was tested by participating laboratories: ampicillin (AMP), ciprofloxacin (CIP), fosfomycin (FOS), nalidixic acid (NAL), nitrofurantoin (NIT), and trimethoprim/sulfamethoxazole (SXT).

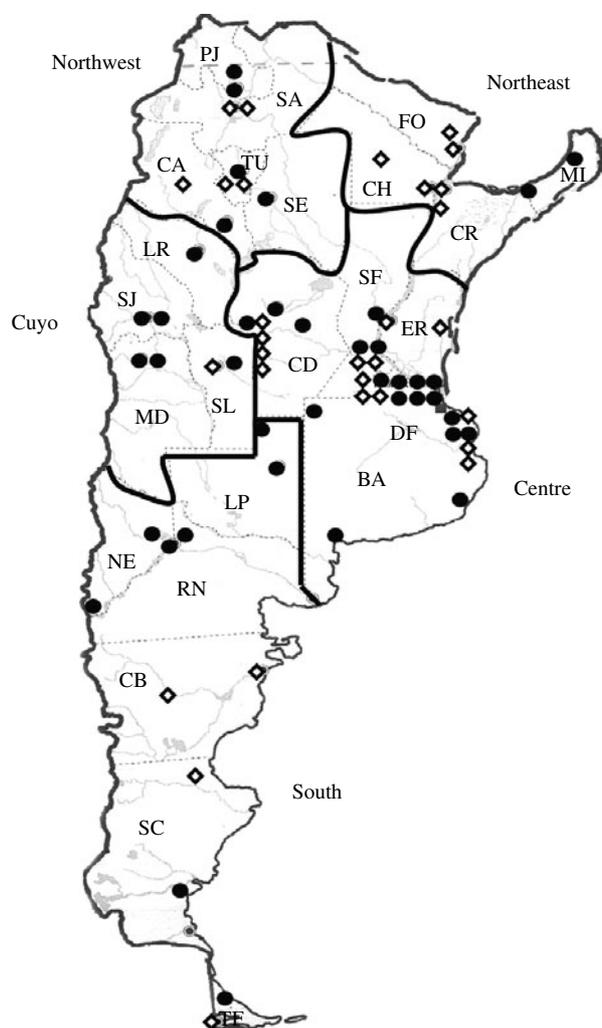


Fig. 1. Geographic distribution of laboratories in the Collaborative Group WHONET-Argentina. The 38 laboratories included in the analyses of this paper are indicated with a solid circle (●), while others are indicated with an open diamond (◇). BA, Buenos Aires (province), CA, Catamarca, CB, Chubut, CD, Córdoba, CH, Chaco, CR, Corrientes, DF, Buenos Aires (federal capital), ER, Entre Ríos, FO, Formosa, LP, La Pampa, LR, La Rioja, MD, Mendoza, MI, Misiones, NE, Neuquén, PJ, Jujuy, RN, Río Negro, SA, Salta, SC, Santa Cruz, SE, Santiago del Estero, SF, Santa Fe, SJ, San Juan, SL, San Luis, TF, Tierra del Fuego, TU, Tucumán.

Resistance was rare to four of these agents (CIP, FOS, NAL, NIT), so we restricted analyses of resistance phenotypes to the two agents to which susceptibility was variable (AMP, SXT).

We looked for outbreaks at three levels of organism identification: (1) genus – *Shigella*; (2) species – *S. flexneri*, *S. sonnei*, *S. dysenteriae*, and *S. boydii*; and (3) resistance phenotype within species: non-susceptible to AMP alone (AMP), non-susceptible to

SXT alone (SXT), non-susceptible to both agents (AMP–SXT), and susceptible to both ('Susceptible'). Resistant and intermediate isolates were categorized as 'non-susceptible' for purposes of these analyses. Insufficient data were available to perform comparable analyses using *Shigella* serotype.

Through the National Diarrhoea and Foodborne Pathogens Laboratory Network, the Enteric Pathogens Service at INEI receives isolates of *Shigella* spp. collected during outbreaks as well as a sample of sporadic strains for serotyping. A subset is further subtyped by pulsed-field gel electrophoresis (PFGE) using standardized protocols from the PulseNet International Network to establish the genetic relatedness of clones and to confirm the occurrence of suspected outbreaks [22, 23].

Space–time permutation scan statistic

The prospective detection of disease outbreaks poses specific challenges in that we know neither where nor when an outbreak may occur, nor its temporal or geographical extent. For the proposed surveillance system, we used a space–time permutation scan statistic [8] as implemented in the free SaTScan™ software [24]. The method automatically adjusts for purely geographic patterns constant over time such as population density and for purely temporal trends nationwide such as day-of-week effects or seasonal patterns of disease occurrence. This automatic adjustment for seasonal variation – through the comparison of simultaneous temporal trends over multiple geographic units – offers a significant advantage over traditional purely temporal approaches which either do not adjust for seasonal trends [6, 25] or which require several years of comparable baseline data [4, 26, 27].

The statistical likelihood that observed signals are due to chance alone is expressed in terms of a recurrence interval [28], which is the inverse of the *P* value. A signal with a recurrence interval of 180 days is of a strength that one would expect to see by chance alone about twice a year and thus consistent with expected random variation. A signal with a recurrence interval of 5000 days on the other hand has a small likelihood of occurring by chance alone (once every 14 years), and thus merits further investigation as an indication of a possible outbreak. High recurrence intervals may also be due to changes in hospital participation, specimen collection practices, or laboratory testing procedures. It is thus critical that statistical signals

Table 1. Frequency of resistance phenotypes* of *Shigella* species isolated from specimens collected July 2006–June 2007

Resistance phenotype	<i>Shigella boydii</i>	<i>Shigella dysenteriae</i>	<i>Shigella flexneri</i>	<i>Shigella sonnei</i>	<i>Shigella</i> , species unknown
Susceptible	5 (15.6%)	4 (57.1%)	158 (9.7%)	79 (23%)	11 (44%)
AMP	0 (0%)	0 (0%)	720 (44.1%)	19 (5.5%)	7 (28%)
SXT	21 (65.6%)	1 (14.3%)	94 (5.8%)	179 (52%)	5 (20%)
AMP–SXT	1 (3.1%)	1 (14.3%)	530 (32.5%)	33 (9.6%)	1 (4%)
Incomplete information	5 (15.6%)	1 (14.3%)	131 (8.0%)	34 (9.9%)	1 (4%)
Total number of isolates	32	7	1633	344	25

* Resistance profile definitions: Susceptible = Susceptible to AMP and SXT; AMP = non-susceptible to AMP, susceptible to SXT; SXT = susceptible to AMP, non-susceptible to SXT; AMP-SXT = non-susceptible to both AMP and SXT; Incomplete information = results for AMP and/or SXT are not available.

detected be further investigated through traditional epidemiological means.

WHONET-SaTScan analysis and signal consolidation

Analyses were performed using a specially adapted version of the WHONET 5.4 software which calls the SaTScan software from within. For each isolate analysed, the temporal data element was specimen collection date; the spatial data element was laboratory latitude and longitude, which were obtained using Google Earth [29]. Only patients' first isolates within a 60-day period were used in the analyses. In our simulated prospective system, we repeated the statistical analysis for each day of the surveillance period 1 July 2006 to 30 June 2007 using the previous 365 days as baseline. We used a maximum cluster length of 30 days and a maximum geographical size of 50% of the observations. The selection of parameters was supported by an optimization exercise (data not shown) which demonstrated that the events detected and conclusions drawn were robust with respect to a range of parameter settings. Recurrence intervals were calculated using 9999 Monte Carlo simulations.

In this study, we defined a signal worthy of further investigation as one with a recurrence interval of ≥ 365 days thus ignoring signals consistent with chance occurrence more often than once a year. Consecutive daily signals were consolidated into 'signal clusters'. There were many instances in which a signal cluster detected at one level of analysis, for example genus, overlapped in time and space with a signal cluster detected at another level of analysis, for example species or resistance phenotype. Such coincident clusters with overlapping isolates were grouped into epidemiological 'events' for purposes of

interpretation. SaTScan-identified events and reported outbreaks were considered concordant if both implicated the same species of *Shigella* in the same province in the same month(s) and involved, when available, the same resistance phenotype and serotype.

Comparative data on shigellosis outbreaks

The national Ministry of Health (MoH) receives reports from provincial health departments on suspected and confirmed outbreaks of shigellosis and other foodborne diseases. Although outbreaks of shigellosis are not subject to mandatory public health reporting requirements in Argentina, local and provincial public health authorities believe that the extent of reporting of known outbreaks to the national level is relatively complete. We compared the statistically identified events from our system with the MoH outbreak registry by noting instances of overlap in time and space and drawing upon additional epidemiological and laboratory data (e.g. PFGE results, epidemiological notes) as available.

RESULTS

Event characterization

There were 2041 isolates analysed in the 12-month period. Fifty-three percent were in children aged <6 years, 20% in children aged 6–18 years, 8% in adults aged >18 years, and age was unknown in 19%. Frequencies of species and resistance phenotypes are presented in Table 1, while the seasonal variability of *Shigella* isolations is depicted in Figure 2. The most common species were *S. flexneri* (80%) and *S. sonnei* (17%).

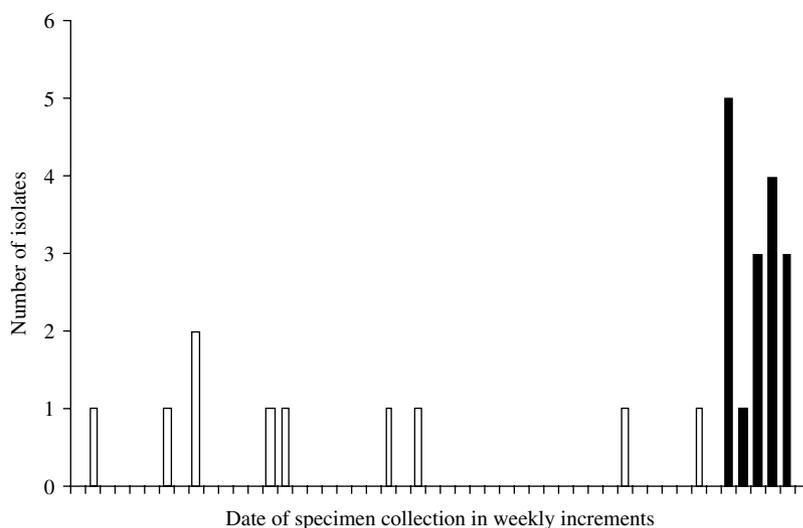


Fig. 2. Frequency distribution of *S. sonnei* isolates non-susceptible to SXT by week for the laboratory in La Pampa associated with event 5. Isolates contributing to the SaTScan event are indicated by solid bars.

Our automated system found 31 statistical signal clusters grouped into 19 epidemiological events comprising a total of 187 isolates (Table 2, Fig. 3). The species most frequently associated with events were *S. sonnei* (47%) and *S. flexneri* (42%); there was one *S. boydii* event. The number of isolates in an event ranged from 2 to 55. Most events were associated with isolates reported by a single laboratory, while four events involved two or three laboratories – three events in Buenos Aires and one in two cities (Neuquén and Cipolletti) on opposite sites of the Limay River. An example of the temporal distribution of specimen collection dates showing both sporadic and clustered isolates is shown in Figure 4 for the laboratory involved in event 5.

Given the clonal nature of most *Shigella* outbreaks, the most specific and epidemiologically meaningful analyses were those conducted with resistance phenotypes. Sixteen of the 19 events were identified in this way, and eight of these were seen only at this level. Two events were detected by species but not by resistance profile. In both of these, susceptibility testing was incomplete, precluding the possibility of detecting this outbreak by resistance profile.

Concordance between SaTScan and reported outbreaks

Outbreaks reported to the MoH during the July 2006–June 2007 surveillance period are shown in Tables 3 and 4 and Figure 3. Two of the six reported

outbreaks were concordant, as defined in the Methods section, with SaTScan events: (1) outbreak F of *S. flexneri* serotype 6 in Río Negro in April with event 13; and (3) outbreak D of *S. sonnei*–SXT in La Pampa in January–February with events 5 and 10. Event 7 in the neighbouring province of Mendoza may have been related to outbreak D in light of a note recorded at the time by local laboratory staff indicating that the suspected source of the outbreak was produce originating in Mendoza.

Molecular typing demonstrated that outbreak E in March in San Luis (which did not contribute data for use in this study), which borders both Mendoza and La Pampa, was caused by the same strain as the one implicated in outbreak D. Thus, outbreaks D and E and events 5, 7, and 10 could reflect a single regional outbreak of *S. sonnei* non-susceptible to SXT involving La Pampa, Mendoza, and San Luis.

Outbreak C of *S. boydii* serotype 2, which took place in Neuquén in January among a group of students returning from a school trip to a distant city, was followed in March by event 12 just across the river in Cipolletti – three cases of the same species, serotype, and resistance pattern. *S. boydii* is relatively rare in Argentina (<2% of *Shigella* isolations) and had rarely been reported in Neuquén or Cipolletti previously, so the circumstances are suggestive, although certainly not conclusive, of some relationship between outbreak C and event 12 despite the several-weeks gap between the two. Local authorities

Table 2. Characteristics of statistical events detected by SaTScan, July 2006–June 2007

Event no.	Analysis type	Organism	Observed serotypes	Province	Period	Date of first signal	No. of isolates	Maximum recurrence interval (days)
1	Genus	<i>S. flexneri</i>	> 3 Serotypes, primarily serotype 2	Buenos Aires (3 labs)	14 Aug.–6 Sept. 2006	6 Sept. 2006	22	455
	Species	<i>S. flexneri</i>	> 3 Serotypes, primarily serotype 2	Buenos Aires (3 labs)	14 Aug.–11 Sept. 2006	25 Aug. 2006	23	> 10 000
2	Resistance profile	<i>S. flexneri</i> :Susceptible	Serotype 2 (3), serotype 3 (1)	Buenos Aires	1–5 Sept. 2006	5 Sept. 2006	4	500
	Genus	<i>S. flexneri</i>	Unserotyped	La Rioja	16–17 Nov. 2006	17 Nov. 2006	6	> 10 000
3	Species	<i>S. flexneri</i>	Unserotyped	La Rioja	16–17 Nov. 2006	17 Nov. 2006	6	1,667
	Genus	<i>S. flexneri</i>	Serotype 2	La Pampa	25 Nov.–13 Dec. 2006	6 Dec. 2006	6	5000
4	Species	<i>S. flexneri</i>	Serotype 2	La Pampa	25 Nov.–13 Dec. 2006	12 Dec. 2006	6	2000
	Resistance profile	<i>S. flexneri</i> :AMP–SXT	Serotype 2	La Pampa	25 Nov.–13 Dec. 2006	6 Dec. 2006	6	> 10 000
5	Resistance profile	<i>S. sonnei</i> :AMP		Santa Fe	19 Dec. 2006–6 Jan. 2007	17 Jan. 2007	2	> 10 000
	Species	<i>S. sonnei</i>		La Pampa	2–31 Jan. 2007	31 Jan. 2007	15	556
6	Resistance profile	<i>S. sonnei</i> :SXT		La Pampa	22 Dec. 2006–31 Jan. 2007	10 Jan. 2007	16	1429
	Species	<i>S. sonnei</i>		La Pampa	8–12 Jan. 2007	12 Jan. 2007	3	5000
7	Resistance profile	<i>S. sonnei</i> :AMP–SXT		La Pampa	8–12 Jan. 2007	12 Jan. 2007	3	833
	Species	<i>S. sonnei</i>		Mendoza	15 Jan.–13 Feb. 2007	9 Feb. 2007	28	909
8	Resistance profile	<i>S. sonnei</i> :SXT		Mendoza	15 Jan.–20 Feb. 2007	7 Feb. 2007	28	> 10 000
	Resistance profile	<i>S. sonnei</i> :SXT		Catamarca	12–25 Feb. 2007	14 Feb. 2007	4	2500
9	Genus	<i>S. flexneri</i> , <i>S. sonnei</i>	<i>S. flexneri</i> serotype 6 (2) <i>S. flexneri</i> serotype 2 (1) <i>S. sonnei</i> (2)	La Pampa	5–6 Feb. 2007	6 Feb. 2007	5	714
10	Resistance profile	<i>S. sonnei</i> :SXT		La Pampa	21 Feb. 2007	21 Feb. 2007	2	> 10 000
11	Resistance profile	<i>S. flexneri</i> :AMP		Jujuy	6–19 Mar. 2007	19 Mar. 2007	9	588
12	Species	<i>S. boydii</i>	Serotype 2 (2), unserotyped (1)	Río Negro	13–27 Mar. 2007	17 Mar. 2007	3	> 10 000
	Resistance profile	<i>S. boydii</i> :SXT	Serotype 2 (2), unserotyped (1)	Río Negro	13–27 Mar. 2007	17 Mar. 2007	3	3333
13	Genus	<i>S. flexneri</i> (27), <i>S. sonnei</i> (4) <i>S. boydii</i> (1), <i>S. dysenteriae</i> (1)	Multiple species and serotypes	Río Negro	27 Mar.–30 May 2007	17 Apr. 2007	33	> 10 000

Species	<i>S. flexneri</i>	Multiple serotypes	Río Negro, Neuquén (3 labs)	27 Mar.–30 May 2007	17 Apr. 2007	40	>10 000
Resistance profile	<i>S. flexneri</i> :Susceptible	Primarily serotype 6	Río Negro	28 Mar.–9 May 2007	17 Apr. 2007	23	>10 000
Species	<i>S. flexneri</i>	Serotype 2	La Pampa	5 Apr.–3 May 2007	16 Apr. 2007	14	588
Species	<i>S. sonnei</i>		Río Negro	18–20 Apr. 2007	19 Apr. 2007	4	>10 000
Resistance profile	<i>S. sonnei</i> :SXT		Río Negro	18–20 Apr. 2007	20 Apr. 2007	4	3333
Resistance profile	<i>S. flexneri</i> :AMP	Serotype 2 (3), unserotyped (2)	Río Negro	17 May–15 June 2007	4 June 2007	5	1000
Resistance profile	<i>S. flexneri</i> :AMP–SXT	Serotype 2 (2), serotype 3 (1) unserotyped (4)	Buenos Aires (2 labs)	5–21 June 2007	21 June 2007	7	526
Resistance profile	<i>S. sonnei</i> :SXT		Río Negro	2–12 June 2007	16 June 2007	3	476
Resistance profile	<i>S. sonnei</i> :AMP		Buenos Aires (2 labs)	7–14 June 2007	14 June 2007	2	370

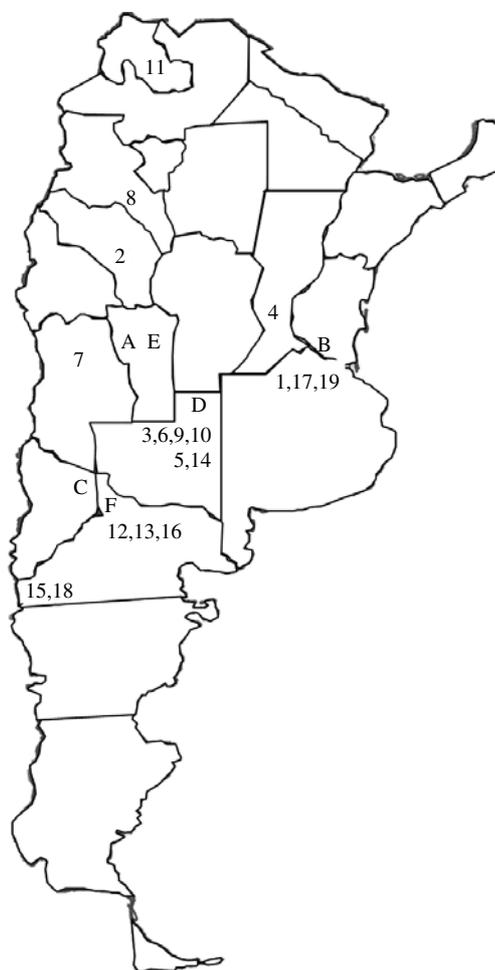


Fig. 3. Location of outbreaks reported to the Ministry of Health (A–E) and statistical ‘events’ detected by SaTScan (1–19), July 2006–June 2007.

felt that the two events were related although molecular confirmation was lacking.

In these simulated prospective analyses, SaTScan-generated signals appeared to be timely: local health authorities were notified of outbreak F on 19 April, 2 days after the first SaTScan-generated signal of event 13 in simulated prospective surveillance, and the first SaTScan signal of event 5 detected outbreak D on 10 January, a full 2 months prior to receipt on 8 March by INEI of the first isolates for confirmation.

Our analyses did not detect outbreaks A and B, and the relationship of three events with outbreak E was suggestive but not conclusive. Outbreaks A and E both occurred in San Luis, a centre excluded from the analyses presented in this study because of late enrolment in the network. The ability to detect outbreak B in Buenos Aires was probably compromised

Table 3. *Shigellosis outbreaks reported to the Ministry of Health, July 2006–June 2007*

Reported outbreak code	Organism	Province	Dates	No. of patients reported	No. of isolates received at INEI	Date of receipt of isolates at INEI	PFGE pattern (no. of isolates)
A	<i>S. flexneri</i> serotype 2	San Luis	Oct. 2006	10	2	14 Nov. 2006	ARJZXX01.0001 (2)
B	<i>S. flexneri</i> serotype 2	Buenos Aires	Dec. 2006	4	3	20 Dec. 2006	ARJZXX01.0002 (2) ARJZXX01.0003 (1)
C	<i>S. boydii</i> serotype 2	Neuquén	Jan. 2007	55	7	17 Jan. 2007	ARJZGX01.0001 (6) ARJZGX01.0002 (1)
D	<i>S. sonnei</i>	La Pampa	Jan.–Feb. 2007	No information	19	8 Mar. 2007	ARJ16X01.0014 (12) ARJ16X01.0013 (2) ARJ16X01.0015 (2)
E	<i>S. sonnei</i>	San Luis	Mar. 2007	4	2	19 Apr. 2007	ARJ16X01.0009 (2)
F	<i>S. flexneri</i> serotype 6	Río Negro	Apr. 2007	No information	6	22 May 2007	ARJZXX01.0004 (3) ARJZXX01.0005 (2) ARJZXX01.0006 (1)

PFGE patterns were labelled according to guidelines of the PulseNet Latin America and Caribbean Network [30]. The first two letters correspond to the country code (AR), then three characters for the bacterial species (JZX: *S. flexneri*; JZG: *S. boydii*; J16: *S. sonnei*), three for restriction enzyme (X01: *Xba*I) and finally the pattern number. For each outbreak listed, all isolates typed were considered to have highly similar PFGE patterns (up to two bands difference) and considered to be clonally related. PFGE patterns ARJ16X01.0014 (outbreak D) and ARJ16X01.0009 (outbreak E) were also highly related, with only two bands different.

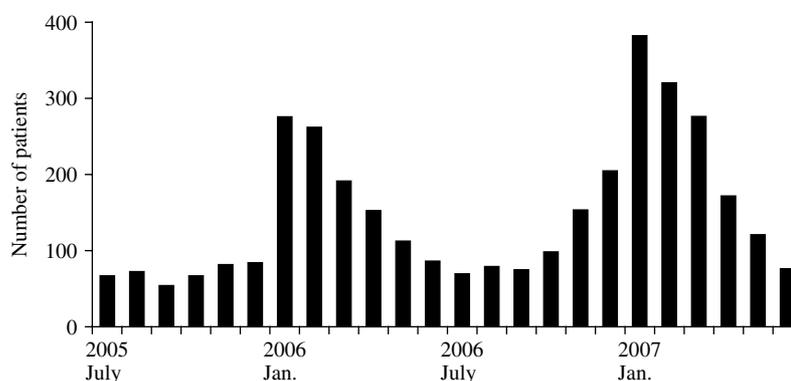


Fig. 4. Frequency distribution of *Shigella* spp. isolates included in the analyses by month July 2005 to November 2006 ($n=29$ laboratories); January 2006 to December 2007 ($n=38$ laboratories).

by the small number of patients ($n=4$) identified in a large urban area

Possible outbreaks identified by SaTScan not reported to the MoH

The epidemiological significance of the 14 SaTScan-identified events without evidence of a relationship to a known outbreak was further investigated through a review of additional details available in isolate line listings and discussions with local public health

authorities. Isolate-level review suggests that 12 of these events were largely produced by clustering of a single *Shigella* strain, and thus were possibly associated with true outbreaks, while discordant isolate details (discrepant species, serotypes, and disk diffusion zone diameters) and low recurrence intervals suggest that events 9 and 19 were spurious statistical findings of low public health interest.

For 11 of these possible outbreaks, local authorities were not aware of any outbreaks with the characteristics indicated in Table 2. They pointed out that most

Table 4. *Shigellosis outbreaks reported to the Ministry of Health and concordant or possibly related events detected by SaTScan*

Outbreak code	Outbreaks reported to Ministry of Health			Events detected by SaTScan			Comment
	Organism	Resistance phenotype	Event codes	Predominant organism	Resistance phenotype		
A	<i>S. flexneri</i> serotype 2	Not available	No concordance	No concordance	San Luis laboratory excluded from SaTScan analyses because of late enrolment		
B	<i>S. flexneri</i> serotype 2	Not available	No concordance	No concordance	Only four patients identified in Buenos Aires		
C	<i>S. boydii</i> serotype 2	Not available	<i>S. boydii</i> serotype 2	SXT	Outbreak C and event 12 in adjacent provinces with a 2-month gap – possibly related		
D	<i>S. sonnei</i>	SXT	5, 7, 10	<i>S. sonnei</i>	Events 5 and 10 concordant with outbreak D. Event 7 in adjacent province in same period – possibly related		
E	<i>S. sonnei</i>	SXT	5, 7, 10	<i>S. sonnei</i>	Outbreak E recognized as a local extension of outbreak D, and thus possibly related to events 5, 7, and 10. The San Luis laboratory was excluded from SaTScan analyses because of late enrolment		
F	<i>S. flexneri</i> serotype 6	Susceptible	13	<i>S. flexneri</i> serotype 6	Concordant		

of the events occurred during the summer months and that any excess isolates at the time were probably attributed to the usual expected rise in shigellosis cases typical of this time of year. Event 15 was interesting in that the reporting laboratory in Bariloche was very familiar with the set of patients identified – four patients with *S. sonnei* non-susceptible to SXT in a 3-day period which had struck local laboratory staff as unusual. A local investigation was initiated at that time. No epidemiological risk factors common to the patients were identified and local authorities concluded that this was not a public health outbreak and thus did not report this event to the MoH. This is an example of complete concordance between a SaTScan event and a known public health investigation, although the investigation concluded that this was not a true outbreak.

DISCUSSION

Public health strategies for the detection of outbreaks depend largely on non-statistical methods, *ad hoc* analyses, and unvalidated thresholds for action. Purely temporal statistics have been used effectively in many countries but with prior assumptions on geographic and temporal units studied and, in the case of pathogens with a strong seasonal component, requirements for several years of comparable baseline data. In this study, we have illustrated the potential use of electronic microbiology laboratory data for real-time disease outbreak detection, highlighting (1) the value of antimicrobial susceptibility test results in the detection of events of public health importance overlooked by current methods and (2) the benefit of space–time models, even when limited historical data are available. A focus on antimicrobial resistance phenotypes offers two important benefits. One is the possibility of supporting local and national resistance containment efforts through the early detection of and response to disease outbreaks due to antimicrobial-resistant pathogens and the prompt selection of appropriate antimicrobials during the outbreak period. The second potential benefit is to enhance the specificity and timeliness of outbreak detection by using the antimicrobial resistance phenotype as a marker to increase the statistical power for outbreak detection.

In this application of space-time outbreak detection algorithms for the detection of *Shigella* outbreaks utilizing the WHONET-Argentina dataset, we found good face validity in that at least two and possibly four of the six known *Shigella* outbreaks were found

by our proposed system. Critically, our system detected an additional 14 events not reported to the MoH, most of which were suggestive of true outbreaks and certainly would have merited further epidemiological investigation and molecular characterization if detected in real time. Most of the suspected outbreaks reflected changes in the relative proportion and numbers of strains previously identified at a particular locale. However, some events (e.g. event 12: *S. boydii* in Río Negro; event 4: *S. sonnei*-AMP in Santa Fe) identified species or phenotypes rarely seen at those specific sites. The proposed system thus seems to have value not only in identifying outbreaks of frequently seen strains but also in the initial appearance of species and resistance phenotypes which are novel or infrequent at a particular location.

In this study, we used historical data to mimic a real-time prospective surveillance system. Such an approach can be used retrospectively as one measure of the effectiveness of public health surveillance and control programmes and also as part of an assessment of the burden of illness attributable to outbreaks. However, a real-time system with immediate confirmation and response would be of greater value.

The ability to implement such a real-time system depends on prompt data entry, analysis, and feedback as well as complete and consistent coverage of all provinces. A pilot project for prospective surveillance of *Shigella* outbreaks was launched in January 2009 for three contiguous provinces with weekly data submissions in close coordination with local authorities responsible for outbreak investigation and control. With the experience gathered in this pilot phase, prospective surveillance is likely to be extended to other pathogens and regions of the country.

The prospects for automating national laboratory-based surveillance for disease outbreak detection seem excellent. Our results highlight the value of antimicrobial resistance phenotypes in cluster detection both as a specific marker for outbreak detection and for early detection and containment of specific resistant pathogens of public health concern.

APPENDIX

Collaborative Group WHONET-Argentina participants contributing data for this study: **Buenos Aires (federal capital):** P. O. Andrés*, D. Ballester*, E. N. Couto, A. L. Fernández*, N. A. Gómez, C. Lucero* N. Orellana, A. Procopio*, M. Quinteros,

M. Vázquez*. **Buenos Aires (province):** R. Cabrera, B. Gatti*, D. Gómez, M. Machain*, A. Pacha, G. Páez*, S. Vaylet*, C. Vescina*. **Catamarca:** M. Ferres*, P. Váldez*, **Córdoba:** M. Bottiglieri, A. M. Littvik*, T. N. López*, L. Wolff De Jacob. **Jujuy:** S. Grosso, M. Toffoli, M. S. Weibel. **La Pampa:** G. Almada*, M. G. Gau De Cornejo*, N. Moreno, A. Pereyra. **La Rioja:** S. Flores De Galimberti*. **Mendoza:** L. Balbi*, M. A. Distefano, B. García*. **Misiones:** S. Grenón, A. M. Miranda, M. Von Specht. **Neuquén:** S. Brasili*, M. R. Núñez*, L. Pianciola*. **Río Negro:** N. Blázquez*, S. De Bunder*, M. C. Carranza*, N. Castro*. **San Juan:** H. Castro*, M. López, O. R. Navarro, R. Reinoso*. **San Luis:** E. M. Fernández*, **Santa Cruz:** H. Cano*, W. Krause*. **Santa Fe:** A. Badano*, N. Borda*, A. Ernst*, E. Méndez*, A. Mollerach*, R. Notario*. **Santiago del Estero:** M. Cragolino, A. M. Nanni De Fuster. **Tierra del Fuego:** M. Vargas*, M. A. Laferrara*, **Tucumán:** M. A. Jure, H. Musa.

(* Denotes individuals from laboratories which have sent isolates to the Enteric Pathogens Service, INEI for serotyping and molecular studies.)

ACKNOWLEDGEMENTS

We extend our gratitude to all members of the Collaborative Group WHONET-Argentina and INEI staff for their ongoing efforts in support of the national surveillance programme. We thank Alejandra Corso, Allyson Abrams, Julie Dunn, and Gonzalo Vázquez-Prokopec for their important contributions to the research. We are also grateful to Susana Bruno and the Service of Antigens and Antisera of INPB-ANLIS 'Carlos G. Malbrán' for the provision of antisera for this study. Support was provided by cooperative agreement GM076672 from the National Institutes of Health under the Models of Infectious Disease Agent Study (MIDAS) programme and research project grant RR025040-02 from the National Institutes of Health.

DECLARATION OF INTEREST

None.

REFERENCES

1. Leal J, Laupland KB. Validity of electronic surveillance systems: a systematic review. *Journal of Hospital Infection* 2008; **69**: 220–229.
2. Overhage JM, Suico J, McDonald CJ. Electronic laboratory reporting: barriers, solutions and findings.

- Journal of Public Health Management and Practice* 2001; **7**: 60–66.
3. **Wurtz R, Cameron BJ.** Electronic laboratory reporting for the infectious diseases physician and clinical microbiologist. *Clinical Infectious Diseases* 2005; **40**: 1638–1643.
 4. **Hutwagner LC, et al.** Using laboratory-based surveillance data for prevention: an algorithm for detecting Salmonella outbreaks. *Emerging Infectious Diseases* 1997; **3**: 395–400.
 5. **Widdowson MA, et al.** Automated, laboratory-based system using the Internet for disease outbreak detection, the Netherlands. *Emerging Infectious Diseases* 2003; **9**: 1046–1052.
 6. **Brown SM, et al.** Binary cumulative sums and moving averages in nosocomial infection cluster detection. *Emerging Infectious Diseases* 2002; **8**: 1426–1432.
 7. **Hacek DM, et al.** Computer-assisted surveillance for detecting clonal outbreaks of nosocomial infection. *Journal of Clinical Microbiology* 2004; **42**: 1170–1175.
 8. **Kulldorff M, et al.** A space-time permutation scan statistic for disease outbreak detection. *PLoS Medicine* 2005; **2**: 216–224.
 9. **Heffernan R, et al.** Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases* 2004; **10**: 858–864.
 10. **Kulldorff M, et al.** Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *American Journal of Public Health* 1998; **88**: 1377–1380.
 11. **Mostashari F, et al.** Dead bird clusters as an early warning system for West Nile virus activity. *Emerging Infectious Diseases* 2003; **9**: 641–646.
 12. **Niyogi SK.** Shigellosis. *Journal of Microbiology* 2005; **43**: 133–143.
 13. **Niyogi SK.** Increasing antimicrobial resistance—an emerging problem in the treatment of shigellosis. *Clinical Microbiology and Infection* 2007; **13**: 1141–1143.
 14. **Andres P, et al.** Extended-spectrum beta-lactamases in *Shigella flexneri* from Argentina: first report of TOHO-1 outside Japan. *International Journal of Antimicrobial Agents* 2005; **25**: 501–507.
 15. **Pichel M, et al.** Short report: analysis of clonal relationship among *Shigella sonnei* isolates circulating in Argentina. *Epidemiology and Infection* 2007; **135**: 681–687.
 16. **Rapoport M, et al.** CMY-2-type plasmid-mediated AmpC beta-lactamase finally emerging in Argentina. *International Journal of Antimicrobial Agents* 2008; **31**: 385–387.
 17. **Rossi A, et al.** Monitoring antibiotic resistance in Argentina. The WHONET program, 1995–1996 [in Spanish]. *Revista Panamericana de Salud Pública* 1999; **6**: 234–241.
 18. **WHONET 5.4.** Software for the management and analysis of microbiology laboratory results (<http://www.who.int/drugresistance/whonetsoftware/>). Accessed 25 February 2009.
 19. **Stelling JM, O'Brien TF.** Surveillance of antimicrobial resistance: the WHONET program. *Clinical Infectious Diseases* 1997; **24** (Suppl. 1): S157–S168.
 20. **Anon.** Performance standards for antimicrobial susceptibility testing. Wayne, PA, USA: Clinical and Laboratory Standards Institute, 2007; publication no. M100-S17 (17th informational supplement).
 21. **Anon.** Recommendations 2007. Paris, France: Comité de l'Antibiogramme de la Société Française de Microbiologie, 2007.
 22. **Ribot EM, et al.** Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. *Foodborne Pathogens and Disease* 2006; **3**: 59–67.
 23. **Swaminathan B, et al.** Building PulseNet International: an interconnected system of laboratory networks to facilitate timely public health recognition and response to foodborne disease outbreaks and emerging foodborne diseases. *Foodborne Pathogens and Disease* 2006; **3**: 36–50.
 24. **SaTScan(TM) v. 7.0.** Software for the spatial and space-time scan statistics, 2006 (<http://www.satscan.org/>). Accessed 25 February 2009.
 25. **Hutwagner LC, et al.** A simulation model for assessing aberration detection methods used in public health surveillance for systems with limited baselines. *Statistics in Medicine* 2005; **24**: 543–550.
 26. **Farrington CP, et al.** A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society Series A* 1996; **159**: 547–563.
 27. **Farrington CP, Beale AD.** Computer-aided detection of temporal clusters of organisms reported to the Communicable Disease Surveillance Centre. *Communicable Diseases Report. CDR Review* 1993; **3**: R78–82.
 28. **Kleinman KP, et al.** A model-adjusted space-time scan statistic with an application to syndromic surveillance. *Epidemiology and Infection* 2005; **133**: 409–419.
 29. **Google Earth.** earth.google.com. Accessed 1 August 2008.
 30. **Swaminathan B, et al.** PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerging Infectious Diseases* 2001; **7**: 382–389.