Emergence of dominant toxigenic M1T1 Streptococcus pyogenes clone during increased scarlet fever activity in England: a population-based molecular epidemiological study

Nicola N Lysney*, Elita Jauneikaite*, Ho Kwong Li, Xiangyun Zhi, Claire E Turner, Mia Mosavie, Max Pearson, Masanori Asai, Ludmila Lobkowicz, J Yimmy Chow, Julian Parkhill, Theresa Lamagni, Victoria J Chalker, Shiranee Sriskandan

Summary
Background Since 2014, England has seen increased scarlet fever activity unprecedented in modern times. In 2016, England’s scarlet fever seasonal rise coincided with an unexpected elevation in invasive Streptococcus pyogenes infections. We describe the molecular epidemiological investigation of these events.

Methods We analysed changes in S pyogenes emm genotypes, and notifications of scarlet fever and invasive disease in 2014–16 using regional (northwest London) and national (England and Wales) data. Genomes of 135 non-invasive and 552 invasive emm1 isolates from 2009–16 were analysed and compared with 2800 global emm1 sequences. Transcript and protein expression of streptococcal pyrogenic exotoxin A (SpeA; also known as scarlet fever or erythrogenic toxin A) in sequenced, non-invasive emm1 isolates was quantified by real-time PCR and western blot analyses.

Findings Coincident with national increases in scarlet fever and invasive disease notifications, emm1 S pyogenes upper respiratory tract isolates increased significantly in northwest London in the March to May period, from five (5%) of 96 isolates in 2014, to 28 (19%) of 147 isolates in 2015 (p=0.0021 vs 2014 values), to 47 (33%) of 144 in 2016 (p=0.0080 vs 2015 values). Similarly, invasive emm1 isolates collected nationally in the same period increased from 183 (31%) of 587 in 2015 to 267 (42%) of 637 in 2016 (p=0.0001). Sequences of emm1 isolates from 2009–16 showed emergence of a new emm1 lineage (designated M1uk)—with overlap of pharyngitis, scarlet fever, and invasive M1uk strains—which could be genotypically distinguished from pandemic emm1 isolates (M1global) by 27 single-nucleotide polymorphisms. Median SpeA protein concentration in supernatant was nine-times higher among M1 UK isolates (190·2 ng/mL [IQR 168·9–200·4]; n=10) than M1global isolates (20·9 ng/mL [0·0–27·3]; n=10; p<0.0001). M1uk expanded nationally to represent 252 (84%) of all 299 emm1 genomes in 2016. Phylogenetic analysis of published datasets identified single M1uk isolates in Denmark and the USA.

Interpretation A dominant new emm1 S pyogenes lineage characterised by increased SpeA production has emerged during increased S pyogenes activity in England. The expanded reservoir of M1uk and recognised invasive potential of emm1 S pyogenes provide plausible explanation for the increased incidence of invasive disease, and rationale for global surveillance.

Funding UK Medical Research Council, UK National Institute for Health Research, Wellcome Trust, Rosetrees Trust, Stoneygate Trust.

Copyright © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction
Scarlet fever is a classic exanthem of childhood caused by the bacterium Streptococcus pyogenes (group A streptococcus) that, until the beginning of the 20th century, was associated with frequent loss of life among children.1 By the start of the 20th century, long before widespread use of antibiotics, the incidence and severity of scarlet fever had begun to fall, a phenomenon that remains largely unexplained.2 One potential (untestable) hypothesis is that the streptococcal bacteria causing the disease might have undergone a pathogenetic change that led to a reduction in the invasive and septic sequelae of scarlet fever.

Since the 1940s, scarlet fever has followed a seasonal springtime pattern—peaking between March and May while remaining less frequent throughout the rest of the year—without the major cyclical epidemics observed in the early 20th century.3 Surges in invasive infections can periodically follow a similar seasonal pattern for reasons that are incompletely understood. In 2014, England had an unexpected surge in scarlet fever infections, with over 15 000 disease notifications—a marked increase in incidence compared with previous decades.4 4 Despite having a major impact on public health resources,5 the increase in infections was not associated with any rise in the incidence of invasive disease. Even greater
Research in context

Evidence before this study
In March to May of 2016, an unexpected elevation in notifications of invasive Streptococcus pyogenes infections in England was seen, coinciding with a national increase in notifications of seasonal scarlet fever (a paediatric exanthem also caused by S pyogenes). Since 2014, scarlet fever notifications in England have reached unexpectedly high levels, peaking between March and May each year, although notifications of invasive S pyogenes infections in 2014 were within expected limits, in contrast to 2016. We aimed to test the hypothesis that the link between scarlet fever and invasive infection patterns might be strain-related and, in the process, identified the emergence of a new M1T1 lineage. We searched PubMed for clinical and laboratory studies published before March 1, 2019, using the search terms “scarlet fever” and “upsurge” or “mortality”, as well as “emm1” or “M1T1” and “streptococcus”. We also searched using the terms “SpeA” or “scarlet fever toxin” or “erythrogenic toxin” and “streptococcus” and “regulation”. We identified studies describing recent and historical trends in scarlet fever incidence, studies that described trends in strain types causing invasive infections, and studies that linked SpeA to dominant strains. We also found studies of toxin expression that reported links with phage induction, growth phase regulation, transcriptional regulators, proteolysis, and host proteins as potential regulators.

Added value of this study
Our study provides a molecular explanation for the association between increased incidence of scarlet fever and increased incidence of invasive S pyogenes infections, by identifying an emergent lineage of M1T1 S pyogenes (M1UK) that expanded rapidly to become the largest single contributor to both non-invasive and invasive infections in 2016. The findings raise the possibility that historical associations between epidemic waves of scarlet fever and invasive infections might also have been linked to strain pathogenicity, in addition to general population susceptibility. Genomic analysis confirmed that the strains that cause scarlet fever are no different to those that cause streptococcal pharyngitis and rarer invasive infections. Increases in one disease could lead to increases in all, particularly if the lineage involved is highly pathogenic. The emergent lineage was characterised by a number of genetic changes that were predictive of increased production of SpeA, and this increased production was confirmed by laboratory testing. Although this might be just one of many changes in the new lineage, increased production of SpeA is predicted to enhance bacterial fitness, as suggested by the increasing dominance of the new lineage in comparison to older M1T1 strains in England. The work highlights that group A streptococcal lineages can differ in pathogenicity.

Implications of all the available evidence
Scarlet fever notifications in England in the period 2014–18 are the highest seen since 1960, and incidence in young children exceeds that reported in other countries. It is uncertain whether the increase in scarlet fever is a result of practice change or other population or environmental factors; the new lineage described is not implicated in the initial upsurge. However, there is a need to curtail this increasing trend. Interventions targeted at the population at risk have the potential to reduce the reservoir of S pyogenes that can seed more harmful invasive infections. Research to identify the most appropriate intervention is underway and practice guidelines for streptococcal pharyngitis might need to take strain variation and wider population effects into account. Increased S pyogenes disease activity could provide a platform for strain evolution and expansion, highlighting an unforeseen consequence of modern epidemics. The genetic changes in the emergent M1T1 clone require detailed laboratory investigation to understand the wider phenotypic changes that have occurred and the molecular basis for these, including transmissibility and response to treatment. It is not known whether the new lineage will recede in due course, as other lineages have done, or if it will remain dominant in the population, and surveillance is needed. Detection of new lineage representatives outside the UK underlines a need for global surveillance and increased vigilance if strains with increased fitness and altered phenotype are detected.
responsible for *S. pyogenes* infections during the 2014–16 scarlet fever seasons regionally, then extended the study nationally, to identify bacterial determinants that might explain the observed increase in invasive disease due to *S. pyogenes* in 2016.

**Methods**

**S. pyogenes** notifications

Cases of suspected scarlet fever are notified by clinicians to Public Health England on the basis of symptoms and signs consistent with scarlet fever, with or without laboratory confirmation of *S. pyogenes* infection. Scarlet fever has been notifiable since 1899 in England. Since 2010, cases of invasive *S. pyogenes* infections have also been notifiable to Public Health England, in accordance with statutory regulations.

**S. pyogenes** isolates and **emm** genotyping

All non-invasive *S. pyogenes* isolates submitted to the Diagnostic Laboratory at Imperial College Healthcare National Health Service Trust (London, UK) between Jan 1, 2009 and Dec 31, 2013, and between March 1 and May 31 each year during 2014–16 were cultured and stored frozen in glycerol (sampling rationale is detailed in the appendix p 29). This laboratory serves northwest London, a population of roughly 2 million people, representing approximately 3% of the population of England. Clinical data were linked to isolates and anonymised in accordance with research ethics approval (number 06/Q0406/20). Isolates were **emm** genotyped (appendix p 29). Laboratories from England and Wales are requested to submit sterile site and invasive *S. pyogenes* isolates to the national reference laboratory, where **emm** genotyping is done on all submitted isolates. All isolates were cultured on Columbia blood agar (Oxoid, Basingstoke, UK) or in Todd-Hewitt broth (Oxoid) at 37°C with 5% CO₂.

**Genome sequencing**

All non-invasive **emm**1 *S. pyogenes* isolates collected from northwest London from 2009 to 2016 were subject to genome sequencing (appendix p 29), as were invasive **emm**1 isolates collected from England and Wales from March to May of 2013 and 2016. Comparative genomic analysis of invasive isolates was done with new (2013 and 2016) and existing (2014–15) genome sequence data (appendix p 12–27). DNA from *S. pyogenes* isolates was prepared, sequenced, analysed, and compared with published data from the UK (2007–12), North America, Nordic regions, and southeast Asia (appendix pp 28–29). All new genome sequence data generated in this study have been submitted to the European Nucleotide Archive (under the accession numbers listed in the appendix [pp 8, 12]).

**Quantification of speA production**

Transcript expression of **speA**, encoding streptococcal pyrogenic exotoxin A (**SpeA**; also known as scarlet fever or erythrogenic toxin A), was quantified by real-time PCR using a standard curve method. **SpeA** protein expression in 5× concentrated overnight supernatants was assessed with western blotting, and compared with a recombinant **SpeA** standard (appendix pp 29, 30). For **SpeA** protein testing, one or two isolates per year (2009–15) per lineage (M1 global vs M1c) were randomly selected by lot, after excluding any strains with mutations in the two-component regulator **covRS** (also known as **csrRS**), which is known to repress a number of virulence factors, including **SpeA**.

**Statistical analysis**

Datasets were compared with a two-tailed Mann Whitney U test for continuous data or a χ² test for categorical data, using GraphPad Prism 5.0 software. *p* values less than 0·05 were considered to indicate statistical significance. To assess trends in invasive disease incidence, Poisson regression in Stata (version 15) was done using mid-year population denominators for England from the Office for National Statistics.

**Role of the funding source**

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

**Results**

The northwest London population had a year-on-year rise in scarlet fever notifications between 2014 and 2016 that was representative of the country as a whole when compared with national notification data (figure 1A). We also compared notifications of invasive *S. pyogenes* infections in northwest London with national data (appendix p 2), which showed a marked increase in *S. pyogenes* invasive disease in 2016 in northwest London during the scarlet fever season that also mirrored the national pattern of increased notifications in 2016. **emm** genotypes of all *S. pyogenes* upper respiratory tract isolates obtained in northwest London during the scarlet fever seasons spanning 2015 and 2016 were ascertained and compared with existing data from 2014. **emm**1 upper respiratory tract strains increased in frequency year-on-year in the scarlet fever seasons, from five (5%) of 96 isolates in 2014, to 28 (19%) of 147 isolates in 2015 (p=0·0021 vs 2014 value), and 47 (33%) of 144 in 2016 (p=0·0080 vs 2015 value); in 2016, **emm**1 became the single most frequent *S. pyogenes* genotype causing upper respiratory tract infections (figure 1B). The increase in **emm**1 strains in 2016 contrasted with the previously reported associations of **emm**3 and **emm**4 strains with the initial upsurge in scarlet fever in 2014. Brief clinical details were supplied for 243 of 387 upper respiratory tract isolates overall, among which 53 (22%) mentioned scarlet fever. Of these scarlet fever-associated isolates, those genotyped...
as emm1 increased from zero (0%) of 17 in 2014, to four (24%) of 17 in 2015, and seven (37%) of 19 in 2016 (p=0.0053 vs 2014 value).

To identify any genetic basis for the expansion of emm1 S pyogenes among upper respiratory tract isolates collected in London, the genomes of all non-invasive emm1 isolates available from northwest London from 2009 to 2016 were sequenced (n=135). Single-nucleotide polymorphism (SNP)-based analysis of emm1 strains pointed to the emergence of a new emm1 lineage (designated M1uk), which could be differentiated from the contemporary emm1 population (M1global) by the presence of 27 core SNPs in regulatory and metabolic genes (figure 2, appendix p 7). The earliest member of the M1uk lineage was identified in 2010, and five intermediate isolates (with 13 or 23 of the unique SNPs) were detected between 2009 and 2012 (appendix p 3). Similarly to M1global, M1uk isolates were identified among all age groups, but included more cases of scarlet fever and evidence of recent transmissions than M1global (appendix p 3). From 2015 onwards, around two-thirds of non-invasive emm1 isolates from northwest London were within the new M1 UK lineage (22 [71%] of 31 isolates in 2015, and 30 [65%] of 46 in 2016). Recombination and pan-genome analyses provided no evidence for gain or loss of transferable elements when comparing M1 global and M1uk strains. Lineage-specific acquisition of antimicrobial-resistance genes was not detected; evidence of the mefA and msrD macrolide-resistance locus was found in just one M1uk isolate, while eight M1global isolates possessed antimicrobial-resistance genes (one isolate with the mefA and msrD locus, one with the tetracycline-resistance gene tetM, and six with the aminoglycoside-resistance gene aph3).

Scarlet fever is a toxin-mediated syndrome, historically associated with the expression of the phage-encoded...
erythrogenic toxin SpeA,13 the gene for which is possessed by all emm1 isolates studied in this investigation. Among the 27 M1UK lineage-defining SNPs, three non-synonymous mutations were identified in the gene rofA encoding the standalone transcriptional regulator RoFA,14 which, alongside a homologue, nra, has been implicated as a repressor of SpeA production in some, but not all, streptococcal genotypes.15,16 Real-time PCR analysis of all sequenced non-invasive emm1 isolates from northwest London (n=135) indicated that the emergent lineage (M1UK) was associated with significantly increased transcription of speA compared with other contemporary emm1 isolates (M1global), suggesting that differential SpeA production was a feature of the M1UK lineage (figure 3A).

To ascertain whether this difference in transcription translated into a difference in SpeA protein expression, we compared ten M1UK (including one intermediate strain harbouring only 13 of the 27 lineage-defining SNPs) and ten M1global isolates and corroborated our finding that M1UK isolates (appendix p 12–27) in England and Wales from March to May each year during 2013–16. Focusing first on invasive disease isolates (appendix p 1), the seasonal increase in notifications matched that observed nationally, emm1 genotyping of all invasive disease isolates referred to the national laboratory showed significant absolute and relative year-on-year increases in emm1, from 183 (31%) of 587 isolates between March and May 2015, to 267 (42%) of 637 in the same period in 2016 (p<0·0001 vs 2015 values), peaking March 2016 (figure 4A).

To ascertain whether invasive S pyogenes infections might be affected by the emergent M1UK lineage, we compared the genome sequences of 552 invasive emm1 isolates (appendix p 12–27) in England and Wales from March to May each year during 2013–16. Focusing first on London, where sequence data from non-invasive isolates were available for comparison, SNP-based phylogenetic analysis showed intermingling of the 31 sterile-site invasive and 135 non-invasive isolates; 16 (84%) of 19 invasive
strains obtained in 2015 and 2016 lay within the emergent M1UK lineage, compared with just five (42%) of 12 obtained between 2013 and 2014 (appendix p 4). Four (13%) of 31 invasive isolates were identical to, or no more than two SNPs different from, non-invasive isolates in the community, consistent with recent transmission.

We then analysed the genome sequences of all 552 emm1 sterile site isolates collected between March and May each year from 2013 to 2016, obtained nationally from across different geographical locations in England and Wales. 425 (77%) of 552 invasive emm1 strains were within the new lineage (figure 4B), which was present in the UK invasive isolate population as early as 2013. Like M1global strains, M1UK strains were phylogenetically distinct from the historical UK emm1 scarlet fever speA-positive isolate NCTC8198 and emm1 speC-positive SF370 (appendix p 5). Longitudinal analysis of all available 1240 UK emm1 sequences (appendix p 28)6–8 obtained from invasive and non-invasive disease cases showed a yearly increase in M1UK such that, by 2016, M1UK strains represented more than 80% of all available emm1 isolates in the UK, outnumbering M1global strains (figure 5A).

Phylogenetic comparison of UK emm1 sequences with available international sequences from North America, Nordic regions, UK, and southeast Asia (appendix p 28)9–11 confirmed that M1UK strains were distinct from the globally disseminated pandemic emm1 strains (figure 5B). A small number of intermediate isolates (possessing at least 13 of the 27 lineage-defining SNPs, including three SNPs in ropA) were identified in the UK and in Nordic countries, including Denmark (n=16), Finland (n=2), and Sweden (n=4, figure 5b, appendix p 6).10 M1UK isolates that possessed all 27 SNPs were, however, unique to the UK, except single sequences isolated in the USA in 20159 and in Denmark in 2012,10 underlining the potential of the new lineage to disseminate internationally (figure 5B, appendix p 6).

Discussion
The modern era of resurgent invasive S pyogenes infections was heralded by reports of virulent strains producing the scarlet fever toxin SpeA17 and subsequently dominated by the emm1 lineage.10 In this study, we showed that the originally polyclonal upsurge of scarlet fever in England has more recently been characterised by the emergence of a new emm1 S pyogenes lineage that produces significantly higher levels of SpeA than other contemporary emm1 strains. The new M1UK lineage showed an apparent fitness advantage within the population, manifest during the scarlet fever seasons of 2015 and 2016. Phylogenetic analysis showed the emergent lineage
to be the dominant cause of invasive *S pyogenes* infections in England in 2016, and indicated that isolates from symptomatic throat infections and scarlet fever represent the major reservoir for invasive infections. The data support the hypothesis that transmission of virulent *emm*1 strains with enhanced ability to cause scarlet fever could underlie the contemporaneous rise in invasive *S pyogenes* disease.

An unprecedented year-on-year increase in scarlet fever notifications, the underlying basis for which remains unclear, has been seen in England since 2014. Among children aged 1–4 years, the incidence of scarlet fever in 2018 reached 1488 per 100000. The increase in scarlet fever activity has followed secular changes in factors such as household structure (including use of childcare) and health-care use, as well as health-care policies, although causal links have not been established. One unforeseen consequence of medical practice change was that the capacity to investigate such an upsurge was undermined by a reduction in diagnostic testing at the national level. In northwest London, however, where strains are collected for epidemiological analysis, a significant increase in genotype *emm*4 pharyngitis strains was observed during the 2014 upsurge in scarlet fever, while *emm*3 was the main genotype associated with physician-reported scarlet fever. *emm*1 was infrequent in 2014, contrasting with the increase we observed between 2015 and 2016, when genotype *emm*1 *S pyogenes* became the dominant cause of upper respiratory tract infections regionally, and invasive disease notifications nationally.

Our genome sequence analysis revealed the emergence of a new *emm*1 lineage, separated from all other *emm*1 strains by 27 unique core SNPs, including three within a gene encoding a potential SpeA regulator, RofA. SpeA is usually the only phage-encoded superantigen in contemporary *emm*1 *S pyogenes*, and has been implicated in the re-emergence of severe invasive *S pyogenes* infections in the 1980s. Although the roles of any specific SNPs were not investigated in the current study, we hypothesise that increased SpeA production by M1UK strains might be an important contributor to the apparent fitness of the new lineage within the nasopharynx. As a phage-encoded superantigen, SpeA is hypothesised to trigger scarlet fever in susceptible children, and has been shown to permit nasopharyngeal infection in humanised models of murine streptococcal infection, plus potential induction of immunity when administered as a toxoid. SpeA can trigger B cell death and abrogate immunoglobulin secretion by the human tonsil, and

---

**Figure 5:** Longitudinal (A) and geographical (B) comparison of M1 lineage with pandemic *emm*1 strains of *Streptococcus pyogenes*. (A) Proportions of M1uk and M1 global isolates among total sequenced invasive and non-invasive *emm*1 *S pyogenes* isolates (n=1240) annually in the UK between 2007 and 2016. (B) M1 lineage in a global context. Maximum likelihood phylogenetic tree constructed from core SNPs (excluding prophage regions) comparing all sequenced UK *emm*1 isolates with the global *emm*1 populations from North America, Nordic countries, and Asia (n=2800 isolates). Shading in grey indicates the emergent lineage M1UK; orange arc indicates intermediate isolates that lie outside M1UK but possess 13 or more of the 27 SNPs present in M1UK, including three SNPs in rofA. UK and international *emm*1 isolates arise throughout the tree, but isolates within the M1UK lineage are exclusively from the UK, except two single isolates from Denmark and the USA (arrows). The scale bar indicates the number of nucleotide substitutions per site. See appendix (p 6) for the unrooted tree. SNPs=single-nucleotide polymorphisms.
produce phage emm scarlet fever 20th century. Although the oldest classically associated with scarlet fever in the early 1) were the first to be serotype M1; genotype monitor. Whether population immunity to SpeA will lead to an eventual decline in the lineage will be of interest to monitor. Thus, production of SpeA might augment the ability of S pyogenes to cause scarlet fever and paediatric pharyngitis. Whether the M1 UK lineage will be suited to other environments is unknown; management of streptococcal sore throat differs greatly between countries,30 as do other important factors such as climate. We previously reported that the M1 global lineage were imported, we speculate that a generalised increase in S pyogenes activity in the wider population— which coincided with England’s scarlet fever upsurges— might have provided the conditions required for adaptation and expansion of emm1 S pyogenes. Whether the M1 UK lineage will be suited to other environments is unknown; management of streptococcal sore throat differs greatly between countries,30 as do other important factors such as climate. We previously reported the emergence of a new emm89 lineage that had lost the capsule locus but gained an active NADase–streptolysin O locus, in addition to four other major recombination events.9 This emm89 lineage has now been identified across several continents.12 The identification of two members of the new M1 UK lineage among isolates outside the UK underlines the potential of such lineages to spread globally. Compared with other genotypes, emm1 S pyogenes has a recognised, heightened association...
with invasive infections. The expansion of such a lineage within the community reservoir of *S. pyogenes* might be sufficient to explain England’s recent increase in invasive infection. Further research to assess the likely effects of *M1* on infection transmissibility, treatment response, disease burden, and severity is required, coupled with consideration of public health interventions to limit transmission where appropriate. Wider national and global surveillance will provide clearer understanding of the lineage’s geographical reach and longer-term fitness, and permit enhanced public health readiness where necessary.

### Contributors

SS, EJ, NNL, CET, VC, and TL contributed to the conception of this project. HKL, XZ, MM, MP, MA, LL, EJ, and JP were responsible for collection of laboratory and genomic data. EJ and NNL interpreted and analysed genome data. EJ and CET undertook bioinformatics analysis of whole genome sequence data. SS, NNL, and EJ prepared the manuscript; NNL and EJ contributed equally. All authors contributed to the interpretation of results and critical review of the manuscript.

### Declaration of interests

JP is a consultant to Next Gen Diagnostics LLC. All other authors declare no competing interests.

### Acknowledgments

This work was supported by grants from the UK Medical Research Council (grant number MR/P022669/1); the UK National Institute for Health Research (NIHR) Health Protection Unit in Healthcare Associated Infections and Antimicrobial Resistance (grant number HPRU-2012-10040); an NIHR Biomedical Research Centre grant (SpyVAC); and the Conors Kerin Foundation. NNL is a Sir Henry Wellcome Postdoctoral Research Fellow funded by the Wellcome Trust (grant number 103197/Z/11/Z). EJ is a Rosetrees/Stonegate 2017 Imperial College Research Fellow (number M683). SS acknowledges the NIHR Biomedical Research Centre grant awarded to Imperial College and the Biomedical Research Centre Infection Bioresource. The authors are grateful to colleagues in the Imperial College National Health Service Healthcare Trust Diagnostic Laboratory and Respiratory and Vaccine-Preventable Bacteria Reference Unit, Public Health England, and Juliana Coelho and Rebecca Guy (National Infection Service, Public Health England).

### References


