Journal of Infection 79 (2019) 435–443

Contents lists available at ScienceDirect

Journal of Infection

journal homepage: www.elsevier.com/locate/jinf

Genomic sequencing of a national *emm*66 group A streptococci (GAS) outbreak among people who inject drugs and the homeless community in England and Wales, January 2016–May 2017

Laura Bubba[a,b,*], Nick Bundle[c,d], Georgia Kapatai[e], Roger Daniel[e], Sooria Balasegaram[c], Charlotte Anderson[c], Vicki Chalker[e], Theresa Lamagni[f], Colin Brown[e], Derren Ready[e], Androulla Efstratiou[a], Juliana Coelho[e]

[a] *Bacteriology Reference Department, Public Health England, 61 Colindale Avenue, NW9 5EQ London, UK*
[b] *European Programme for Public Health Microbiology Training (EUPHEM), European Centre for Disease Prevention and Control (ECDC), Gustav III:s boulevard 40, 169 73 Solna, Stockholm, Sweden*
[c] *United Kingdom Field Epidemiology Training Programme, Public Health England, 61 Colindale Avenue, NW9 5EQ London, UK*
[d] *Field Epidemiology Services, National Infection Service, Public Health England, Wellington, Waterloo Rd, Lambeth, London SE1 8UG, London, UK*
[e] *Respiratory and Vaccine Preventable Bacteria Reference Unit, National Infection Service, Public Health England, 61 Colindale Avenue, NW9 5EQ London, UK*
[f] *Healthcare-Associated Infection & Antimicrobial Resistance Department, National Infection Service, Public Health England, 61 Colindale Avenue, NW9 5EQ London, UK*

ARTICLE INFO

Article history:
Accepted 11 August 2019
Available online 13 August 2019

Keywords:
*emm*66
Group A streptococcus
Streptococcus pyogenes
Outbreak
Whole genome sequencing
People who inject drugs
Homeless
People reporting problematic alcohol use
England
Wales

SUMMARY

An outbreak of an uncommon emm type (emm66.0) of group A streptococcus (GAS) occurred in England and Wales between January 2016 and May 2017, involving 52 individuals who were homeless or injecting drugs users. In order to investigate the outbreak, epidemiological and network analysis were performed; moreover 55 isolates (32 outbreak, 5 non-outbreak and 13 historical – 2005–2015) were tested with whole genome sequencing (WGS), antimicrobial resistance determination, Bayesian evolutionary analysis (BEAST).

Forty one isolates (including 32 outbreak strains) belonged to a single *emm*66.0 clade (average SNP difference 6.6; range 0–16 SNPs) separate from the other isolates and two strains previously considered part of the outbreak (SNP average: 5876; range 93–8417 SNPs). Antibiotic resistance was not detected in the outbreak clone. No common source of infection was identified. WGS confirmed expansion of an *emm*66.0 clone in a hard-to-reach population and enabled refinement of the initial case definition.

Introduction

Streptococcus pyogenes or group A streptococcus (GAS) can cause several mild to life-threatening infections. Age, intravenous drug use, problematic alcohol use, immunosuppression and presence of chronic disease (diabetes mellitus, cancer, or HIV) are the main risk factors for invasive disease among adults.[1] GAS infections are defined as invasive (iGAS) based on isolation of the pathogen from blood or other sterile sites. Globally, the occurrence of iGAS has increased during the last decade with an estimate incidence from 2 to 4 per 100,000 in developed countries.[2]

In September 2016, a cluster of seven iGAS cases, of which 3 were attributed to type *emm*66, was detected by the local Health Protection Team (HPT) among people who inject drugs (PWID) or were street homeless in a city in the south of England. A local outbreak control team (OCT) was established in October 2016, whose interim finding until January 2017 were previously reported.[3] From surveillance data in England and Wales, *emm*66 iGAS infection had rarely been reported prior to 2016, in either the general community or high risk populations.[3] There are limited reports of type *emm*66 internationally.[4,5]

This study was undertaken with the objective of describing *emm*66 GAS outbreak genetic phylogeny and identifying antimicrobial resistance determinants within isolates collected during the outbreak caused by a rare *emm* type in PWID and homeless people. The outbreak isolates have been compared with those collected in the previous 10 years (2005–2015) in order to determine if the outbreak was attributed to multiple variants of *emm*66, introduction of a new strain or expansion of a circulating historical strain.

* Corresponding author at: Bacteriology Reference Department, Public Health England, 61 Colindale Avenue, NW9 5EQ London, UK.
E-mail addresses: laura.bubba@phe.gov.uk, lauretta.bubba@gmail.com (L. Bubba).

https://doi.org/10.1016/j.jinf.2019.08.009
0163-4453/

Table 1
Description of the isolates and analysis performed.

Isolate type	Isolate Total	N. of isolates	*emm* type	Collection period	Epidemiological data collection	Questionnaire issued	WGS, AMR analysis
Outbreak case	52	32	*emm* 66.0	1 Jan 2016–11 Feb 2017	Yes	Yes[b]	Yes
		20	*emm* 66.0	12 Feb 2017–23 May 2017	Yes	Yes[b]	No
Non-outbreak case	12	4	*emm* 66.0	1 Jan 2016–11 Feb 2017	Yes	No	Yes
		1	*emm* 66.1				
		7	*emm* 66.0	12 Feb 2017–23 May 2017	Yes	No	No
Historical	18	14	*emm* 66.0	1 Jan 2005–31 Dec 2015	No[a]	No	Yes
		4	*emm* 66.1				

[a] Basic demographic information only, obtained from laboratory iGAS referral forms.
[b] Questionnaire delivered to 32/52 outbreak cases.

Methods

Case numbers and analyses are listed in Table 1.

Case definitions and epidemiological information

'Outbreak cases' were defined as individuals with confirmed GAS type *emm*66.0 (invasive or non-invasive) in England and Wales, with sample date from 1 January 2016 to 23 May 2017, who were, or were epidemiologically linked to someone who was, PWID, homeless or reporting problematic alcohol use. All outbreak cases fell within 10 SNPs on average, when WGS results were available.

'Non-outbreak cases" were defined as having confirmed GAS type *emm*66 infection (invasive or non-invasive) in England and Wales, with sample date from 1 January 2016 to 23 May 2017, without the risk factors listed in the case definition or known epidemiological links to outbreak cases.

'Historical cases' were defined as any GAS *emm*66 case in England and Wales with sample date prior to the outbreak 1 January 2005 to 31 December 2015.

'SNP outbreak cluster' was defined as any isolate with WGS results collected between 2005 and February 2017 genetically clustered within the same 10 SNPs average of the defined outbreak cases cluster.

Thus isolates from cases that were epidemiologically defined as non-outbreak and historical cases, were considered as part of 'SNP outbreak cluster' whether they comply the definition.

Individuals with re-infections due to the same strain as identified on WGS were only counted once, using the date of the initial infection.

Epidemiological information on possible route of transmission and behaviour were collected and analysed as described in Bundle et al.[3] Following the outbreak alert,[6] non invasive GAS infections were also included in the analysis.

The age and sex distribution of outbreak and non-outbreak cases were compared using Mann-Whitney and Fisher's exact tests, respectively. The epidemiological analysis was undertaken in R v3.2.2[7] and Stata v14.[8] Network analysis was carried out using yED version 3.17[9] to evaluate common exposures and geographical links between outbreak cases.

Microbiological analysis

Identification and typing

Initial microbiological analysis was performed at local level and in cases of invasive infection, isolates were sent to the Public Health England (PHE) reference laboratory for further characterization. Risk factor status such as PWID was sporadically documented on isolate referral forms.

Routine *emm* typing and subtyping was undertaken on referred isolates as described in the CDC guidelines.[10]

Trends in background emm *type distribution*

Reference laboratory data from January 2005 to December 2015 were analysed to determine historic trends in iGAS *emm* types and the number of iGAS samples referred from PWID. The *emm* type distribution among PWID from 2016–2017 was reviewed.

Whole genome sequencing (WGS) analysis

Whole genome sequencing (WGS) was performed on a subset of available isolates collected from 32/52 outbreak cases, with sample collection date between 1st January 2016 and 12th February 2017 (Table 1). In the WGS analysis 5/12 non-outbreak cases and all 18 historical *emm*66 isolates collected at the national reference laboratory from 2005 to 2015 were included (Table 1).

DNA was extracted using the QIAsymphony SP automated instrument (Qiagen) and QIAsymphony DSP DNA Mini Kit after a pre-lysis step of the bacteria with mutanolysin and lysozyme. DNA library was made using Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) and sequencing performed using a HiSeq 2500 System (Illumina, San Diego, CA, USA) and the 2×100-bp paired-end mode. The quality of all sequences and the output as FASTQ files were evaluated as previously described.[11] A maximum likelihood (ML) tree was generated by first mapping the obtained reads with a complete and annotated genome of an unrelated *emm*66.0 strain from northern France.[12] SNPs were extracted by mapping with BWA to the reference genome[12] and variants using GATK. Then a multiple FASTA file including all sequences analysed using RAxML was created.[11] Only core genome was included into the SNPs analysis. Single nucleotide polymorphisms (SNP) and Bayesian evolutionary (BEAST) analyses were performed according to published methods.[13]

Antimicrobial resistance (AMR) and multilocus sequence type analysis (MLST)

Phenotypic antibiotic susceptibility was evaluated using agar dilution[14] with antibiotics and concentrations listed in supplementary Table 1. Multilocus sequence typing (MLST) profile was also performed from the WGS analysis.

Genomic analysis for the presence/absence analysis of genetic markers of antimicrobial resistance was undertaken using GeneANNOTation;[15] presence of a resistance gene was based on >90% nucleotide identity with the reference full length sequence.

Results

Epidemiological results

A total of 52 cases fulfilled the outbreak case definition whereas 12 fitted into the non-outbreak group.

Males were dominant among both outbreak (39/52) and non-outbreak cases (12/12) cases with no difference in the sex distribution ($P = 0.1044$). Outbreak cases were significantly younger (median age 38 years; range 25–68) than the non-outbreak cases (median age 69 years; range 35–91 years) ($P = 0.0004$).

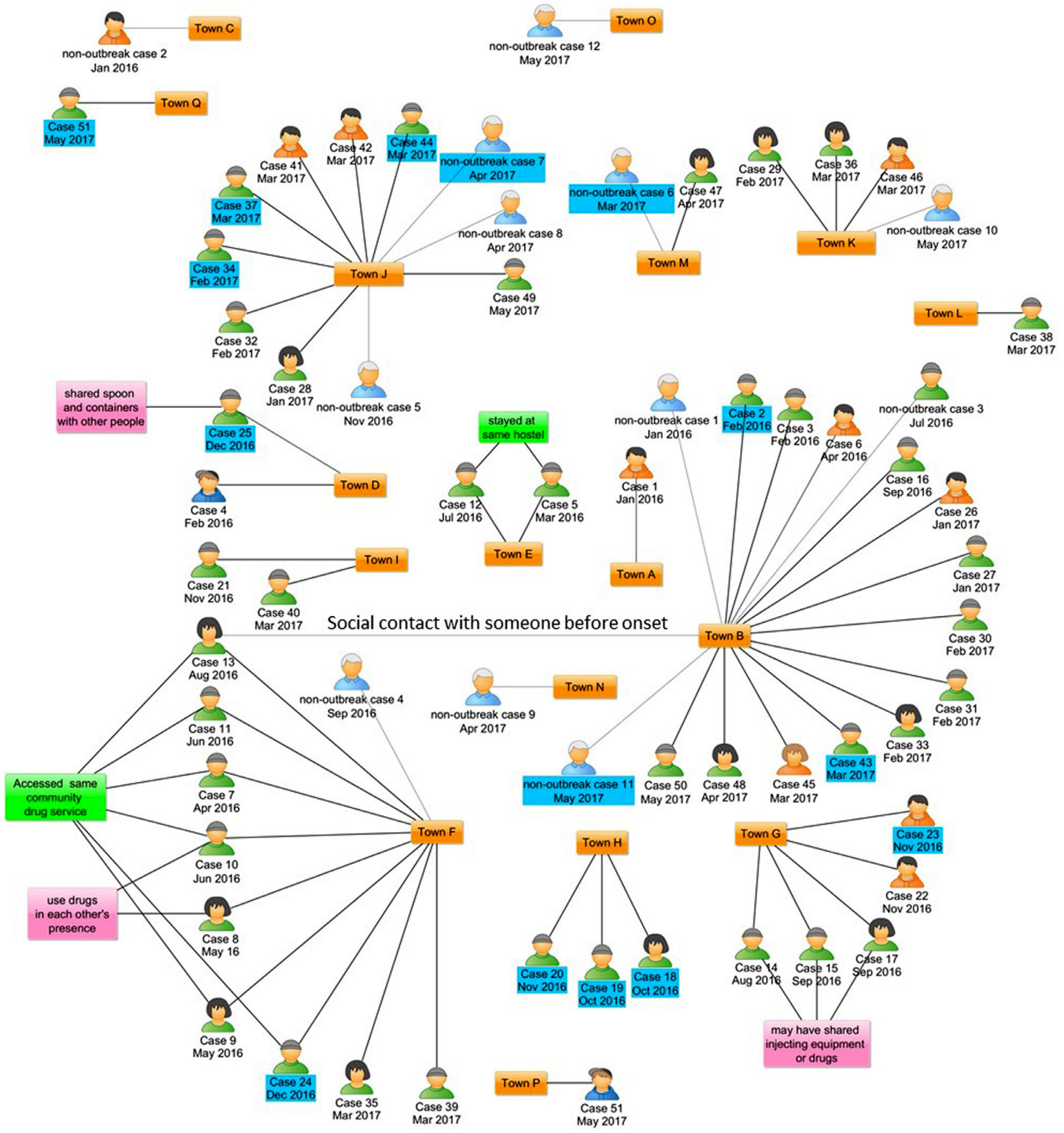

Fig. 1. Network map summarising links between cases by town, place and social contact, outbreak and non-outbreak cases of GAS and iGAS *emm*66.0 and *emm*66.1, England and Wales, 5 January 2016 to 21 May 2017.
Primary risk factor shown by shirt colour: green PWID (may share other factors); orange homeless only; blue problematic alcohol uses only; grey presence of risk factors unknown. Infection shown by shading of text – Blue: non-invasive GAS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Overall, 41/52 (79%) of the outbreak cases were PWID, of whom 24 were also homeless.

The first cluster of cases occurred in Town B in south–east England with steady geographical spread to other regions of England and South Wales and temporal clustering in many of the affected towns (Fig. 1). Network analysis (Fig. 1) revealed a number of links between outbreak cases within, but not between, towns. A potential link was identified for Case 15 (from Town F) who reported social contact with someone from Town B prior to illness. Town J, Town B, Town K and Town F had both outbreak and non-outbreak cases whereas Town C, Town N and Town O had only single non-outbreak case.

Microbiological findings

Trends in emm *type distribution*

A total of 20405 isolates were *emm*-typed between January 2005 and December 2015 by the reference laboratory, of which 177 (0.9%) had PWID on the referral form. The type distribution of iGAS isolates among the remainder, showed a predominance of *emm*1 (4981/20228; 25%) followed by *emm*3 (12%), *emm*89 (11%) and *emm*12 (9%). Among GAS/iGAS known to be from PWID from the same period *emm*89 was predominant (30/177; 17%), followed by *emm*87 (15/177, 8%), *emm*75 (13/177, 7%) and *emm*94 (3%). The distribution of *emm* types among patients identified as PWID changed when comparing the 2005–15 and 2016–17 periods, with a notable predominance of *emm*66 due to its emergence in 2016 through this outbreak (Fig. 2). *Emm*66 made up 2.3% (64/2767) of all GAS referrals to the national laboratory in 2016 and 2017 (until 23 May). In 2016, *emm*66 was the most detected type (39%) among PWID, followed by *emm*89 (14%).

Typing results

Typing with Sanger sequencing was performed on all 64 isolates collected during the outbreak period. All 52 outbreak cases and 11 non-outbreak cases were all typed as *emm*66.0 and one non-outbreak case was typed as *emm*66.1. Historical cases were typed with the same method: among the 18 historical isolates, 13 were *emm*66.0 and 5 were *emm*66.1; eight were collected in 2015 and 10 from 2005 to 2014. Information on risk factors was available for only two historical isolates: one was a homeless individual with isolate collected in 2013 and typed as *emm*66.1, the other was PWID with sample collected in 2015 and typed as *emm*66.0.

WGS analysis

WGS analysis was performed on a subset of isolates and led to a revision of the case definitions as it was previously reported into the initial description of the outbreak investigation.[3]

Following the current case definitions, the SNP outbreak cluster consisted of all outbreak cases ($n = 32$), 3 non-outbreak cases and 6 historical isolates collected in 2015 and typed with Sanger method as *emm*66.0. (Figs. 3 and 4). Overall, an average of 6.6 SNPs difference (range: 0–16 SNPs) was observed within this cluster, and 95.2 SNPs difference on average (range: 90–102 SNPs) with the three 2013 *emm*66.0 isolates included into the same branch (Figs. 3, 4 and Supplementary Table 2).

The reference strain was 32 SNPs difference on average from the SNP outbreak cluster (Fig. 4).

The remaining isolates, including the *emm*66.0 historical isolates from 2005 to 2013, the two *emm*66.1 historical isolates from 2015 and the *emm*66.1 non-outbreak isolate from 2016, were distinguished from the SNP outbreak cluster by more than 8000 SNPs on average (range: 7853–8420) (Fig. 3, Supplementary Table 2).

A total of seven 0 SNPs difference clusters were observed in six different geographical locations. Five 0 SNPs difference clusters were observed within one Town, whereas the remaining two 0 SNPs clusters included isolates from two and three different towns, respectively. The average SNP differences within the towns was less than 5 SNPs (range: 0.5–4.3); while between towns it ranged from 0.5 to 14.4 SNPs. Samples from Town G showed a range of 7.4–13.4 SNPs difference with the other cities and 0.6 SNP difference on average within the city: all cases were street homeless with three cases linked in time (August–September 2016) and through anecdotal reports of sharing needles or drugs (Fig. 1). In Town F, all seven cases included in the analysis showed a minimal SNPs difference ranging from 0 to 6 SNPs (average: 1.6 SNPs), clustering in time (April–December 2016), and six of them were linked through the same community drug service. The remaining case (no.9) reported “use of drugs in each other’s presence” with case no.11 (Fig. 1), the two showing 0 SNPs and infection less than one month apart. The two Town E cases were linked by sharing the same hostel, and the SNPs analysis showed 1 SNP difference.

The evolutionary rate was calculated as 9 bases per year according to BEAST analysis of all non-identical *emm*66.0 sequences included in the SNP outbreak cluster.

Antimicrobial resistance and MLST results

Phenotypic or genetic markers of resistance were not detected in any outbreak isolates. The same MLST profile (8,7,8,8,9,3,1), identifying the sequence type 44, was observed in all isolates clustered into the SNP outbreak cluster (supplementary Table 2).

Discussion

We detected and characterised the genomic sequences of a national outbreak caused by *emm*66.0 strains and involving 52 cases among PWID, homeless individuals and people reporting problematic alcohol use with sample collection date from January 2016 to May 2017. No single source of infection has been identified. The genomic analysis led to detection of a specific outbreak lineage and helped refine the case definition. In fact, the initial case definition considered part of the outbreak all individuals with one or more risk factors, epidemiological links and isolate resulted *emm*66 by Sanger method analysis.[3] Therefore, initially the outbreak included also one case resulted *emm*66.1, that analysed with WGS method result a different clone with more than 8000 SNPs (in Fig. 3 reported as non-outbreak case 2) and one individual with a sample typed as *emm*66.0, but with 93 SNPs difference from the outbreak cluster (in Fig. 3 reported as non-outbreak case 3)

This under-served population has already been described as susceptible to invasive GAS infections and several studies have associated the use of illicit drugs, homelessness or problematic use of alcohol with an increased risk of invasive GAS infection.[1,16,17] An outbreak of iGAS/GAS among homeless individuals in Alaska associated with another rare type, *emm*26.3, was recently described.[18] A descriptive study on a Spanish population with iGAS identified that the risk of invasive infection increased where both injecting drug use and homelessness were present;[17] this combination of risk factors was found in 46% (24/52) of our outbreak cases, although it was not possible to test for an association between homelessness, injecting drug use and iGAS infection since risk factor information was not available for our comparator historical patients. Outbreak cases were representative of the wider UK PWID population in terms of sex, age and we identified no notable changes in drug using practice in the period before illness.[19–22] Furthermore, there was no difference in terms of age, sex, mortality (only 1/52 cases died) and those described in previous studies on this population.[2,20–22]

It has been reported that the *emm* type distribution among PWID population and those without reported risk factors may differ.[23] In our setting, comparison of the *emm* distribution data, collected from the iGAS surveillance data before the outbreak detection (2005−2015), indicated that *emm*1 (30%) was predominant amongst the general community and that *emm*89 (14%) was the most frequent *emm* type among the reported PWID. In 2016–2017 the most common *emm* type among PWID was *emm*66, representing about 39% of all GAS infections reported in this population from January 2016, which may have been partially enhanced by active case-finding.

The high similarity (6.6 SNPs difference on average) of sequences from historical *emm*66.0 isolates collected in 2015 and the currently investigated SNP outbreak cases cluster suggested this *emm* clone may have appeared in the UK in 2015. Conceivably, it may have started to circulate in the south of England in 2013 and slowly spread throughout the south-centre of England

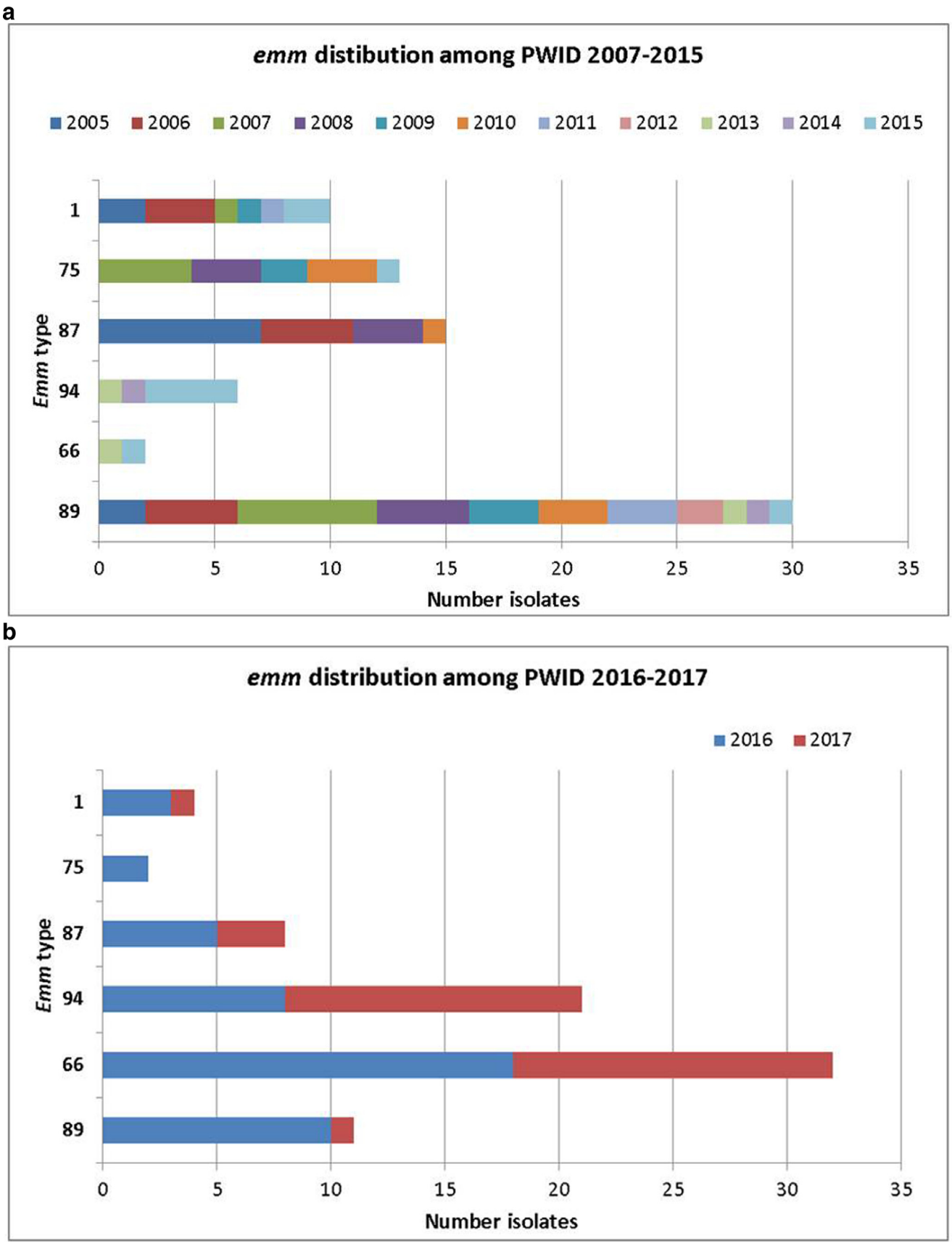

Fig. 2. Distribution of the six most common *emm* types causing iGAS in PWID, England and Wales, from January 2005 to December 2015 (A) and from 2016 to 2017 (B).

Fig. 3. Maximum likelihood phylogenetic tree including all 55 isolates sequenced both emm66.0 and emm66.1, 2005–2017. In bold are reported the outbreak, non-outbreak and historical cases fulfilling the SNP outbreak case definition, resulting all isolates within 6.6 SNPs difference on average. In italics is reported the reference strain used to build the phylogenetic tree. The historical and non-outbreak strains excluded from the outbreak cluster because showing more than 10 SNPs difference on average are also reported in regular format. The bootstrap value is reported in correspondence of the nodes. The square bracket indicates the isolates collected from outbreak cases.

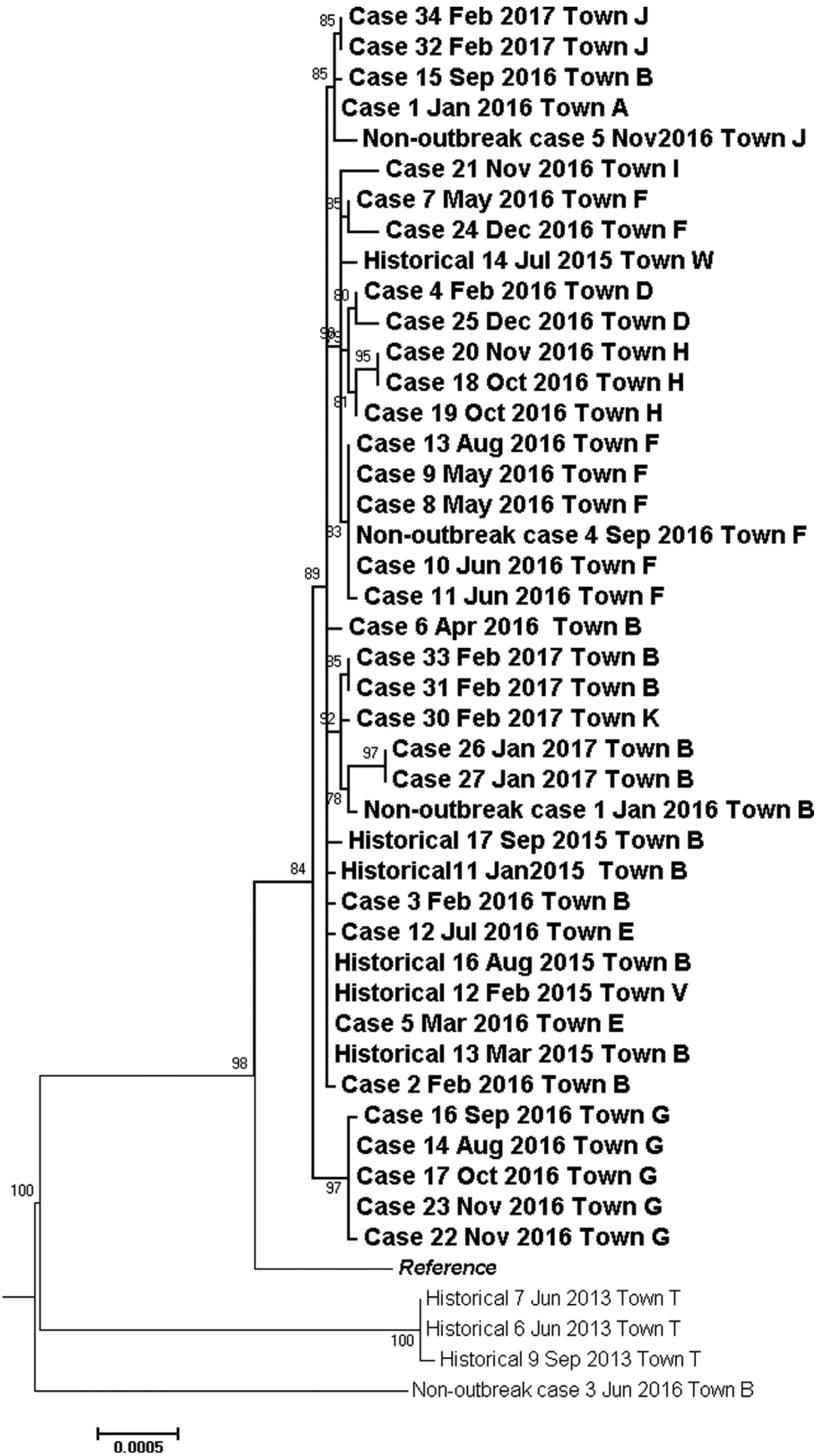

Fig. 4. Maximum likelihood phylogenetic tree including the 32 outbreak cases, the 3 non-outbreak cases and 6 historical isolates collected in 2015 that clustered together with 6.6 SNPs difference on average. In this subtree, the reference strain with 32 SNPs difference on average with the outbreak clade was reported along with 3 historical isolates collected in 2013 and one non-outbreak case with 96 and 93 SNPs difference on average with the outbreak clade.

and into Wales, continuing to spread further north since the outbreak was detected (data not shown). Lower SNP variation was observed within towns (<5 SNPs difference) suggesting local circulation. Inferred direct transmission was observed in a cluster in Town F with six patients linked by the same community drug service, of which two persons reported use of drugs in "each otherś presence", from which bacterial isolates were 0 SNP apart. Other possible sources of infection, such as contaminated drugs or a common dealer are highly unlikely for a number of reasons, including the temporal and geographical spread of cases. The non-availability of drug batch samples and wrappers for testing preclude the possibility to investigate differences to usual drug taking methods; however, the requirement of acid/heat treatment before the use of drugs would certainly compromise the ability of the pathogen to survive.

The close clustering (average 32 SNP difference) of the reference strain taken from Brittany, France in 2013[12] with the outbreak strains was unexpected, and considering the evolutionary rate of 9 bases per year calculated with the BEAST analysis, it is plausible that the UK introduction was of an ancestor of the French *emm*66.0 strain. It must be noted however, that conclusions from SNP variation in context with contemporary circulating isolates may be difficult and in the present study comprised only historical isolates dating from 2005, lacking non-invasive contemporary isolates.

It has been demonstrated that different *emm* types can have various levels of SNP variation with the historical isolates[13,22–24] and that the evolution of strains can also lead to highly successful new strains, such as non-capsulated *emm*89 strain described by Turner et al.[24] Therefore, for the SNP case definition a 10 SNPs difference average was chosen as cut-off to describe the outbreak strains, as described elsewhere.[22] Finally, the clonal expansion was supported from the MLST and AMR analyses; in fact, all isolates fitting the SNP outbreak case definition showed the identical MLST profile and sequence type and AMR profile. Our investigation has some limitations, including difficulties in interviewing the affected population, lack of routine reporting of risk factors and the limited number of isolates available from the general population. In particular, the interviews were especially difficult to conduct for cases identified retrospectively. Accurately establishing networks of contacts poses challenges for investigation and control to disrupt transmission. This problem, and the spread in time and location, may at least in part explain why we were unable to find evidence to support direct transmission between cases in different towns. Asymptomatic carriage by third parties was also a possible source of transmission that could not be readily investigated.

Risk factors and comorbidities are not routinely collected in national surveillance; therefore, the status of PWID with homelessness was reported only for one iGAS *emm*66.0 historical patient collected in 2015, though in two more cases of the same year, the site of infection (left groin abscess and ankle tissue) suggested PWID status.

It is important to consider that in our population SNP and BEAST analyses were limited by access to a small number of isolates overall and very few from the general population, due to the rarity of the *emm*66 strain, therefore it was not possible to compare the outbreak strain with the concurrent *emm*66.0 strains circulating in asymptomatic or non-invasive infections in the general and/or risk populations.

It is also likely that additional cases of infection with the present clone were not detected and that the identified infections represent an underestimate of the overall infection in the population with this clonal lineage of GAS.

In conclusion, the *emm*66.0 outbreak was attributed to a distinct clone that was most likely introduced in England and Wales in 2015, spreading preferentially among PWID, causing sporadic infections in the general population. No data were available, however, on symptom-free carriage of this strain.

WGS analysis enabled description of the extension of the outbreak, the redefinition of the case definition, increased understanding of the circulating strain and evaluation of its evolutionary rate, highlighting the value of this technique in the outbreak investigation.

The route of transmission has not been identified and drug contamination has been considered a very unlikely hypothesis. Plausible potential person-to-person transmission routes include respiratory/close contact and/or oral activity linked with the drug use and/or sharing injecting equipment/sexual activity.

Future cluster and outbreak analysis may benefit from the use of WGS, for example in scenarios with limited access to epidemiological data or scarce knowledge of the pathogen, as shown in this outbreak investigation where a hard-to-reach population and a less common streptococcal type were involved.

The importance of hygienic and safe injecting practices in PWID, and early seeking healthcare for wound infections remain important preventative measures in reducing the risk of invasive infections. Drug and alcohol services and others including HPTs working with the affected populations should share information to promote detection.

Declaration of Competing Interest

None.

Acknowledgment

We would like to thank all outbreak investigation teams, listed in alphabetical order, for their role in investigating the outbreak: Surrey and Sussex and South West Health Protection Teams: Rachel Cloke, Max Courtney-Pillinger, Sarah King, Rachel Kwiatkowska and Jill Rajan-Iyer; National Infection Service, Colindale: Monica Desai, Rachel Glass, Vivian Hope; Chenchal Dhami Field Epidemiology Services: Maya Gobin, Charles Beck, Daiga Jermacane (EPIET Alumna).

The authors would like to thank Aftab Jasir for her critical reading of the paper and her precious suggestions and Natalie Groves for her help with accession number submission.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jinf.2019.08.009.

References

1. Factor SH, Levine OS, Schwarts B, Harrison LH, Farley MM, et al. Invasive group A streptoccoccal disease:risk factors for adults. *Emerg Infect Dis* 2003;**9**(8):970–7.
2. Efstratiou A, Lamagni T Epidemiology of streptococcus pyogenes. *Streptococcus pyogenes: Basic biology to clinical manifestations*. Ferretti JJ, Stevens DL, Fischetti VA, editors, Oklahoma CityOK: University of Oklahoma Health Sciences Center; 2016. Available from: *https://www.ncbi.nlm.nih.gov/books/NBK343616/* 2016.
3. Bundle N, Bubba L, Coelho J, Kwiatkowska R, Cloke R, King S, et al. Ongoing outbreak of invasive and non-invasive disease due to group A streptococcus (GAS) type *emm*66 among homeless and people who inject drugs in England and Wales, January to December 2016. *Euro Surveill.* 2017;**22**(3) pii=30446.
4. Luca-Harari B, Darenberg J, Neal S, Siljander T, Strakova L, Tanna A, et al. Clinical and microbiological characteristics of severe streptococcus pyogenes disease in Europe. *J Clin Microbiol* 2009;**47**(4):1155–65.
5. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global *emm* type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect Diseas* 2009;**9**(10):611–16.

6. PHE brief note 2 November 2016 http://phenet.phe.gov.uk/Our-Organisation/Health-protection-briefings/Briefing%20notes%202016/Increase-in-invasive-Group-A-Streptococcal-infections-in-people-who-inject-drugs.pdf.
7. R Project for Statistical Computing: https://www.r-project.org/.
8. STATA software. StataCorp.. *Stata Statistical Software: Release 14*, College Station, TX: StataCorp LP; 2015. Available at: *https://www.stata.com/*.
9. yED software. Available at: https://www.yworks.com/products/yed.
10. Centres for disease control and prevention, streptococcus laboratory. Assigning *emm* Types and Subtypes Atlanta, GA, USA2007 [14 December 2016]. Available at: https://www.cdc.gov/streplab/assigning.html.
11. Chalker V, Jironkin A, Coelho J, Al-Shahib A, Platt S, Kapatai G, et al. Genome analysis following a national increase in scarlet fever in England 2014. *BMC Genom* 2017;**18**(1):224.
12. Meygret A, Vincent P, Moullec S, Nacazume J, Adnani Y, Lavenier D, et al. Genome sequence of the uncommon streptococcus pyogenes M/*emm*66 strain STAB13021, isolated from clonal clustered cases in French Brittany. *Genome Announc* 2016;**4**(4):e00689-16. doi:10.1128/genomeA.00689-16.
13. Duchêne S, Holt KE, Weill FX, Le Hello S, Hawkey J, Edwards DJ, Fourment M, Holmes EC. Genome-scale rates of evolutionary change in bacteria. *Microb Genom* 2016;**2**:e000094. doi:10.1099/mgen.0.000094.eCollection.
14. Wikler MA, Cockerill FR, Bush K, Dudley MN, Eliopoulos GM, Hardy DJ, et al. Methods for dilution Antimicrobial Susceptibility tests for bacteria that grow aerobically; approved standard- eighth edition. *Clinical and laboratory standards institute methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically approved standard M07-A8 (8), CLSI* 2009;**29**(2).
15. ARGANNOT: http://en.mediterranee-infection.com/article.php?laref=283%26titre=arg-annot.
16. Tyrrell GJ, Lovgren M, St Jean T, Hoang L, Patrick DM, Horsman G, et al. Epidemic of group a streptococcus M/emm59 causing invasive disease in Canada. *Clin Infect Dis* 2010;**51**(11):1290–7.
17. Sierra JM1, Sánchez F, Castro P, Salvadó M, de la Red G, Libois A, et al. Group A streptococcal infections in injection drug users in Barcelona, Spain: epidemiologic, clinical, and microbiologic analysis of 3 clusters of cases from 2000 to 2003. *Medicine (Baltimore)* 2006;**85**(3):139–46.
18. Mosites E, Frick A, Gounder P, Castrodale L, Li Y, Rudolph K, et al. Outbreak of invasive infections from subtype emm26.3 group A streptococcus among homeless adults-Anchorage, Alaska, 2016–2017. *Clin Infect Dis* 2017. doi:10.1093/cid/cix921.
19. Public Health England, shooting up: infections among people who inject drugs in the UK, 2016. Updated in November 2017. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/663003/Shooting_Up_2017_report.pdf.
20. Nakagawa I, Kurokawa K, Yamashita A, Nakata M, Tomiyasu Y, Okahashi N, et al. Genome sequence of an M3 strain of streptococcus pyogenes reveals a large scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res* 2003;**13**(6A):1042–55.
21. Ogura K, Watanabe S, Kirikae T, Miyoshi-Akiyama T. Complete genome sequence and comparative genomics of a streptococcus pyogenes emm3 strain M3-b isolated from a Japanese patient with streptococcal toxic shock syndrome. *J Genomics* 2017;**5**:71–4.
22. Fittipaldi N, Olsen RJ, Beres SB, Van Beneden C, Musser JM. Genomic analysis of *emm*59 group A streptococcus invasive strains, United States. *Emerg Infect Dis.* 2012;**18**(4):650–2.
23. Curtis SJ, Tanna A, Russell HH, Efstratiou A, Paul J, Cubbon M, et al. Invasive group A streptococcal infection in injecting drug users and non-drug users in a single UK city. *J Infect* 2007;**54**:422e426.
24. Turner CE, Abbott J, Lamagni T, Holden MT, David S, Jones MD, et al. Emergence of new highly successful acapsular group a streptococcus clade of genotype *emm*89 in the United Kingdom. *MBio* 2015;**6**(4):e00622.