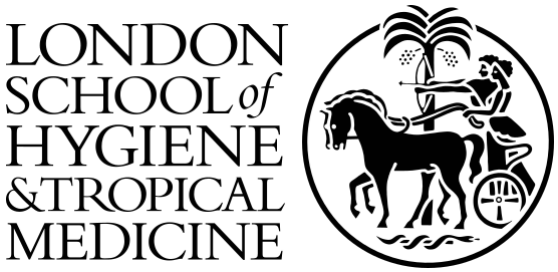


Sexually Transmitted Infections

Identifying and interpreting spatial temporal variation in diagnoses of infectious syphilis amongst men, England: 2009 to 2013

Journal:	*Sexually Transmitted Infections*
Manuscript ID	sextrans-2015-052306.R1
Article Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	PETERSEN, JAKOB; University of Essex, Institute for Social & Economic Research Gibin, Maurizio; European Commission, Joint Research Centre Sile, Bersabeh; Public Health England, Health Protection Services Simms, Ian; Public Health England, Health Protection Services
Keywords:	EPIDEMIOLOGY (GENERAL), HEALTH SERV RESEARCH, SURVEILLANCE, SYPHILIS

Identifying and interpreting spatial temporal variation in diagnoses of infectious syphilis amongst men, England: 2009 to 2013

Authors: Jakob Petersen(1), Maurizio Gibin(2), Bersabeh Sile(3), Ian Simms(3)

1. Institute for Social and Economic Research, University of Essex, Colchester, Essex, CO43SQ, UK, +44(01206)873683, jpeterb@essex.ac.uk
2. European Commission, Joint Research Centre, Isidra, Italy
3. HIV & STI Dept, Health Protection Services, Public Health England, London, UK

Key words
EPIDEMIOLOGY (GENERAL), HEALTH SERV RESEARCH, SURVEILLANCE, SYPHILIS

Abstract

Objectives

Spatial clusters and variations in the trajectory of local epidemics were explored in relation to sexual orientation, demographic factors, stage of syphilis infection and HIV serostatus.

Methods

Kulldorff's scan statistics (SaTScan) was used to distinguish endemic and temporary clusters using a two-stage analysis.

Results

Endemic areas were found in London, Manchester, Brighton and Blackpool. Up to 40% of diagnoses were found within an 11km radius of central London. Of men diagnosed with syphilis in London 80% were men who have sex with men (MSM). Annual incidence in London increased from 24 (23-26 95% CI) cases per 100000 male population in 2009 to 36 (34-38) in 2013. In comparison with clusters, endemic areas were characterised by: a significantly higher ($p<0.05$) proportion of MSM (83% compared to 73%), an increased HIV positivity (41% v 15%), age 35 to 44 (34% v 23%), lower proportion of patients born in the UK (50% v 79%), and a lower proportion of primary stage infection (40% v 47%). Space-time clusters outside endemic areas occurred in urban and rural areas, diagnoses fluctuating below 10 per month. Exponential increases in diagnoses resembling point source outbreaks were seen at two locations.

Conclusion

Control of syphilis in endemic areas has proved elusive and clusters present unique intervention opportunities. Investigating the diversity of local epidemics provides information that can be used to predict outbreak structure, plan and evaluate sexual health services, and guide public health investigation, hypothesis generation and research.

Introduction

Spatial clustering has been a key feature of the English infectious syphilis epidemic since it re-emerged in 1997 [1,2]. Increasing incidence has been characterised by outbreaks and foci, infection quickly becoming endemic in London, Manchester and Brighton [1]. Between 2009 and 2013, diagnoses in men increased by 18% (from 2507 to 2970). Most diagnoses were seen in men who have sex with men (MSM), diagnoses in this group increasing by 41% from 1697 to 2393 during this time. In contrast diagnoses seen in women decreased by 19% (345 to 279)[3]. The epidemic has been focused on infection in white MSM aged 25 to 34 many of whom have HIV co-infection and high numbers of sexual partners [1,2]. Localised clusters have followed a similar pattern although some have centred around young, socially vulnerable heterosexuals [2,4]. Contextual information collected from these clusters has provided insight into the complex local and national public health challenges associated with the syphilis epidemic but detection has relied on the vigilance of local sexual health and public health professionals. Effective control relies on early identification and case management. The aim of this study was to explore spatial clusters and variations in the trajectory of local epidemics in relation to sexual orientation, demographic factors, stage of syphilis infections and HIV serostatus.

Methods

Data sources

Diagnoses of infectious (primary, secondary, early latent) syphilis were extracted from the GUMCADv2 surveillance dataset for the calendar years 2009 to 2013 inclusive [5]. This anonymised dataset included the date of syphilis diagnoses and stage of infection as well as patient age, sexual orientation, ethnicity, country of birth and census tract of residence (2001 Middle Layer Super Output Area or MSOA01). Patients were excluded if their address was not within England (n=780, 5% of diagnoses). Population weights were used to convert the mid-year population estimates for men between 2009 and 2013 (denominator spatial unit) from the Census 2011 to the Census 2001 geography (numerator spatial unit) [6,7]. Data were geocoded to the Office of National Statistics 2011 built-up area subdivisions [8].

Statistical methods

A detailed description of the statistical methods is provided in the Web Appendix. Briefly, Kulldorff's scan statistics method (SaTScan 9.3) was used [9,10]. The methodology emulates a surveillance scenario in which monthly scans for emerging space-time clusters could be evaluated as potential outbreaks. SaTScan creates a large number of circles which assess potential clusters ranging from single census tracts to circles containing 50% of the base population. The expected number of cases was calculated from the denominators assuming a Poisson distribution. The most likely cluster was found by comparing the likelihood ratio (observed/expected) within each circle with that outside. Significance levels and recurrence intervals of the spatial clusters were obtained from Monte-Carlo replications. Covariates were not included since denominator data for MSM were not routinely available. Figure 3 was created by smoothing and mapping data using kernel density estimation [11].

Space-time cluster analysis - SaTScan

A two tier analytical strategy was adopted to prevent endemic areas from 'swamping' smaller outbreaks. Endemic areas were identified using spatial cluster analyses carried out for five calendar years in a row [12]. Data from endemic areas were then removed and prospective space-time cluster analysis was carried out for each month between January 2010 and December 2013 scanning across

a 6 months' time window. Since syphilis outbreaks can be sustained for several months the temporal window was selected at 1 to 6 months. For areas where clusters were detected time trends were smoothed using the Epanechnikov kernel function and plotted together with the results of the monthly prospective space-time analyses.

Results

Comparison of the characteristics of diagnoses made in London with the Rest of England

Between 2009 and 2013, 12521 diagnoses were reported, 6082 (49%) of which were seen in London (Table 1). Incidence in London was three times that seen in England overall. The proportion of men diagnosed with syphilis who were MSM varied from 76% to 87% in London compared to 70% to 81% in England. Diagnoses in patients born abroad accounted for nearly half of all diagnoses in London (43 to 51%) whereas for England diagnoses were higher in UK born men (59 to 65%). Modal age (30 years) was the same for men born abroad whilst the age distribution of UK born men was bimodal; peaks being seen in the late 20s and early 40s within and outwith London respectively (Figure 1). Bimodality was observed in both the MSM and non-MSM group in initial analyses. Diagnoses in men of white ethnic identity accounted for 67% to 69% of all diagnoses in London and 75 to 77% in England. Diagnoses in black men varied from 10% to 14% in London and 6 to 8% in England. In England, the majority of diagnoses of infectious syphilis consisted of primary infection, whilst there was a move to a more equal distribution between primary, secondary and early latent in London between 2009 and 2013. The proportion of syphilis diagnoses co-infected with HIV increased from 34% to 47% between 2009 and 2013. For England co-infection increased from 26% to 37% over the same period.

Table 1 Characteristics of diagnoses of infectious syphilis in men, London and England: 2009 to 2013.

	London					England Total				
	2009	2010	2011	2012	2013	2009	2010	2011	2012	2013
Diagnoses (n)	968	975	1253	1337	1549	2251	2223	2527	2640	2880
Sexual orientation										
MSM (%)	77.1	76	85.9	86.6	87.3	69.8	70.4	78.2	79.5	81
Other males (%)	22.9	24	14.1	13.4	12.7	30.2	29.6	21.8	20.5	19
Age (mean, years)	36.8	37.8	37.3	36.8	37.9	36.7	37.0	36.4	36.6	37.2
Age group (years)										
<25 (%)	9.9	8.4	7.7	9.8	6.7	13.6	13.0	14	14.4	11.8
25-34 (%)	34	31.5	33.7	34.9	33.5	30.7	31.5	33.1	33.1	33.1
35-44 (%)	35.5	34.1	37	32.2	34.3	32.1	30.6	31.1	27.9	29
45+ (%)	20.5	25.9	21.6	23	25.5	23.4	24.5	21.5	24.4	25.9
Ethnicity										
White (%)	67.9	68.9	69.2	67.2	67.2	76.7	76.2	76.2	75.4	74.8
Mixed (%)	4.3	4.1	4.9	5.2	5	2.7	2.7	3.4	3.4	3.7
Asian (%)	4.3	4.4	4.5	5.2	4.8	4.1	3.9	4.2	4.7	4.4
Black (%)	10.6	13.6	10.2	11.1	11.7	6	7.8	6.8	6.8	7.2
Other (%)*	12.8	8.9	11.1	11.3	11.3	10.5	9.4	9.4	9.7	9.9
Birthplace										
UK (%)	42.5	45	42.2	41.5	40.7	64.7	65.4	62.2	61	58.9
Abroad (%)	44	43	50.3	49.4	50.7	25.9	26.6	32.3	31.9	33.1
Other (%)	13.5	12	7.5	9.1	8.7	9.4	8	5.5	7.1	8
Infection stage										
Primary (%)	51.2	41.4	36.8	36.1	42.3	47	41.8	42.2	40.6	43
Secondary (%)	23.2	30.4	33.9	34.6	27.2	27.3	30.6	29.8	31.4	27.7
Early latent (%)	25.5	28.2	29.3	29.2	30.5	25.7	27.6	27.9	28	29.3
HIV serostatus										
Positive (%)	34.4	41.1	44	44	46.7	26.4	29.1	30.8	33.8	36.5
Annual incidence										
Patients (n)	943	945	1192	1283	1491	2207	2174	2449	2560	2799
/100k male pop	24.1	23.8	29.4	31.3	35.8	8.6	8.4	9.4	9.7	10.5
95% CI	22.6;25.7	22.3;25.3	27.8;31.1	29.6;33	34;37.7	8.3;9	8.1;8.8	9;9.7	9.4;10.1	10.2;10.9

* *Other* includes all other ethnic groups and records for which is information was missing.

Table 2 Characteristics of diagnoses in endemic and cluster areas, men: 2009 to 2013

	Cases (n)	MSM (%)	HIV-Positive (%)	Mean age (years)	<25 years (%)	25-34 years (%)	35-44 years (%)	45+ years (%)	Born UK (%)	Primary syphilis (%)	Secondary syphilis (%)	Early latent (%)
Endemic areas	**7514**	**83**	**41**	**37.0**	**10**	**34**	**34**	**23**	**50**	**40**	**30**	**30**
London	6082	83	43	37.3	8	34	35	23	42	41	30	29
Greater Manchester	1040	81	34	34.8	16	37	29	17	84	37	25	37
Brighton	252	70	45	39.4	10	24	35	31	73	33	35	32
Blackpool	140	85	21	36.0	22	26	26	25	94	34	24	43
Space-time clusters	**1146**	**73***	**15***	**36.7**	**19***	**36**	**23***	**22**	**79***	**47***	**30**	**23***
West Yorkshire	384	74	19*	35.3	16*	38	24*	21	82*	49*	27	24
Tyneside	232	74	9*	32.2	26*	38	25	11*	72*	35	32	33
Bournemouth	125	83	18*	38.8	8	36	30	25	77*	46	34	20
Bristol	74	82	11*	34.8	26*	34	15*	24	80*	73*	19	8*
Nottingham	72	51*	8*	36.3	17	33	24	26	85*	43	38	19
Herefordshire	57	81	16*	39.4	19	19	26	35	81*	70*	26	4*
Southampton	54	76	22*	37.5	17	39	19	24	65*	67*	17	17
Kingston-upon- Hull	54	50*	11*	33.3	30*	28	17*	26	82*	48	35	17
Plymouth	49	76	14*	38.1	22*	27	8*	43*	96*	18*	51*	31
Southend	45	69*	22	34.2	16	47	16*	16	73	47	27	27
Rest of England	**3861**	**65***	**18***	**36.7**	**19***	**29***	**25***	**27***	**81***	**47***	**29**	**24***
Total	**12521**	**76**	**32**	**36.8**	**13**	**32**	**30**	**24**	**62**	**43**	**29**	**28**

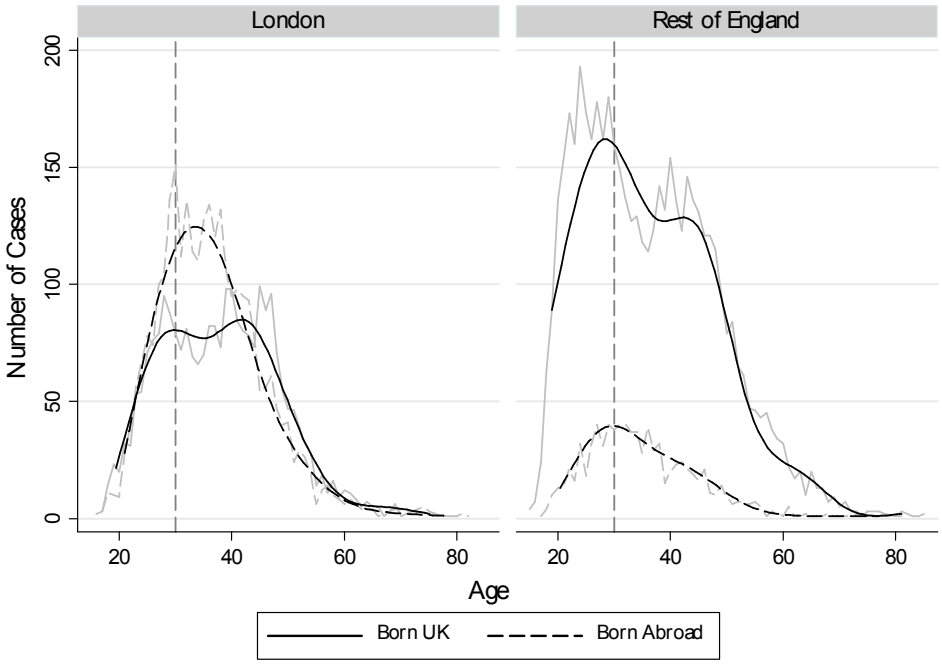
value above average

* proportion of non-endemic area significantly different from endemic area sub-total (95% level).

Characteristics of space-time endemic areas and clusters

Analysis of endemic areas and clusters revealed considerable heterogeneity (Table 2; Figures 2 and 3). Up to 40% of all diagnoses were resident within a 11km radius in Central London. Manchester, the West Yorkshire conurbation, Brighton, Tyneside, Blackpool and Bournemouth included 8% (1040), 3% (384), 2% (252), 2% (232), 1% (140) and 1% (125) of diagnoses, respectively. Sixty percent of diagnoses were seen in endemic areas whereas 9% and 31% seen in space-time clusters and the rest of England respectively (Table 2). There was a strong annual increase in diagnoses in London and a rising trend in Brighton whereas diagnoses in Manchester fell (Figure 4). Other endemic areas saw a lower but consistent burden of diagnoses: Brighton and Blackpool experienced weeks during which no diagnoses were made.

Figure 1 Diagnoses of infectious syphilis in men by birthplace & region: 2009 to 2013*



*Vertical line = age 30 years (mode).

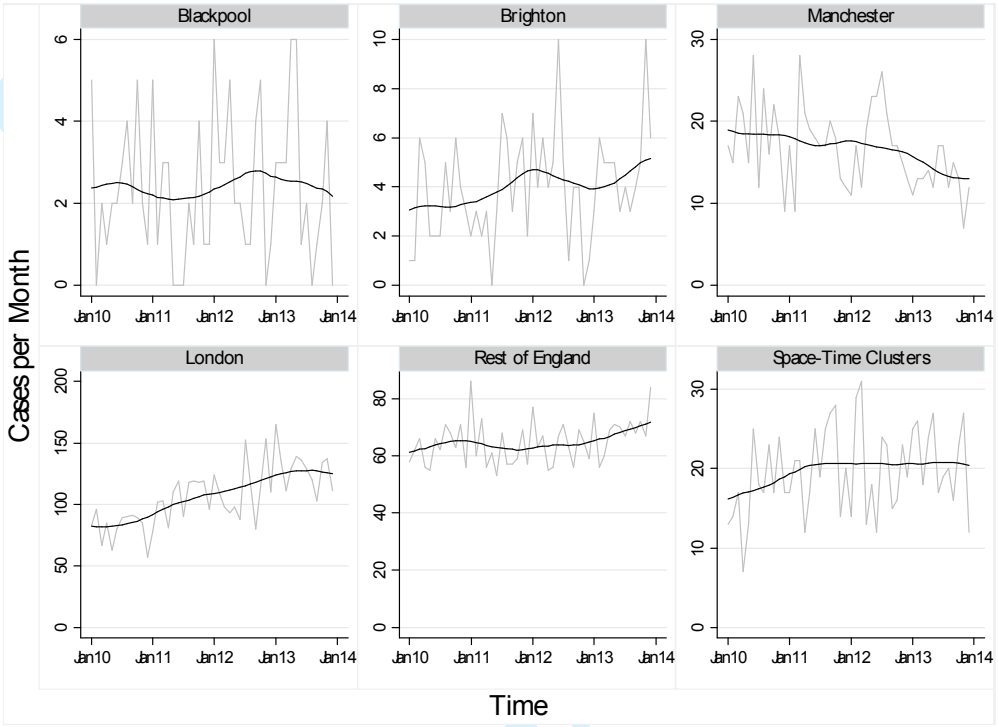
Figure 2 Space-time clusters of male infectious syphilis outbreaks: 2010 to 2013

[Map – England – attached]

Figure 3 Case density map of male syphilis cases in central London: 2009 to 2013*

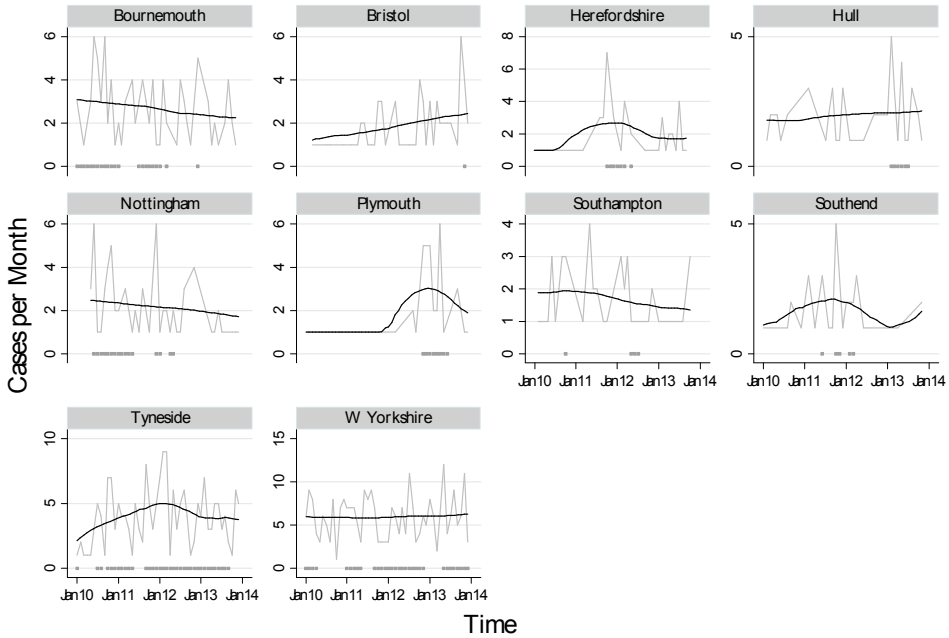
[Map – London - attached]

Figure 4 Monthly cases of infectious syphilis in males by endemic areas and the rest of England: 2010 to 2013*†



* Solid black lines represent the observations smoothed using Epanechnikov kernel function.
† The range on the y-axes varies between panels.

Figure 5 Monthly cases of infectious syphilis in men, clusters: 2010 to 2013*†



***** Area units were defined by ONS's built-up area polygons for urban areas and by local authority for rural areas.
† Solid black lines represent the observations smoothed using Epanechnikov kernel function.
…. Months where 1 to 6 months space-time clusters were detected.

In comparison with endemic areas, space-time clusters only included 1146 diagnoses (9%). Clusters occurred in urban and rural locations diagnoses fluctuating below 10 cases per month (Figure 5, the small squares on the baseline indicate when space-time clusters were identified). The exponential increase in diagnoses seen in Herefordshire and Plymouth resemble point source outbreaks. Clusters were significantly associated with a higher proportion of diagnoses made in men born in the UK and lower proportions of MSM, a lower proportion of HIV co-infection, and a higher proportion of <25 year olds, and a lower proportion of 35 to 44 year olds (Table 2).

On completion of the two tier analysis, 3861 diagnoses (31%) could not be classified into an endemic or cluster areas. Overall the characteristics of the sporadic cases were similar to those diagnoses made in cluster areas.

Discussion

Spatial-temporal variation is a critical feature of the infectious syphilis epidemic seen in men in England. London bears the heaviest burden of diagnoses in terms of incidence, density and absolute numbers. The majority of diagnoses were seen within endemic areas but these four locations differed in terms of the characteristics of each local epidemic. In contrast, space-time clusters only represent a tenth of diagnoses, developed over several months and like endemic areas showed substantial variation in the stage of infection, HIV-positivity, sexual orientation, age, and place of birth.

The main weakness of the investigation was the lack of denominator data for the MSM population which in England is focused on London, Manchester, Brighton and Blackpool. This is likely to have attenuated the strength of the detected relationships. The advantage of SaTScan analysis is that it does not require an *a priori* hypothesis about cluster location, size, or duration. The software adjusts for inhomogeneous population density, addresses multiple testing, preserves patient confidentiality at small area level and adjusts for inhomogeneous population density within the base population. In addition, since SaTScan analysis and kernel density maps are not dependent on the denominators created by administrative units, the analysis was not influenced by the modifiable areal unit problem [13,14].

During the period covered by this study a number of outbreak investigations amongst MSM and young heterosexuals were undertaken by local outbreak control teams across England. Some investigations were published but many could not be put into the public domain because of the need to protect the confidentiality of the patients who were involved in small scale investigations [15,16]. Of the space-time clusters detected all had been the subject of investigations but the the analysis did not identify two investigations based on MSM living in East Anglia [17]. The clusters may have been pushed below the limit of detection because of the small number of diagnoses involved in the outbreak. The size of the epidemic may also have been underestimated for example due to the limited success of partner notification and variations in patient clinical management between locations. Although all diagnoses conformed to British Association for Sexual Health and HIV standards, testing frequency varies across England [18]. Genitourinary medicine services within London generally undertake more intensive patient management, a process which could increase the syphilis diagnostic rate compared to HIV negative patients. This could account for the increased detection of early latent syphilis within London.

The epidemic explored here has the characteristics of a metapopulation. Essentially, populations in which syphilis is endemic sustain the epidemic and seed sexual networks in smaller satellite populations through bridging populations creating local clusters. Infection cycles relatively independently within these populations [19]. Smaller populations are more likely to become extinct, elimination occurring due to density dependent or independent factors or stochastic events. The sporadic diagnoses that could not be classified as endemic or cluster could have been recent infections that had the potential to start localised epidemics. Alternatively they may have been more long standing infections from which local sustained transmission had not emerged.

Since 2000 the epidemiology of syphilis, *Neisseria gonorrhoeae*, lymphogranuloma venereum (LGV), hepatitis C, and sexually transmissible infections such as *Shigella flexneri*, *S. sonnei* and VTEC0117 has changed significantly, influenced by population flow, sero-adaptive behaviour amongst HIV positive MSM, antimicrobial resistance together with advances in diagnostic techniques and therapeutic agents [20–22]. Chemsex drugs have become widely used and, since their first launch in early 2009, location based sexual networking applications have become an increasingly popular method of meeting sex partners and organising sex parties [23]. In effect sexual network structure is changing from a density dependent factor into a density independent factor thereby increasing the potential for the infection transmission. Such developments make infection control increasingly challenging particularly in locations where infection is endemic. Clusters can vary in size and usually take months to control. Outbreaks may be highlighted by local sexual health professionals or through exceedance reporting at Public Health England (PHE) Colindale, PHE Field Epidemiology Services and laboratories. PHE has introduced automated spatio-temporal detection tools based on SaTScan and *R* to monitor the epidemiology of gonorrhoea in England [24,25]. This tool, which is available to public health professionals, is used by local services as part of their outbreak detection strategy to increase awareness to existing and emerging sexual health challenges and inform service planning. The method is to be extended to include infectious syphilis. Outbreak detection methods also need to be fine-tuned to enable the detection and characterisation of infection clusters amongst young heterosexuals and women in low population density areas so we can better understand why outbreaks occur and assess the potential for vertical transmission.

Heterogeneity across the syphilis epidemic emphasises the importance of planning local control and prevention strategies. Here we have demonstrated how spatio-temporal analysis can be used to explore the characteristics that make locations inherently similar or unique in space and time. Such information could be used to consider the structure of future outbreaks, plan, and evaluate sexual health services, and provide a starting point for public health investigation, hypothesis generation and research. The future challenge is adapting such analytical and visualisation techniques for use with real-time GUMCADv2 reporting to provide the evidence base that enables health care professionals to respond to changes within the developing epidemic.

Key points

- Spatial-temporal variation is a critical feature of the infectious syphilis epidemic and emphasises the importance of planning local control and prevention strategies.
- Most diagnoses were seen within four endemic areas, London, Manchester, Brighton and Blackpool. The characteristics of each of these epidemics were different.
- The heaviest burden of diagnoses in terms of incidence, density and absolute numbers was seen in London.
- Space-time clusters, which included 10% of all diagnoses, were found in urban and rural areas.

- Space-time clusters developed over several months and varied in terms of stage of infection, HIV-serostatus, sexual orientation, age, and place of birth.

Acknowledgements
We would like to thank Dr Gwenda Hughes, Holly Mitchell, Martina Furegato, Stephen Duffell, Dr Nkemdiri Wheatley and Dr Hamish Mohammed for providing access to the genitourinary medicine clinic activity dataset (GUMCAD).

Funding
This work was supported by Public Health England.

Conflict of interest
None.

Contributorship statement
JP, MG and IS were closely involved with the design, conduct, analysis, presentation and interpretation of the study findings. BS extracted and quality assured the GUMCAD dataset.

References

1 Simms I, Fenton KA, Ashton M, *et al.* The re-emergence of syphilis in the United Kingdom: the new epidemic phases. *Sex Transm Dis* 2005;**32**:220–6.

2 Jebbari H, Simms I, Conti S, *et al.* Variations in the epidemiology of primary, secondary and early latent syphilis, England and Wales: 1999 to 2008. *Sex Transm Infect* 2011;**87**:191–8. doi:10.1136/sti.2009.040139

3 PHE. Sexually transmitted infections and chlamydia screening in England, 2013. 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/345181/Volume_8_number_24_hpr2414_AA_stis.pdf

4 Simms I, Bell G, Hughes G. Infectious syphilis in young heterosexuals: responding to an evolving epidemic. *Int J STD AIDS* 2011;**22**:481–2. doi:10.1258/ijsa.2011.011206

5 Savage EJ, Mohammed H, Leong G, *et al.* Improving surveillance of sexually transmitted infections using mandatory electronic clinical reporting: the genitourinary medicine clinic activity dataset, England, 2009 to 2013. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2014;**19**:20981.

6 Simpson L, Yu A. Public access to conversion of data between geographies, with multiple look up tables derived from a postal directory. *Comput Environ Urban Syst* 2003;**27**:283–307. doi:10.1016/S0198-9715(02)00018-2

7 Office for National Statistics SF. National Statistics Postcode Products. Off. Natl. Stat. 2008.http://www.ons.gov.uk/ons/guide-method/geography/products/postcode-directories/-nspp-/index.html (accessed 27 Feb2015).

8 Office for National Statistics. Built-up Area/Built-up Area Sub-division. Off. Natl. Stat. 2013.http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/built-up-areas---built-up-area-sub-divisions/index.html (accessed 27 Feb2015).

9 Kulldorff M. A spatial scan statistic. *Commun Stat - Theory Methods* 1997;**26**:1481–96. doi:10.1080/03610929708831995

10 Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *J R Stat Soc Ser A Stat Soc* 2001;**164**:61–72. doi:10.1111/1467-985X.00186

11 ESRI. *ArcGIS 10.2.2*. Environmental Systems Research Institute, Redlands, CA 2014.

12 Gesink DC, Sullivan AB, Miller WC, *et al.* Sexually transmitted disease core theory: roles of person, place, and time. *Am J Epidemiol* 2011;**174**:81–9. doi:10.1093/aje/kwr035

13 Openshaw S. *The modifiable areal unit problem*. Geo Abstracts, University of East Anglia 1984.

14 Simms I, Gibin M, Petersen J. Location, location, location: what can geographic information science (GIS) offer sexual health research? *Sex Transm Infect* 2014;:sextrans – 2014–051695. doi:10.1136/sextrans-2014-051695

15 Morgan E, Blume A, Carroll R. A cluster of infectious syphilis among young heterosexuals in south-east Hampshire. *Int J STD AIDS* 2011;**22**:512–3. doi:10.1258/ijsa.2011.010536

16 Moussa R, Sundkvist T, Emmett L. Investigation of a cluster of syphilis among heterosexuals in an English town. *Int J STD AIDS* 2011;**22**:521–2. doi:10.1258/ijsa.2011.011010

17 Welfare W, Lacey H, Lighton L, *et al.* An outbreak of infectious syphilis among young heterosexuals in an English town. *Int J STD AIDS* 2011;**22**:519–20. doi:10.1258/ijsa.2011.011031

18 Kingston M, French P, Goh B, *et al.* UK National Guidelines on the Management of Syphilis 2008. *Int J STD AIDS* 2008;**19**:729–40. doi:10.1258/ijsa.2008.008279

19 Blanchard JF, Aral SO. Emergent properties and structural patterns in sexually transmitted infection and HIV research. *Sex Transm Infect* 2010;**86 Suppl 3**:iii4–9. doi:10.1136/sti.2010.046037

20 Simms I, Gilbart VL, Byrne L, *et al.* Identification of verocytotoxin-producing Escherichia coli O117:H7 in men who have sex with men, England, November 2013 to August 2014. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2014;**19**.

21 Malek R, Mitchell H, Furegato M, *et al.* Contribution of transmission in HIV-positive men who have sex with men to evolving epidemics of sexually transmitted infections in England: an analysis using multiple data sources, 2009-2013. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2015;**20**.

22 Simms I, Field N, Jenkins C, *et al.* Intensified shigellosis epidemic associated with sexual transmission in men who have sex with men - Shigella flexneri and S. sonnei in England, 2004 to end of February 2015. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2015;**20**.

23 Bourne A, Reid D, Hickson F, *et al. The Chemsex Study: drug use in sexual settings among gay and bisexual men in Lambeth, Southwark & Lewisham*. 2014. http://sigmaresearch.org.uk/reports/item/report2014a

24 R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: : R Foundation for Statistical Computing 2015. https://www.R-project.org

25 Hungerford D, Vivanco R. Development of spatio-temporal cluster detection tool for sexually transmitted infections in the North West of England. University of Warwick, UK: 2014.

Web-resources for Methods used in:

Identifying and interpreting spatial temporal variation in diagnoses of infectious syphilis amongst men, England: 2009 – 2013.

Data source

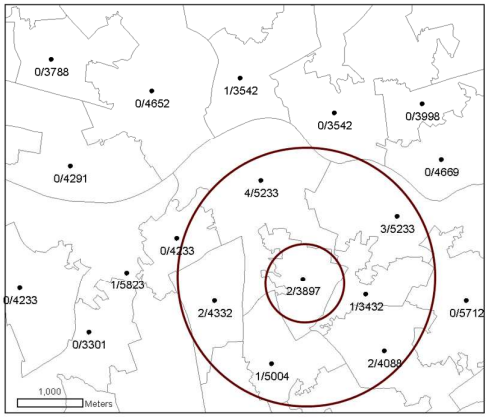
The NHS Genitourinary Medicine Clinic Activity Dataset (GUMCAD) was set up in 2008 to create a single, national database to monitor trends in STI diagnoses, to determine specific risk groups and to assess local service needs [1].

Statistical methods

A variety of space-time cluster detection techniques are available, e.g. SaTScan, ClusterSeer, Geosurveillance, and Surveillance for R; of which SaTScan is the most popular technique [2]. SaTScan has been used in hundreds of epidemiological and simulated health service datasets [3] and Kulldorff's prospective scan statistics method with discrete Poisson model (SaTScan 9.3) was used here to explore high risk clusters [4,5]. SatScan has a number of strengths: no *a priori* hypotheses about cluster location, size or duration is needed; the software adjusts for inhomogeneous population density; takes multiple testing into account and provides a single test statistics for the most likely cluster. Potential secondary clusters are also ranked.

At the start of the analytical process SaTScan creates concentric circles around a centroid. In the example, the centroid of each Middle Super Output Area is located (Figure). SaTScan sums the number of cases and population-at-risk inside and outside of the circle. Candidate circles with a higher ratio of observed-to-expected are selected. To assess how likely the results could have been generated by chance alone the clusters are tested using a large number of Monte-Carlo replications.

Figure Likelihood ratio test: identifying cases and at-risk-population. Two of the many concentric circles of varying radii evaluated are shown (fictive data).



● Middle Layer Super Output Area centroid

Space-time cluster analysis - SaTScan

For space-time analysis cluster cylinders were constructed in which height represented the time dimension. The analyses can be computationally intensive especially if a large number of geographical data points are included. Here, an exploratory analysis indicated that it was possible to

detect the same clusters at MSOA01 as at LSOA01 level and so, to reduce computation time, the analysis was performed at MSOA01 level. Large, relatively low risk clusters tend to have higher significance levels than the smaller high-risk clusters that they may contain [6]. Data from the smaller clusters with higher risk are therefore suppressed by default from the output if they fall inside a larger significant cluster. Initially, such high-risk clusters were explored by progressively restricting the maximum base population from the default to ≤0.125% (equivalent to a base population of around 30,000 men per year) by changing the output criteria of the analyses with the ≤50% base population restriction. No additional tests were carried out. The changes to clusters outside London were minimal below the 20% population threshold.

A purely spatial cluster analysis was carried out for each calendar year to identify persistent clusters over time. Persistent clusters were found each year in London, Manchester, Blackpool and Brighton. Spatial clusters were found in Bournemouth for three out of the four years studied. These clusters were interpreted as the major core areas.

Prospective space-time cluster analysis was carried out monthly between January 2010 and December 2013 to identify outbreak areas outside the core areas. The total number of consecutive tests was 48 and only clusters with recurrence intervals at or above 50 have been reported [5]. Population thresholds were evaluated and set to 1% of person-time. For these analyses core areas were removed from both the numerator and denominator side. The temporal window was selected as 1-6 months as syphilis outbreaks are known to build up over order several months. This two-tier strategy was used to avoid smaller and temporal outbreaks to be swamped by the larger and endemic core areas. Localities were identified by overlaying a map of the cluster cylinder bases with the boundary sets from ONS's 2011 built-up area and built-up area subdivisions.

Web resource references

1 Savage EJ, Mohammed H, Leong G, *et al.* Improving surveillance of sexually transmitted infections using mandatory electronic clinical reporting: the genitourinary medicine clinic activity dataset, England, 2009 to 2013. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2014;**19**:20981.

2 Robertson C, Nelson TA. Review of software for space-time disease surveillance. *Int J Health Geogr* 2010;**9**:16. doi:10.1186/1476-072X-9-16

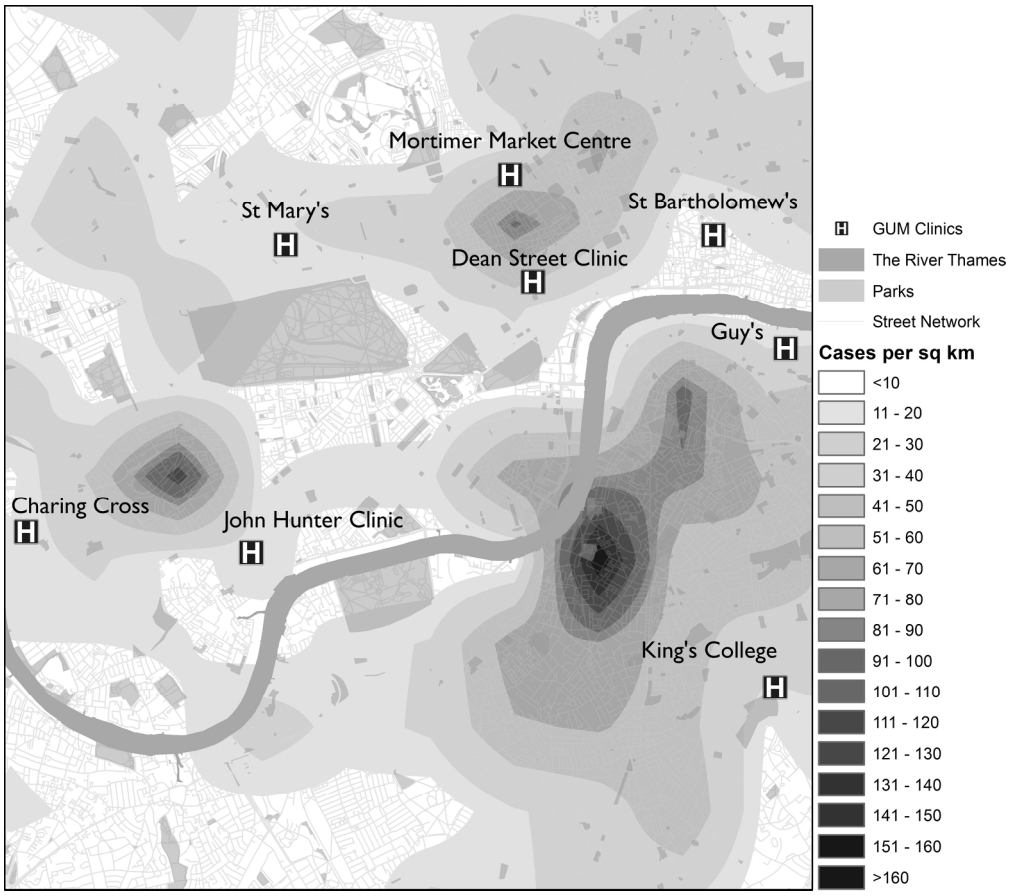
3 Kulldorff M. SaTScan User Guide. 2015.http://www.satscan.org/ (accessed 7 Nov2015).

4 Kulldorff M. A spatial scan statistic. *Commun Stat - Theory Methods* 1997;**26**:1481–96. doi:10.1080/03610929708831995

5 Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *J R Stat Soc Ser A Stat Soc* 2001;**164**:61–72. doi:10.1111/1467-985X.00186

6 Chen J, Roth RE, Naito AT, *et al.* Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. *Int J Health Geogr* 2008;**7**:57. doi:10.1186/1476-072X-7-57

297x420mm (300 x 300 DPI)



191x169mm (300 x 300 DPI)