WGS to determine the extent of *Clostridioides difficile* transmission in a high incidence setting in North Wales in 2015

David W. Eyre¹,²*, Robert Shaw²†, Helen Adams³, Tracey Cooper⁴, Derrick W. Crook², Rhonda-Marie Griffin³, Phil Mannion⁵, Mari Morgan⁶, Trefor Morris⁷, Michael Perry⁷, Sophie Jones⁷, Tim E. A. Peto², Jonathan Sutton³, A. Sarah Walker², Dafydd Williams³ and Noel Craine⁸

¹Big Data Institute, University of Oxford, England, UK; ²Nuffield Department of Medicine, University of Oxford, England, UK; ³Betsi Cadwaladr University Health Board, Ysbyty Gwynedd, Bangor, Wales, UK; ⁴Betsi Cadwaladr University Health Board, Ysbyty Glan Clwyd, Bangor, Wales, UK; ⁵Public Health Wales, Microbiology, Ysbyty Glan Clwyd, Rhyl, Wales, UK; ⁶Public Health Wales, Health Protection, Capital Quarter, Cardiff, Wales, UK; ⁷Public Health Wales, Microbiology, Heath Hospital, Cardiff, Wales, UK; ⁸Public Health Wales, Health Protection, Ysbyty Gwynedd, Bangor, Wales, UK

*Corresponding author. Tel: +44 1865 220855; E-mail: david.eyre@ndm.ox.ac.uk orcid.org/0000-0001-5095-6367
†These authors contributed equally.

Received 23 July 2018; returned 7 October 2018; revised 6 November 2018; accepted 16 November 2018

**Objectives:** Rates of *Clostridioides* (*Clostridium*) *difficile* infection (CDI) are higher in North Wales than elsewhere in the UK. We used WGS to investigate if this is due to increased healthcare-associated transmission from other cases.

**Methods:** Healthcare and community *C. difficile* isolates from patients across North Wales (February–July 2015) from glutamate dehydrogenase (GDH)-positive faecal samples underwent WGS. Data from patient records, hospital management systems and national antimicrobial use surveillance were used.

**Results:** Of the 499 GDH-positive samples, 338 (68%) were sequenced and 299 distinct infections/colonizations were identified, 229/299 (77%) with toxin genes. Only 39/229 (17%) toxigenic isolates were related within 2 SNPs to 1 infection/colonization from a previously sampled patient, i.e. demonstrated evidence of possible transmission. Independent predictors of possible transmission included healthcare exposure in the last 12 weeks (*P* = 0.002, with rates varying by hospital), infection with MLST types ST-1 (ribotype 027) and ST-11 (predominantly ribotype 078) compared with all other toxigenic STs (*P* < 0.001), and cephalosporin exposure in the potential transmission recipient (*P* = 0.02). Adjusting for all these factors, there was no additional effect of ward workload (*P* = 0.54) or failure to meet cleaning targets (*P* = 0.25). Use of antimicrobials is higher in North Wales compared with England and the rest of Wales.

**Conclusions:** Levels of transmission detected by WGS were comparable to previously described rates in endemic settings; other explanations, such as variations in antimicrobial use, are required to explain the high levels of CDI. Cephalosporins are a risk factor for infection with *C. difficile* from another infected or colonized case.

**Introduction**

The use of WGS in endemic settings has revealed that only the minority of hospital and community *Clostridioides* (*Clostridium*) *difficile* infections (CDIs) are acquired from other symptomatic cases.¹,² However, how acquisition from cases varies with increased *C. difficile* incidence is not known. Despite declines in CDI incidence over the last 15 years,³ North Wales has among the highest CDI incidence in the UK; in 2015–16 CDI incidence was 51.1 per 100 000 population, compared with a Wales-wide rate³ of 40.1, 25.8 in England⁴ and 31.2 in Scotland⁵ (calculated using total reported cases³–⁵ and mid-2015 population estimates⁶). Surveillance methodologies in England⁴ and Wales⁵ are broadly similar. Reporting in Scotland⁵ differs by including patients ≥15 years old, compared with ≥2 years in England and Wales.

To investigate the relatively high incidence of CDI in North Wales, a prospective WGS study was initiated to test the hypothesis that this was due to increased within-hospital *C. difficile* transmission. We also investigated whether risk factors for transmission could be found, in order to identify potential infection control and other preventative interventions.
Methods

Setting

Wrexham Maelor Hospital, Glan Clwyd Hospital and Ysbyty Gwynedd are three district general hospitals providing secondary-level care to the entire region of North Wales. These hospitals serve a population of 694,473 (mid-year 2015 estimate), living in a mix of urban and remote rural settings. All hospital and community samples submitted for C. difficile testing from these hospitals, smaller community hospitals in the same region and GP surgeries are processed by a single laboratory at Glan Clwyd Hospital. These hospitals are randomly identified as hospitals A, B and C to anonymize these hospitals, smaller community hospitals in the same region and GP hospitals are randomly identified as hospitals A, B and C to anonymize these hospitals, smaller community hospitals in the same region and GP surgeries are processed by a single laboratory at Glan Clwyd Hospital. These hospitals are randomly identified as hospitals A, B and C to anonymize samples.

Microbiology

Faecal samples submitted for C. difficile testing underwent glutamate dehydrogenase (GDH) testing using C. DIFF Chek-60 (TechLab, Blacksburg, VA, USA). Positive samples underwent C. DIFF QUK CHECK COMPLETE (TechLab) to confirm the GDH result and detect the presence of C. difficile toxins A and B by enzyme immunoassay. Samples were saved, selectively cultured for C. difficile and identified as described previously and isolates obtained under WGS. GDH-positive patients were considered infected or colonized, and those who were faecal toxin-positive were considered to be infected (i.e. have CDI). Cases were denoted healthcare facility-associated, community-associated or indeterminate using standard surveillance definitions. Cases were assigned to a given hospital based on inpatient exposure in the last 12 weeks, excluding the 48 h immediately prior to diagnosis.

Sequencing

DNA was extracted after subculture of a single colony and sequenced using Illumina HiSeq 2500. Sequence data were processed as previously, mapping sequenced reads to the C. difficile 630 reference genome. Sequences were compared using SNPs, obtaining differences between sequences from maximum-likelihood phylogenies, corrected for recombination using CipalFrameML. Sequence reads were also assembled de novo with Velvet and VelvetOptimiser. Toxin genes, tcdA and tcdB, were identified from de novo assemblies using BLAST searches, on the basis of >1 kb of sequence identity to each gene. MLST types were predicted from de novo assemblies. Sequence data have been deposited under NCBI Sequence Read Archive BioProject PRJNA412541.

Genomic analysis

Based on C. difficile evolutionary rates and within-host diversity, >95% of transmission pairs sampled ≤123 days apart are expected to have ≤2 SNPs between them and cases up to 124–364 days apart ≤3 SNPs, but with 3 SNPs uncommon. Therefore, during the 5.5 months of the study ≤2 SNPs were expected between the large majority of transmitted isolates. Where patients had multiple samples, subsequent isolates >10 SNPs different to previous isolates from the same patient were considered to represent a new acquisition of a distinct C. difficile strain. This higher SNP threshold almost completely excludes isolates being from the same infection within the study. We used a previously described correction factor to adjust for sequencing only a subset of GDH-positive samples, assuming sequenced and non-sequenced cases transmit onwards at the same rate and cases are missing at random.

Risk factor analysis

Data from paper and electronic patient records were extracted into a Public Health Wales data warehouse. Data were available on admissions and ward movements for infected/colonized patients from the three district general hospitals and from smaller community hospitals and nursing homes. Additional data were available on prescribing and ward workload (ward admissions per day) and from mandatory audits of cleaning compliance within the hospital setting.

Multivariate logistic regression was used to identify independent predictors of a case being genetically related to >1 previous cases within ≤2 SNPs, selecting a final model from factors shown in Table 1 using backwards elimination with an exit P value of >0.1. Multiple fractional polynomials were used to allow for non-linear effects of continuous factors. Following initial model selection, each excluded variable was added back to the model one at a time and retained if its Wald P value was <0.1. Interactions between main effects in the final model were retained where the interaction P was <0.01. All analyses were performed using Stata 14.1 (Stata Corp, College Station, TX, USA).

Ethics

Ethics approval was not required as the work formed part of the Betsi Cadwaladr University Health Board’s response to C. difficile infection. Sequencing was carried out on C. difficile isolates following routine isolation.

Results

Between 1 February 2015 and 16 July 2015, 499 C. difficile GDH-positive samples were obtained from 417 patients. One hundred and eighty-two (36%) samples from 159 patients were faecal toxin positive and considered to represent infections. One patient had evidence of a genetically distinct second infection. Of these 160 CDIs, 33 (21%) were community-associated, i.e. had no healthcare facility exposure for >12 weeks, representing a rate of 4.8 per 10 000 population per year. One hundred and eighteen (74%) were healthcare-facility associated (healthcare exposure within 4 weeks) and nine (6%) indeterminate (healthcare exposure 4–12 weeks ago), together representing a rate of 5.7 per 10 000 bed-days. Monthly CDI incidence, with historic rates, is shown in Figure 1.

Of the 499 GDH-positive samples, 338 (68%) underwent WGS [144/182 (79%) faecal toxin-positive samples and 194/317 (61%) faecal toxin-negative samples]. Rates of GDH-positive sample retrieval were similar by hospital: 95/136 (70%), 55/81 (68%) and 92/134 (69%) at hospitals A, B and C, respectively; and 5/6 (83%) for patients exposed to both hospital A and C. Six (86%) of the seven samples from patients with only community hospital exposure were retrieved and 85/135 (63%) of samples from patients without recent hospital exposure.

Considering all GDH-positive samples, irrespective of faecal toxin status, the 338 sequenced samples contained 299 distinct infections/colonizations in 290 patients. Of these, 229/299 (77%) had detectable toxin genes on WGS and, within these potentially toxigenic isolates, 114/229 (50%) were from consistently faecal toxin-positive patients, 103/229 (45%) from consistently faecal toxin-negative patients and 12/229 (5%) from patients with both...
faecal toxin-positive and -negative results from different samples. Of the 70 distinct colonizations without detectable toxin genes on WGS, 65 (93%) were consistently faecal toxin-negative, 4 (6%) were faecal toxin-positive and 1 (1%) had both faecal toxin-positive and -negative results on different samples.

Genetic and epidemiological links between samples

Of the 299 sequenced distinct infections/colonizations, 43 (14%) were within ≤2 SNPs of previously sampled patient, i.e. had evidence of possible transmission (Figure 2). Thirty-nine (91%) of these 43 genetically-linked cases were toxigenic (i.e. had toxin genes) such that 39/229 (17%) distinct toxigenic infections/colonizations were within ≤2 SNPs of previously sampled patient. Figure 3 shows the relationship between donor and recipient faecal toxin status. Faecal toxin-positive cases were not more likely to have a faecal toxin-positive donor; instead, faecal toxin-negative recipients had predominantly positive donors, and some faecal toxin-positive recipients had faecal toxin-negative donors ($P = 0.006$ versus no relationship between donor and recipient toxin status). Of the 43 potentially transmitted infections, 26 (60%) had a single possible source, 9 (21%) had two possible sources and 4 (9%), 3 (7%) and 1 (2%) had three, four and five possible sources respectively. The median (IQR) range time from the most recently sampled case within ≤2 SNPs of the potential recipient was 21 (7–47) days.

Healthcare exposure in the 12 weeks prior to diagnosis was an important predictor of genetic linkage to a previous case; 40/217 (18%) patients with healthcare exposure were genetically linked to a previous case, compared with 3/82 (4%) without ($P = 0.001$, Figure 4a). Rates of genetic linkage to previous cases varied at the three hospitals: 9/80 (11%), 11/51 (22%) and 20/75 (27%) at hospitals A, B and C, respectively ($P = 0.04$, Figure 4a). Transfers between hospitals were uncommon; 5 patients were exposed to both hospitals A and C, and 6 patients only to smaller community hospitals; none of these 11 patients were genetically linked to a previous case. Genetic linkage did not correspond to the overall rates of

<table>
<thead>
<tr>
<th>Table 1. Risk factors for genetic linkage (≤2 SNPs) with a previous case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetically unlinked (N = 256)</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>n or median</td>
</tr>
<tr>
<td>Any hospital exposure in last 12 weeks</td>
</tr>
<tr>
<td>none</td>
</tr>
<tr>
<td>in hospital A</td>
</tr>
<tr>
<td>in hospital B</td>
</tr>
<tr>
<td>in hospital C</td>
</tr>
<tr>
<td>in both hospitals A and C</td>
</tr>
<tr>
<td>in community hospital only</td>
</tr>
</tbody>
</table>

MLST type

| | | | |
| other | 168 66% | 24 56% | 1.00 baseline | <0.001 |
| ST-1 | 9 4% | 9 21% | 7.00 2.53–19.38 | 7.61 2.50–23.16 |
| ST-11 | 13 5% | 6 14% | 3.23 1.12–9.30 | 2.27 0.67–7.68 |
| non-toxigenic | 66 26% | 4 9% | 0.42 0.14–1.27 | 0.36 0.11–1.17 |
| Sex, female | 166 65% | 28 65% | 0.91 0.46–1.80 | 1.03 1.00–1.05 |
| Age, years | 79 69–86 | 82 71–88 | 1.02 1.00–1.05 | 0.06 |
| Recipient faecal toxin positive | 101 39% | 28 65% | 2.86 1.46–5.63 | 0.002 |
| Inpatient days in last 90 days | 12 3.5–25 | 17.5 8–41 | 1.02 1.00–1.04 | 0.03 |
| Any antibiotic | 142 55% | 28 65% | 1.36 0.69–2.67 | 0.37 |
| Fluoroquinolone | 21 8% | 5 12% | 1.47 0.52–4.14 | 0.46 |
| Cephalosporin, 2nd/3rd generation | 6 2% | 5 12% | 5.48 1.59–18.85 | 6.03 1.42–25.50 |
| β-Lactam/β-lactamase inhibitor | 61 24% | 15 35% | 1.54 0.78–3.06 | 0.22 |
| Meropenem | 10 4% | 4 9% | 2.52 0.75–8.44 | 0.13 |
| Pneumopump inhibitor | 35 14% | 6 14% | 1.02 0.40–2.60 | 0.96 |
| Laxative | 18 7% | 3 7% | 0.99 0.28–3.52 | 0.99 |
| Cleaning audit, per day below target | 10 2–24.5 | 16.5 7–36 | 1.01 1.00–1.03 | 0.06 |
| Admissions, per admission exposed to | 56 21.5–110 | 85 37–141 | 1.00 1.00–1.01 | 0.07 |

Antibiotic and proton pump exposures are ever receiving the relevant agent in the 90 days prior to diagnosis, and laxative exposure in the 30 days prior to diagnosis. Cleaning audit exposure is the total number of days in the 90 days prior to diagnosis spent on a ward that had failed to meet the audit standard at the last available audit. Ward workload was judged by the total number of other patient admissions that occurred during all in-patient days in the 90 days prior to diagnosis.

$a$These hospital exposure categories had no genetic links and so an OR cannot be calculated.

The 70 distinct colonizations without detectable toxin genes on WGS, 65 (93%) were consistently faecal toxin-negative, 4 (6%) were faecal toxin-positive and 1 (1%) had both faecal toxin-positive and -negative results on different samples.

**Genetic and epidemiological links between samples**

Of the 299 sequenced distinct infections/colonizations, 43 (14%) were within ≤2 SNPs of ≥1 infections/colonizations from a previously sampled patient, i.e. had evidence of possible transmission (Figure 2). Thirty-nine (91%) of these 43 genetically-linked cases were toxigenic (i.e. had toxin genes) such that 39/229 (17%) distinct toxigenic infections/colonizations were within ≤2 SNPs of ≥1 infections/colonizations. Figure 3 shows the relationship between donor and recipient faecal toxin status. Faecal toxin-positive cases were not more likely to have a faecal toxin-positive donor; instead, faecal toxin-negative recipients had predominantly positive donors, and some faecal toxin-positive recipients had faecal toxin-negative donors ($P = 0.006$ versus no relationship between donor and recipient toxin status). Of the 43 potentially transmitted infections/colonizations, 26 (60%) had a single possible source, 9 (21%) had two possible sources and 4 (9%), 3 (7%) and 1 (2%) had three, four and five possible sources respectively. The median (IQR) range time from the most recently sampled case within ≤2 SNPs of the potential recipient was 21 (7–47) days.

Healthcare exposure in the 12 weeks prior to diagnosis was an important predictor of genetic linkage to a previous case; 40/217 (18%) patients with healthcare exposure were genetically linked to a previous case, compared with 3/82 (4%) without ($P = 0.001$, Figure 4a). Rates of genetic linkage to previous cases varied at the three hospitals: 9/80 (11%), 11/51 (22%) and 20/75 (27%) at hospitals A, B and C, respectively ($P = 0.04$, Figure 4a). Transfers between hospitals were uncommon; 5 patients were exposed to both hospitals A and C, and 6 patients only to smaller community hospitals; none of these 11 patients were genetically linked to a previous case. Genetic linkage did not correspond to the overall rates of
Figure 1. *C. difficile* incidence in North Wales 2013–15. Public Health Wales surveillance data are for faecal toxin-positive CDI cases.

Figure 2. Proportion of cases genetically linked within ≤2 SNPs, classified by surveillance definitions and epidemiological relationships between linked cases. Cases sharing the same ward or hospital did so with their potential donor between the dates of their diagnoses, or prior to the diagnosis of either case. For cases sharing the same ward within 28 days, the potential recipient spent time on the same hospital ward after the discharge of an already diagnosed donor.
healthcare-associated/indeterminate GDH-positive C. difficile colonization/infection at hospitals A, B and C, which were 16.8, 11.3 and 12.2 per 10,000 bed-days respectively, or to faecal toxin-positive CDI, occurring at 7.0, 5.2 and 5.2 per 10,000 bed-days, respectively.

Of the 43 genetically linked cases, 11 (26%) shared time and space on the same hospital ward with their potential donor between the dates of their diagnoses, 9 in a district general hospital and 2 in a community hospital (Figure 2). A further 2/43 (5%) patients shared time and space on the same ward location at different times within the 28 days prior to diagnosis (5 in a district general hospital, 2 in a community hospital and 1 in a nursing home). Finally, 2/43 (5%) patients without any other link shared time in the same district general hospital between the dates of their diagnoses, but not specific wards. Thus 20/43 (47%) potential recipients had no recent or concurrent shared healthcare exposure with any previous case within 2 SNPs even at the broadest level of the hospital, and accounting for smaller community hospitals and nursing homes.

The most commonly occurring toxigenic STs were: ST-6 (30/229 toxigenic infections/colonizations, 13%); ST-2 (27, 12%); ST-8 (21, 9%); and ST-44 (18, 8%), all from C. difficile clade 1;1; and ST-11 (19, 8%, equivalent most commonly to ribotype 078) and ST-1 (18, 8%, ribotype 027). Rates of genetic linkage were higher in ST-1 and ST-11 than the combined group of all other toxigenic STs (Figure 4b, P < 0.001). Rates of genetic linkage were lower for nontoxigenic C. difficile despite all tested patients having diarrhoea.

Similar percentages of sequenced infections/colonizations after the first 3 months of the study were within <2 SNPs of a previous case (20/123 (16%) versus 23/176 (13%) before, P = 0.27) even though cases earlier in the study may have been less likely to have had their source sampled. We applied a previously published correction11 to adjust for having only sequenced 68% of C. difficile-positive samples. This provided a corrected estimate for the percentage of cases after the first 3 months of the study that were genetically linked to a prior case of 24% (i.e. 20/123×1/0.68). Restricting only to potentially toxigenic cases, this figure was 27% (16/67×1/0.68).

**Risk factors for transmission**

Independent risk factors for genetic linkage within <2 SNPs to a previous case (Table 1) included healthcare exposure in the last 12 weeks, in hospital A [OR 3.15 (95% CI 0.77–12.9)], in hospital B [5.63 (1.40–22.7)] and in hospital C [10.1 (2.75–37.4)], compared with no healthcare exposure (P = 0.002). C. difficile genotype was also associated with genetic linkage (P < 0.001); compared with all other toxigenic STs, ST-1 cases were independently more likely to be linked to a previous ST-1 case [OR 7.61 (95% CI 2.5–23.2)] and there was some evidence for similar associations for ST-11 [2.27 (0.67–7.68)]. Older patients were somewhat more likely to be genetically linked to a previous case (P = 0.06). Second/third-generation cephalosporin exposure in the last 90 days in the potential transmission recipient increased the risk of genetic linkage [OR 6.03 (95% CI 1.42–25.5, P = 0.02)]; however, only 5/43 (12%) of cases and 6/256 (2%) of controls were exposed. Adjusting for all these factors, within the limits of the power of the study, we found no evidence for any additional effects on transmission of ward workload (P = 0.54), or failure to meet cleaning audit targets (P = 0.25).

**Population risk factors for CDI**

We considered explanations other than increased transmission for rates of CDI in North Wales. The majority of antibiotics in the UK are prescribed in primary care by GPs. Rates of community antibiotic use (in the second quarter of 2015) were higher in North Wales (296.7 items per 1000 STAR-PUs) and Wales overall (296.9 per

---

**Figure 3.** Relationship between potential transmission donor and recipient faecal toxin status. A mixed toxin status patient had ≥1 faecal toxin-positive and ≥1 faecal toxin-negative samples. Overall P value = 0.006.
1000 STAR-PUs), compared with England (243 per 1000 STAR-PUs) (Figure 5). Comparing acute hospital total antibiotic use in DDDs per 1000 bed-days in 2015, for the 17 acute hospitals in Wales, Ysbyty Gwynedd had the second highest rate, Ysbyty Glan Clwyd the sixth highest and Wrexham Maelor the twelfth.

Similarly, age is another CDI risk factor. The population in North Wales is older than Wales as a whole; 22.6% of the population are 65 years old, compared with 20.4% in Wales and 17.9% in England (mid-2016 data).

**Discussion**

Despite high CDI incidence in North Wales, based on WGS results, only 39/229 (17%) of toxigenic infections/colonizations could have
been plausibly acquired from another case. Adjusting for only sequencing 68% of isolates, this proportion was still only 27%. This is higher than in a study of six English hospitals, where rates of genetic linkage to previous cases were between 7% and 24% by hospital and 20% overall. However, these differences are insufficient to explain CDI incidence being nearly double in North Wales compared with England. Therefore, higher incidence is likely to be driven predominantly by factors other than lapses in infection control.

Antimicrobial exposure is an important CDI risk factor. Rates of antibiotic use in primary care are higher in Wales than in England, but similar in North Wales to Wales overall, potentially explaining some of the differences between North Wales and England, but not between North Wales and elsewhere in Wales. Additionally, two of the three hospitals in North Wales are among the highest users of antibiotics of all the acute hospitals in Wales.

Similarly, increasing age is another important risk factor for CDI and the population in North Wales is older than in Wales as a whole and in England. Other factors may also be important; the area of the country served by the three hospitals contains extensive areas of livestock farming. Disease-causing C. difficile strains have been isolated from livestock, with overlap seen between isolates from CDI cases, healthy humans and livestock on WGS. However, a large-scale environmental survey 20 years ago in South Wales identified relatively little C. difficile in livestock. Asymptomatic patients are another potential source of C. difficile infection; however, it is not known if rates of C. difficile colonization differ across the UK.

Recent healthcare exposure was an important risk factor for potential acquisition from a previous case; 40/43 (93%) genetically related cases were in hospital in the 12 weeks prior to their diagnosis. The median (IQR) time between genetically related cases was

Figure 5. Antibiotic prescribing in primary care in England and Wales; items prescribed per 1000 STAR-PUs. Data are presented for July–September 2015. Areas shaded are Welsh Unitary Authorities and English Clinical Commissioning Groups. Source of data is reference 18.
21 (7–47) days. However, shared space and time on the same hospital ward could only explain a minority of genetically related cases, and nearly half of such cases had no healthcare contact, including allowing for shared time in hospital resulting in overlap outside of wards, e.g. diagnostic areas. Additionally, although our study is only moderately powered, we found no signal that failure to meet cleaning audit targets or high levels of patient turnover were associated with more transmission. However, the proportion of cases linked to a previous case varied between 11% and 27% at the three main hospitals, suggesting potential for reductions in overall incidence. Supporting the previously described role in transmission of GDH-positive patients without detectable faecal toxin, 76 77 78 79 (18%) toxigenic C. difficile acquisitions could only be linked to consistently toxin-negative sources. Therefore, all patients with toxigenic C. difficile should be a focus of infection control efforts, not just those with detectable faecal toxin. ST-1 (ribotype 027) and ST-11 (ribotype 078) were associated with higher rates of genetic linkage, replicating previous findings from England 27 and for ST-1 from Canada. 28 The underlying reasons for this may be multifactorial, including more severe disease 19 leading to greater environmental contamination, enhanced environmental persistence and also a greater likelihood of clinically detectable disease in transmission recipients.

Antimicrobials are risk factors for CDI. 22 We investigated more specifically the effect of recent antimicrobial exposure on acquisition of C. difficile from another case. Second/third-generation cephalosporin exposure, but not antibiotics in general or any other specific antibiotic class, increased the risk of being a transmission recipient. The effect of cephalosporins may reflect intrinsic resistance in C. difficile, 30 and more variable susceptibility to other antibiotics in the population studied.

The main limitation of this study is that only 68% of samples tested were available for sequencing; this was due to a failure by the research team to ensure all samples taken for diagnostic purposes were successfully processed prior to sequencing within the study. This will have reduced the observed rates of linkage to previous cases, as demonstrated in previous simulations. 11 However, by applying a correction factor for missing data we were able to estimate the true proportion of cases linked. As rates of sample retrieval for sequencing were similar between the three hospitals, the differences observed in linkage rates are unlikely to have been differentially affected by sample retrieval rates at each site. The small number of samples, 5/299 (2%) infections/colonizations, that were faecal toxin positive but yielded isolates that lacked toxin genes on sequencing may have arisen as a result of mixed infections, laboratory error or a false-positive faecal toxin assay. Mixed infections are a potential additional source of underestimates of the extent of transmission from other cases, but previous work suggests this is uncommon in C. difficile. 11

This study was based on prospective storage of samples, culture of isolates and sequencing in response to a period of high CDI incidence. An alternative approach that may allow similar methods to be applied more widely is the storage of C. difficile GDH-positive faecal samples, e.g. on a rolling annual basis. These can then be cultured and sequenced retrospectively if increased incidence is noted, as demonstrated recently in six English hospitals. 11 The development of surveillance systems that interpret CDI incidence and sequencing data and present it back to clinicians in a timely manner is essential to guide local and national infection prevention and control responses.

In summary, despite relatively high CDI incidence in North Wales, levels of transmission detected by WGS were comparable to previously described rates in endemic settings; other explanations, including variations in antimicrobial use, are required to understand the reasons for the high levels of CDI.

Acknowledgements
We thank Dai Griffiths, Kim Bowden, Helen Booth, Gary Porter Jones, Melissa Van Der Bijl and Sarah Davies for their support.

Funding
The study was funded by the North Wales Awyrr Las charitable funds and by NHS (Wales) Research and Development support costs. D. W. E. is an NIHR Clinical Lecturer and a Robertson Foundation Fellow. T. E. A. P. is an NIHR Senior Investigator.

Transparency declarations
None to declare.

References


