Molecular epidemiology and antimicrobial resistance phenotype of paediatric bloodstream infections caused by Gram-negative bacteria in Oxfordshire, UK

Sam Lipworth[1,2], Karina-Doris Vihta[1,3], Tim Davies[1,2], Sarah Wright[2], Merline Tabirao[2], Kevin Chau[1], Alison Vaughan[1], James Kavanagh[1], Leanne Barker[1], Sophie George[1], Shelley Segal[2], Stephane Paulus[2], Lucinda Barrett[2], Sarah Oakley[2], Katie Jeffery[2], Lisa Butcher[2], Tim Peto[1,2,4,5], Derrick Crook[1,2,4,5], Sarah Walker[1,4,5], Seilesh Kadambari[2,6]* and Nicole Stoesser[1,2]*

*Joint Senior Authors

[1]Nuffield Department of Medicine, University of Oxford, UK

[2]Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[3]Department of Engineering, University of Oxford, UK

[4]NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford in partnership with Public Health England, Oxford, United Kingdom

[5]NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom

[6]Department of Paediatrics, University of Oxford, UK

Keywords: neonatal sepsis, paediatric bloodstream infections, whole genome sequencing, antimicrobial resistance

Running Title: Sequencing of paediatric Gram-negative bloodstream infections

Corresponding author: Dr S Lipworth, Samuel.lipworth@ndm.ox.ac.uk

Alternate corresponding author: Dr N Stoesser Nicole.stoesser@ndm.ox.ac.uk

Abstract:

Objectives: Gram-negative organisms are common causes of bloodstream infection (BSI) during the neonatal period and early childhood. Whilst several large studies have characterised these isolates in adults, equivalent data (particularly incorporating whole genome sequencing) is lacking in the paediatric population.

Methods: We performed an epidemiological and sequencing based analysis of Gram-negative bloodstream infections in children <18 years old between 2008 and 2018 in Oxfordshire, UK.

Results: 327 isolates (296 successfully sequenced) from 287 patients were included. The burden of infection was predominantly in neonates (124/327[38%]). Most infections were caused by *Escherichia coli (149/327[46%])/Klebsiella* spp. (69/327[21%]) and *Enterobacter hormaechei (34/327[10%])*. There was no evidence of an increasing incidence of *E. coli* BSIs (IRRy 0.96, 95%CI 0.90-1.30, p=0.30) and for *Klebsiella* spp. there was some evidence that the incidence decreased slightly (IRRy 0.91, 95%CI 0.83-1.00, p=0.06). Similarly the proportion of antimicrobial resistant (across all antimicrobial classes evaluated) isolates did not change over time, though we did identify some evidence of sub-breakpoint increases in gentamicin resistance IRRy 1.86, 95%CI 1.33-2.58, $p_{heterogeneity}$=0.002. The population structure of *E. coli* BSI isolates in neonates and children mirrors that in adults with a predominance of STs 131/95/73/69 and the same proportion of O-antigen serotypes covered by the ExPEC-4V vaccine. In most cases there was no evidence of transmission/point-source acquisition and whole genome sequencing was able to refute a previously suspected *Serratia marcescens* outbreak.

Conclusion: Our findings support continued use of current empirical treatment guidelines and suggest that O-antigen targeted vaccines may have a role in reducing the incidence of neonatal sepsis.

47 **Introduction**

48 Gram-negative bloodstream infections (GNBSI) are a common cause of significant morbidity and
49 mortality globally in neonates and young children[1–4]. Their incidence has increased in both the UK
50 and US over the past decade, particularly in very low birth-weight neonates (VLBW, defined as
51 <1500g)[5,6]. Their association with antimicrobial resistance (AMR) has been highlighted by a recent
52 study in the United States of 721 *E. coli* isolates (including 393 bloodstream infection [BSI]-
53 associated isolates), which found high rates of non-susceptibility to commonly used empirical
54 antibiotics including ampicillin (66.8%) and gentamicin (16.8%), as well as an extended beta-
55 lactamase (ESBL) phenotype in 1 in 20 cases[7]. A recent study of 2,483 neonates with culture-
56 confirmed sepsis in low and middle-income countries showed that *Klebsiella* spp. was the
57 predominant pathogen causing multidrug-resistant neonatal sepsis[8]. In Greece, a retrospective
58 observational study in 16 neonatal intensive care units (NICUs) revealed almost half (45%; 36/80) of
59 *Klebsiella* spp. were resistant to either gentamicin or amikacin[9]. The ability of many Gram-negative
60 bacilli (GNB) to readily acquire and exchange genetic material (particularly AMR genes [ARGs]) via
61 mobile genetic elements means that the proliferation of drug-resistant strains remains a constant
62 threat.

63

64 The molecular epidemiology of *E. coli* and *Klebsiella* spp. isolates causing invasive infection in adults
65 has been characterised in large sequencing studies.[10,11] These have demonstrated the emergence of
66 particular AMR-associated sequence types (e.g. *E. coli* ST131)[12], the genetic homogeneity of isolates
67 causing community and nosocomial onset infections suggesting a common reservoir[13], the potential
68 for vaccines to play a role in reducing the incidence of these infections[14,15], and the emerging threat
69 of the convergence of multidrug resistance and hypervirulence in *Klebsiella* spp[16]. To our knowledge,
70 no study to date has systematically evaluated the molecular epidemiology of *E. coli*/*Klebsiella* spp.
71 and other common causes of GNBSI in a paediatric population; published studies focus

72 predominantly on evaluations of outbreaks caused by AMR-associated strains and/or on neonates
73 (see above)[8,17,18]. In this study we therefore aimed to investigate sequencing data from a relatively
74 large collection of sequentially acquired, unselected bloodstream isolates from neonates and
75 children presenting to hospitals in Oxfordshire, UK, over the past decade.

76

77 **Methods**

78 **Isolate selection**

79 Oxford University Hospitals NHS Foundation Trust is a large healthcare provider in the South East of
80 England serving a paediatric population of approximately 142,000 across four hospitals (of which
81 two have emergency and acute general paediatric medicine, and one provides all
82 neonatal/paediatric critical care and specialist paediatric services for the region). The microbiology
83 laboratory additionally provides a service to all regional community healthcare providers. All *E. coli*
84 and *Klebsiella* spp. isolates (deduplicated to one morphotype per 90-day period) from Oct-2008 to
85 Nov-2018 collected from blood cultures of patients <18 years old on the day of collection were
86 included in the study. The same selection criteria were applied to other GNB from August-2011 to
87 September-2018 which were excluded from the initial period due to resource limitations. Prior to
88 2013 antimicrobial susceptibility testing was performed using disk diffusion; after this the Phoenix
89 BD system was used with European Committee on Antimicrobial Susceptibility Testing (EUCAST)
90 breakpoints. Amikacin susceptibility phenotyping was not routinely performed prior to 2013.

91

92 **Sequencing procedures**

93 Frozen stocks were sub-cultured onto Columbia blood agar and incubated overnight at 37°C. DNA
94 extractions were performed using the QuickGene DNA extraction kit (Autogen, MA, USA) as per the
95 manufacturer's instructions (with an additional mechanical lysis step – FastPrep, MP Biomedicals,

96 CA, USA; 6m/s for 40 secs, done twice). Sequencing was performed using Illumina HiSeq
97 2500/3000/4000/MiSeq instruments as described previously[19]. All sequencing data have been
98 deposited under NCBI accession number PRJNA604975.

99 **Bioinformatics**

100 *De novo* assembly was performed using Shovill (v1.0.4)[20]. Reads were mapped to sequence type
101 (ST[21,22]) specific references using Snippy (v4.6.0)[23] (Table S1). For the four *E. coli* STs with the largest
102 number of isolates (i.e. *E. coli* ST131/95/73/69), we created core genome alignments using Snippy-
103 core with the --mask auto setting, padded with the reference base at invariant positions; these
104 whole genome alignments were used as input to Gubbins (v 2.3.4)[24]. Such recombination-corrected
105 phylogenies were also created for *E.hormaechei* (the most common non-*E. coli*/*Klebsiella* species
106 detected in our study) and *S. marcescens* (because there was thought to have been an outbreak in
107 our neonatal intensive care unit in 2016). We also used genomic distances calculated by Mash[25]
108 (using -k21 -s 100,000 for within-study comparison of isolates and -k21 -s 1000 for comparison of
109 isolates in this study to our collection of sequences of adult bloodstream infection isolates[26] from the
110 same region over the same time period due to computational feasibility). Annotation against
111 reference databases (VFDB/ResFinder) was performed using ABRicate (v2.3.4)[27] with genes called as
112 being present if there was ≥80% coverage and DNA identity compared to the reference. Sequence
113 types were predicted using the MLST tool (v2.19)[28]. For *Klebsiella* spp. isolates speciation and
114 virulence gene detection (appendix) was performed using Kleborate (v2.0.4)[29]. Detailed QC metrics
115 and raw Abricate/Kleborate output have been uploaded to Figshare
116 (https://figshare.com/projects/Paediatric_GNBSIs_in_Oxfordshire/135254).

117

118 **Definitions**

119 We defined isolates as being likely of neonatal origin if they originated from infants i) in their first 30
120 days of life, or were ii) on the neonatal intensive care unit, or iii) under the care of a neonatologist
121 on the day the blood culture was taken. For analytical purposes we classified other children as <12
122 months, 1-4 years, 5-9 years and 10-17 years of age. Early-onset infection was defined as disease
123 within the first 72 hours of life[30]. We calculated a 'resistance score' using a previously described
124 method[31] (the sum of the number of resistance gene categories carried out of amoxicillin, co-
125 trimoxazole, cefotaxime, gentamicin and ciprofloxacin). We further categorised BSIs according to
126 healthcare exposure prior to onset as follows: nosocomial (>48hours after admission to hospital),
127 'quasi-nosocomial' (within 30 days of last discharge), 'quasi-community' (31-365 days since last
128 discharge) and community (>365 days since last discharge)[32]. Genetic relatedness in the form of
129 single nucleotide polymorphism (SNP) thresholds definitively associated with transmission is variably
130 defined for the species evaluated; based on recent studies, we considered a threshold of >20 SNVs
131 between isolates as highly unlikely to be representative of a transmission event[33].

132

133 **Epidemiology/statistics**

134 Routinely collected healthcare data were acquired via pseudonymised linkage in the Infections in
135 Oxfordshire Research Database (IORD). IORD has generic Research Ethics Committee, Health
136 Research Authority and Confidentiality Advisory Group approvals (19/SC/0403, 19/CAG/0144) as a
137 de-identified electronic research database. Data on suspected infectious focus (only available for *E.*
138 *coli/Klebsiella* spp.) were acquired via linked local infection control records which had been
139 submitted to Public Health England as part of the mandatory surveillance programme. Likely source
140 was identified by infectious disease/microbiology physicians using best clinical judgement or
141 designated as unknown where there was significant uncertainty. For each species, we modelled the
142 number of bloodstream infections (BSIs) per year using negative binomial regression, with the total
143 number of paediatric admissions in each year used as an offset to account for changes in the

144 population over time. Only complete years (i.e. excluding 2008 and 2018) were included in this part
145 of the analysis. Median quantile regression was used to model the relationship between age and
146 virulence gene carriage/antimicrobial gene "resistance score"; this analysis was performed in STATA
147 version 15. All other statistical analysis was performed in R v4.0.3[34,35].

148

149 **Results**

150 **Microbiology of neonatal and paediatric Gram-negative bloodstream infections in Oxfordshire**

151 Of the 327 GNBSI isolates cultured during the study period from individuals aged <18 years, 149
152 (46%) were identified as *E. coli* and 69 (21%) as *Klebsiella* spp.; the remaining 109 (33%) belonged to
153 other species (n.b. latter category only collected from 2011). There was no evidence of a change in
154 the incidence of *E. coli* bloodstream infections (BSIs) either overall (incidence rate ratio per year [i.e.
155 the relative increase/decrease in incidence per year] IRRy: 0.96, 95%CI: 0.90-1.03, p=0.30) or in
156 neonates (IRRy: 0.97, 95%CI: 0.86-1.06, p=0.39). This was also the case for other Gram-negative
157 species (i.e. non *E. coli* and *Klebsiella spp.*) both overall (IRRy: 1.07, 95%CI: 0.93-1.24, p=0.33) and in
158 the neonatal group (IRRy: 0.83, 95%CI: 0.65-1.07, p=0.19). Conversely there was some evidence of a
159 decrease in the overall number of *Klebsiella* spp. BSIs in the same period (overall IRRy: 0.91, 95%CI
160 0.83-1.00, p=0.06; neonates IRRy: 0.80, 95%CI: 0.69-0.92, p=0.002).

161

162 Neonates had the highest number of GNBSI (124/327 [38%], of which 8 were early-onset) followed
163 by other infants <1y (66/327 [20%]) (Table 1). There were 62/327 (19%) cases in the 1st-4th years of
164 life, 28/327 (9%) in the 5th-9th and 47/327 (14%) in the 10th-17th. A higher proportion of isolates came
165 from males than would be expected from the birth:sex ratio (201/327 [61%], multinomial goodness
166 of fit p <0.001). BSIs in neonates, children aged 1-4 and 10-17 were predominantly nosocomial or

167 quasi-nosocomial whereas in those aged 3-12 months and 5-9 years they were predominantly
168 community or quasi-community onset (Table 1).

169

170 In our network of hospitals, amoxicillin and gentamicin are used as empirical treatment for neonatal
171 sepsis. In this patient group, 87/119 (73%) isolates were resistant to amoxicillin and 10/122 (8%) to
172 gentamicin (missing data for five and two isolates respectively); 7/118 (6%) isolates (4/56 *E. coli* and
173 3/27 *Klebsiella* spp.) were resistant to both. In older children (>1 month), ceftriaxone is the empirical
174 agent used in our setting and in this population 24/183 (13%) isolates were resistant (missing data
175 for 20 isolates). The proportion of resistant isolates appeared broadly stable over time (figure 1/S1).
176 For gentamicin however, there was some evidence that proportion of susceptible isolates with an
177 MIC >1mg/L (IRRy 1.86, 95%CI 1.33-2.58) increased compared to those with an MIC <=1mg/L (IRRy
178 1.13, 95%CI 1.04-1.22) $p_{heterogeneity}$=0.002 (EUCAST gentamicin breakpoint for resistance >2mg/L).
179 There was no evidence of substantial differences in phenotypic profiles between treating specialties
180 (Figure S2). Likewise, the proportion of resistant isolates was similar for nosocomial vs. community-
181 onset cases for all agents except amoxicillin where nosocomial isolates were proportionally more
182 resistant (99/118 (84%) vs 115/174 (66%), table S2), reflecting the higher burden of
183 *Klebsiella*/*Enterobacter* spp. BSIs in this patient group (77/126 (61%) vs 101/201 (50%); both these
184 genera are normally considered intrinsically resistant to amoxicllin.

185

186 Data on likely primary infective focus was only available for 78 isolates (66 *E. coli*/12 *Klebsiella* spp.),
187 of these 18 (23%) were of gastrointestinal origin, 8 (10%) from invasive lines, 20 (26%) from the
188 urinary tract, 2 (3%) from the chest and 30 (38%) from unclear/other sources. Seventeen children
189 (six neonates) had infections with >1 species and five (two neonates) had polyclonal infections (>1 ST
190 of the same species).

191

192 **Molecular epidemiology of *Escherichia coli* causing BSI, including AMR and virulence gene burden**

193 Of the 149 *E. coli* isolates cultured, 133 (87%) were successfully retrieved for sequencing (Figure 2).
194 The four dominant *E. coli* STs (ST131 n=16, ST95 n=16, ST73 n=20 and ST 69 n=15) were the same as
195 those in adults[26]. There was no evidence that the population structure observed differed from that
196 in the adult population in Oxfordshire (which is also the same as that observed globally[26];
197 multinomial goodness of fit: p=0.44). The most prevalent O-antigen types were O6 (n=21), O1
198 (n=14), O2 (n=12), O16 (n=10) and O25 (n=8). The proportion of isolates with an O-antigen included
199 in the ExPEC-4V vaccine[36,37] (which has been evaluated in a phase-II study in adults), was similar
200 when comparing all children (55/133, 41%) and neonates (25/53, 47%) to that seen in the overall
201 population in our previous study[14] (1499/3278 (46%); chi-squared: p=0.4 for both).

202

203 Carriage of genes conferring resistance to ampicillin was common (65/133, 49% isolates carrying 82
204 alleles), with bla_{TEM-1B} (37/82, 45%) being the most prevalent allele. Eight isolates carried a
205 gentamicin resistance gene (5 *aac(3)-IIa* [3 in ST131, 1 each in STs 73 and 23] and 3 *aac(3)-IId* [1 each
206 in STs 95, 131 and 69]) and 18/133 (14%) carried a gene conferring resistance to ceftriaxone (mostly
207 ESBL genes, namely $bla_{CTX-M-15}$ [n=9], $bla_{CTX-M-14b}$ [n=3] and $bla_{SHV-102}$ [n=5]); Figure 3). Whilst ST131 was
208 the dominant ESBL-gene-carrying ST (7/16, 44%), other STs carrying these genes included ST73
209 (4/20), ST2141 (3/3), ST12 (2/7) and STs 23 and 38 (1 each). Quantile regression revealed some
210 association between increasing age and increased resistance score (change in median per year,
211 CMy=0.12, 95% CI 0.04-0.19, p =0.002). Conversely there was some evidence that virulence gene
212 carriage (number of such genes detected by Abricate in the VFDB) was higher in younger children
213 (CMy = -0.53 95%CI -1.03 - -0.033, p=0.04 Figure S3).

214

215 **Molecular epidemiology of *Klebsiella* spp. causing BSI, including AMR and virulence gene burden**

216 Of the 69 *Klebsiella* spp. isolates cultured, 60 (87%) were successfully retrieved for sequencing. The
217 predominant *Klebsiella* spp. in both neonates and older infants/children was *Klebsiella pneumoniae*
218 (n=28/60 [47%]), though a diverse selection of related species were also occasionally isolated (*K.*
219 *oxytoca*: n=12 [20%], *K. michiganensis*: n=8 (13%), *K. aerogenes*: 5 (8%), *K. grimontii*: n=3 (5%), *K.*
220 *ornithinolytica*: n=2 (3%) and *K. variicola*: n=2 (3%)). As we have previously observed in adults and in
221 contrast to the more clonal population seen in *E. coli*, STs causing *Klebsiella* spp. BSI were diverse
222 (n=32 STs in total). There were 11 isolates to which no ST could be assigned. We recently observed
223 that four O-antigens (O2v2, O1v1, O3b and O1v2) were found in 75% isolates from all *K. pneumoniae*
224 BSIs in Oxfordshire over a ten-year period; amongst paediatric BSIs this figure was 61% (17/28;
225 Fisher's exact test: p=0.15).

226

227 Half of *Klebsiella* spp. isolates (30/60) carried the yersiniabactin virulence factor (*ybt*); these were
228 mostly non-*K. pneumoniae* (12/12 *K. oxytoca*, 7/8 *K. michiganensis*, 2/2 *K. ornitholytica* and 3/3 *K.*
229 *grimontii*). One *K. aerogenes* isolate additionally carried the genotoxin colibactin (*clb*). There was no
230 difference in the proportion of E. coli/Klebsiella spp. isolates carrying a gene conferring ceftriaxone
231 resistance (18/133, 14% vs 7/60, 12%, p=0.82). Only 4/60 (7%) isolates carried a gentamicin
232 resistance gene (four *aac(3)-IIa* and one *aac(3)-Ia*, similar to *E. coli* (8/133, 6%); Fisher's exact test:
233 p=0.90).

234

235 **Other Gram-negative species**

236 Of the 104 non-*E. coli*/*Klebsiella* spp. ("other") GNB cultured in this study, 77 (74%) were successfully
237 sequenced. The predominant species in both neonates and older children was *E. hormaechei*
238 (overall: 34/75, [44%]; neonatal: 14/25 [56%]). This was also the only other species carrying ESBL-

239 producing genes ($bla_{CTX\text{-}M\text{-}15}$ in four isolates, Figure 3); these isolates also carried the $bla_{TEM\text{-}1B}$ beta-
240 lactamase as well as *aac(3)-IIa* conferring gentamicin resistance. Two isolates carrying class-D beta-
241 lactamases predicted to confer carbapenem resistance were identified ($bla_{OXA\text{-}23}$ in *Acinetobacter*
242 *baumannii*, $bla_{OXA\text{-}427}$ in *Aeromonas caviae*).

243

244 **Genomic relatedness of isolates**

245 Analysis of recombination-corrected phylogenies for the major *E. coli* STs (131/95/69) demonstrated
246 no clear evidence of direct or temporally related indirect transmission between patients (with the
247 closest genomic distance between patients = 32/87/28 single nucleotide variants [SNV] for
248 ST131/95/69 respectively). There was a cluster of three near-identical ST73 isolates from two
249 patients (3/6 SNPs apart), however the isolation date was ~2 years apart and these patients had
250 never been admitted to hospital at the same time. There was a cluster of three multidrug-resistant
251 *E. hormaechei* isolates separated by a maximum of 37 SNPs. All three patients had been admitted in
252 the same year but not at the same time and to different wards. Other *E. hormaechei* isolates were
253 genomically diverse (Figure 4). Three *S. marcescens* from the neonatal ICU in 2016 were grouped
254 together in the phylogeny (Figure S4) however the relatively large distances between them
255 (31/81/85 SNPs), made recent transmission unlikely.

256

257 There was no difference in the distributions of genomic distances (reflected here by Mash distances)
258 between isolates from patients admitted i) in the same year, ii) to the same ward in the same year,
259 or iii) to the same ward with overlapping admission dates and iv) overall (i.e. all admissions to any
260 ward at any time point) (Kruskal-Wallis rank sum test p=0.30, Figure S5). Finally, when considering *E.*
261 *coli* and *Klebsiella* isolates in this study, there was no difference between the Mash distances to the
262 nearest paediatric genomic neighbour (median: 0.00093 [IQR: 0.00036-0.0023]) compared to the

263 nearest adult neighbour (median: 0.0011 [IQR: 0.00041-0.0024], Wilcoxon rank sum p=0.33. Figure
264 S6).

265

266 **Discussion**

267 *E. coli* was the main causative agent of paediatric and neonatal Gram-negative bloodstream infection
268 in our study over a ten year period. The population structure of *E. coli* isolates causing invasive
269 infection in children mirrors that seen in our centre and globally for adults. We found no evidence of
270 significant nosocomial transmission of isolates causing bloodstream infection, suggesting that in our
271 setting invasive isolates are acquired from the environment, represent colonising gastrointestinal
272 flora[38], or as a result of transmission from parents/other close contacts. Our findings suggest O-
273 antigen targeted vaccines (currently in phase II/III trials in adults) might have the potential to reduce
274 the incidence of invasive neonatal and paediatric *E. coli* infections, in the former potentially by
275 immunisation of pregnant women. Most isolates remained susceptible to at least one first-line
276 antibiotic justifying continuation of current empirical treatment guidelines which should include an
277 aminoglycoside in neonates[30]. There was however some evidence of increasing (sub-breakpoint)
278 gentamicin MIC which warrants ongoing surveillance, and 1 in 20 GNBSIs were caused by isolates
279 resistant to the ampicillin+gentamicin combination.

280

281 The incidence of these infections has remained stable in Oxfordshire over the past decade in
282 children, in contrast to the increase observed in adults. The substantial burden of disease due to
283 paediatric GNBSI occurs in neonates. The absence of evidence of an increasing burden of AMR-
284 associated disease and demonstrable outbreaks likely points to good infection control practice and
285 antimicrobial stewardship, and in our setting we were able to replicate results from a previous
286 study[31] suggesting greater carriage of virulence genes in *E. coli* isolates from younger children and

287 neonates but higher resistance scores in older children. The latter finding might be explained by
288 increased hospital contact and environmental exposure in this group whereas reasons for and
289 clinical significance of virulence gene acquisition is unclear. Continued epidemiological surveillance
290 with targeted whole genome sequencing to identify strain and/or plasmid outbreaks may help to
291 inform early infection control interventions and would seem particularly justified in these age groups
292 given the high associated morbidity and mortality. In addition we would encourage consideration of
293 clinical trials of existing O-antigen targeted vaccines for prevention of neonatal sepsis.

294

295 Our data highlight the importance of *E. hormachei* and as a potential causative agent of neonatal
296 and paediatric sepsis as well as its potential to cause multidrug-resistant disease in this setting.
297 Several small clusters of *E. hormaechei* (part of the *E. cloacae* complex) bloodstream infection in
298 neonates have been described in the literature, with enteral feeding and milk formulations as well as
299 poor general infection control procedures thought to be implicated[39–41].

300

301 Our data highlights the value of whole genome sequencing for the investigation of potential
302 outbreaks on neonatal units. WGS analysis of three isolates from a suspected (on the basis of
303 epidemiology and pulsed-field gel electrophoresis) *S. marcescens* outbreak suggested recent
304 transmission was actually highly unlikely. Had this analysis been performed in real-time it may have
305 reassured clinicians and saved time and resource. Apart from a small cluster of three patients with *E.*
306 *hormaechei* infections where acquisition from a common source was possible based on genomic and
307 epidemiological analysis (though the patients were never admitted to the same ward), there was
308 relatively high diversity amongst isolates effectively ruling out transmission/point-source acquisition
309 resulting in BSI. However, unwell neonates are often transferred across a geographic region for
310 specialist care and so it is essential that active surveillance mechanisms (collecting and analysing

311 real-time microbiological data) are implemented across networks in order to inform infection
312 prevention measures appropriately.

313

314 Notably, whilst we found a lack of clear evidence of transmission between patients in this study, BSIs
315 are likely to represent an extremely insensitive marker of transmission events, as most transmission
316 with these organisms would be expected to lead to colonisation and not necessarily invasive
317 infection. Furthermore whilst the short-read sequencing used in this study allows us to confidently
318 exclude outbreaks of strains, we cannot address the question of potential transmission of AMR
319 genes and virulence factors on plasmids. Additional limitations include the fact that we were not
320 able to sequence all cultured isolates, the relatively sparse availability of data on source attribution
321 and the single region nature of the study. We did not sample other reservoirs potentially
322 contributing to transmission, including the unit environment.

323

324 In summary, our study demonstrates a flat overall incidence trend and stable numbers of AMR-
325 carrying GNBSI isolates in children, contrasting with what has been seen in adult populations in the
326 same setting[26]. Pertinently our data supports the ongoing use of an aminoglycoside as part of
327 empirical treatment guidelines in neonatal sepsis. The disproportionate impact of GNBSI on
328 neonates and young infants should encourage active microbiological surveillance across neonatal
329 networks and strategies to prevent disease through vaccine trials in pregnant women.

330

331

332 **Figure legends**

333 Figure 1 – Change in minimum inhibitory concentrations (MIC) of isolates over time.

334 Figure 2 - Species (and sequence type, ST, where shown) for sequenced isolate over time. Each
335 point represents the isolation date of a sequenced isolate. The grey shaded area represents the fact
336 that sequencing of non-*E. coli*/*Klebsiella* spp. isolates did not begin prior to 2012.

337 Figure 3 – Carriage of antimicrobial resistance genes by isolates sequenced in the study. The tree
338 was created using mashtree and gene presence is shown in red. The black/green/maroon/orange
339 bars at the top represent carriage of genes producing enzymes with activity (as predicted by
340 Resfinder) against aminoglycosides/amoxicillin/cefotaxime (+amoxicillin)/ceftriaxone
341 (+cefotaxime/amoxicillin) respectively.

342

343 Figure 4 – Recombination corrected phylogeny of *Enterobacter hormaechei* isolates sequenced in
344 this study. Tree tip colours show the year of isolates, red bars denote presence/absence of
345 antimicrobial resistance genes. The scale of the tree is shown in SNPs.

346

347 **Table legend**

348 Table 1 – microbiological and patient characteristics of patients with Gram-negative bloodstream
349 infections in Oxfordshire stratified by age of onset. P values represent row-wise comparisons using
350 Fisher exact tests for categorical and Kruskal-Wallis tests for continuous variables.

351

352 **Ethics Statement**

353 The Infections in Oxfordshire Research Database (IORD; https://oxfordbrc.nihr.ac.uk/research-
354 themes-overview/antimicrobial-resistance-and-modernising-microbiology/infections-in-oxfordshire-
355 research-database-iord/) has generic Research Ethics Committee, Health Research Authority and
356 Confidentiality Advisory Group approvals (19/SC/0403, 19/CAG/0144) which facilitate the pseudo-

357 anonymised linkage of routinely collected NHS electronic healthcare record data from the Oxford
358 University Hospitals NHS Foundation Trust Clinical Systems Data Warehouse and research data (e.g.
359 sequencing data) from the Antimicrobial Resistance and Modernising Microbiology Theme of the
360 Oxford NIHR Biomedical Research Centre, Oxford. IORD links records by a specific, random, number
361 ensuring that no patient-identifiable information is shared with researchers using this resource. We
362 sequenced bacterial isolates from bloodstream infections that are routinely stored by the John
363 Radcliffe Hospital Microbiology laboratory. In the UK, bacterial isolates (such as those sequenced in
364 this study) routinely cultured from human clinical samples do not require ethical approval for
365 analysis under the provisions of the Human Tissue Act as they do not contain any material
366 considered to be human tissue.

367

368 **References**

369 1. Malande OO, Nuttall J, Pillay V, Bamford C, Eley B. A ten-year review of ESBL and non-ESBL
370 Escherichia coli bloodstream infections among children at a tertiary referral hospital in South
371 Africa. *PLoS One*. 2019;14(9):e0222675.

372 2. Powell EC, Mahajan PV, Roosevelt G, et al. Epidemiology of Bacteremia in Febrile Infants Aged
373 60 Days and Younger. *Ann Emerg Med*. 2018;71(2):211-216.

374 3. Shakur SM, Whitehall J, Mudgil P. Pediatric bloodstream infections in metropolitan Australia.
375 *World J Pediatr*. 2019;15(2):161-167.

376 4. Agyeman PKA, Schlapbach LJ, Giannoni E, et al. Epidemiology of blood culture-proven bacterial
377 sepsis in children in Switzerland: a population-based cohort study. *Lancet Child Adolesc Health*.
378 2017;1(2):124-133.

379 5. Stoll BJ, Puopolo KM, Hansen NI, et al. Early-Onset Neonatal Sepsis 2015 to 2017, the Rise of
380 Escherichia coli, and the Need for Novel Prevention Strategies. *JAMA Pediatr*.
381 2020;174(7):e200593.

382 6. *Annual Publication of Epidemiological Commentary*. Public Health England; 2020.
383 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_dat
384 a/file/940716/Annual_epidemiology_commentary_April_2019_March_2020.pdf

385 7. Flannery DD, Akinboyo IC, Mukhopadhyay S, et al. Antibiotic Susceptibility of Escherichia coli
386 Among Infants Admitted to Neonatal Intensive Care Units Across the US From 2009 to 2017.
387 *JAMA Pediatr*. 2021;175(2):168-175.

388 8. Sands K, Carvalho MJ, Portal E, et al. Characterization of antimicrobial-resistant Gram-negative
389 bacteria that cause neonatal sepsis in seven low- and middle-income countries. *Nat Microbiol*.
390 2021;6(4):512-523.

391 9. Gkentzi D, Kortsalioudaki C, Cailes BC, et al. Epidemiology of infections and antimicrobial use in
392 Greek Neonatal Units. *Arch Dis Child Fetal Neonatal Ed*. 2019;104(3):F293-F297.

393 10. Kallonen T, Brodrick HJ, Harris SR, et al. Systematic longitudinal survey of invasive Escherichia
394 coli in England demonstrates a stable population structure only transiently disturbed by the
395 emergence of ST131. *Genome Res*. Published online July 18, 2017. doi:10.1101/gr.216606.116

396 11. Ellington MJ, Heinz E, Wailan AM, et al. Contrasting patterns of longitudinal population
397 dynamics and antimicrobial resistance mechanisms in two priority bacterial pathogens over 7
398 years in a single center. *Genome Biol*. 2019;20(1):184.

399 12. Petty NK, Ben Zakour NL, Stanton-Cook M, et al. Global dissemination of a multidrug resistant
400 Escherichia coli clone. *Proc Natl Acad Sci U S A*. 2014;111(15):5694-5699.

401 13. Lipworth SIW, Vihta KD, Chau K, et al. Ten year longitudinal molecular epidemiology study of
402 Escherichia coli and Klebsiella species bloodstream infections in Oxfordshire, UK. *Genome Med*.
403 Published online 2021.

404 14. Lipworth S, Vihta KD, Chau KK, et al. Ten years of population-level genomic Escherichia coli and
405 Klebsiella pneumoniae serotype surveillance informs vaccine development for invasive
406 infections. *Clin Infect Dis*. Published online January 7, 2021. doi:10.1093/cid/ciab006

407 15. Royer G, Clermont O, Condamine B, et al. O-Antigen Targeted Vaccines Against Escherichia coli
408 May Be Useful in Reducing Morbidity, Mortality, and Antimicrobial Resistance. *Clin Infect Dis*.
409 2022;74(2):364-366.

410 16. Wyres KL, Nguyen TNT, Lam MMC, et al. Genomic surveillance for hypervirulence and multi-
411 drug resistance in invasive Klebsiella pneumoniae from South and Southeast Asia. *Genome*
412 *Med*. 2020;12(1):11.

413 17. Giuffrè M, Cipolla D, Bonura C, et al. Outbreak of colonizations by extended-spectrum β-
414 lactamase-producing Escherichia coli sequence type 131 in a neonatal intensive care unit, Italy.
415 *Antimicrob Resist Infect Control*. 2013;2(1):8.

416 18. Stoesser N, Giess A, Batty EM, et al. Genome sequencing of an extended series of NDM-
417 producing Klebsiella pneumoniae isolates from neonatal infections in a Nepali hospital
418 characterizes the extent of community- versus hospital-associated transmission in an endemic
419 setting. *Antimicrob Agents Chemother*. 2014;58(12):7347-7357.

420 19. De Maio N, Shaw LP, Hubbard A, et al. Comparison of long-read sequencing technologies in the
421 hybrid assembly of complex bacterial genomes. *Microb Genom*. 2019;5(9).
422 doi:10.1099/mgen.0.000294

423 20. Seemann T. Shovill. Accessed May 18, 2020. https://github.com/tseemann/shovill

424 21. Wirth T, Falush D, Lan R, et al. Sex and virulence in Escherichia coli: an evolutionary perspective.
425 *Mol Microbiol*. 2006;60(5):1136-1151.

426 22. Diancourt L, Passet V, Verhoef J, Grimont PAD, Brisse S. Multilocus sequence typing of Klebsiella
427 pneumoniae nosocomial isolates. *J Clin Microbiol*. 2005;43(8):4178-4182.

428 23. Seemann T. *Snippy*. Github; 2015. Accessed September 12, 2017.
429 https://github.com/tseemann/snippy

430 24. Croucher NJ, Page AJ, Connor TR, et al. Rapid phylogenetic analysis of large samples of
431 recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*.
432 2015;43(3):e15.

433 25. Ondov BD, Treangen TJ, Melsted P, et al. Mash: fast genome and metagenome distance
434 estimation using MinHash. *Genome Biol*. 2016;17(1):132.

435 26. Lipworth S, Vihta KD, Chau K, et al. Molecular epidemiology of Escherichia coli and Klebsiella
436 species bloodstream infections in Oxfordshire (UK) 2008-2018. *medRxiv*. Published online 2021.

437 27. Seemann T. *Abricate*. Github Accessed July 5, 2019. https://github.com/tseemann/abricate

438 28. Seemann T. *Mlst*. Github Accessed July 12, 2019. https://github.com/tseemann/mlst

439 29. Holt K. Kleborate. Accessed May 22, 2020. https://github.com/katholt/Kleborate

440 30. *Neonatal Infection: Antibiotics for Prevention and Treatment*. National Institute for Clinical
441 Excellence Accessed May 4, 2021. https://www.nice.org.uk/guidance/ng195

442 31. Burdet C, Clermont O, Bonacorsi S, et al. Escherichia coli bacteremia in children: age and portal
443 of entry are the main predictors of severity. *Pediatr Infect Dis J*. 2014;33(8):872-879.

444 32. Vihta KD, Stoesser N, Llewelyn MJ, et al. Trends over time in Escherichia coli bloodstream
445 infections, urinary tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998–2016:
446 a study of electronic health records. *Lancet Infect Dis*. 2018;18(10):1138-1149.

447 33. Wyres KL, Hawkey J, Mirčeta M, et al. Genomic surveillance of antimicrobial resistant bacterial
448 colonisation and infection in intensive care patients. *BMC Infect Dis*. 2021;21(1):683.

449 34. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2019.
450 https://www.R-project.org/

451 35. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. *Journal of Open Source*
452 *Software*. 2019;4(43):1686. doi:10.21105/joss.01686

453 36. Frenck RW Jr, Ervin J, Chu L, et al. Safety and immunogenicity of a vaccine for extra-intestinal
454 pathogenic Escherichia coli (ESTELLA): a phase 2 randomised controlled trial. *Lancet Infect Dis*.
455 2019;19(6):631-640.

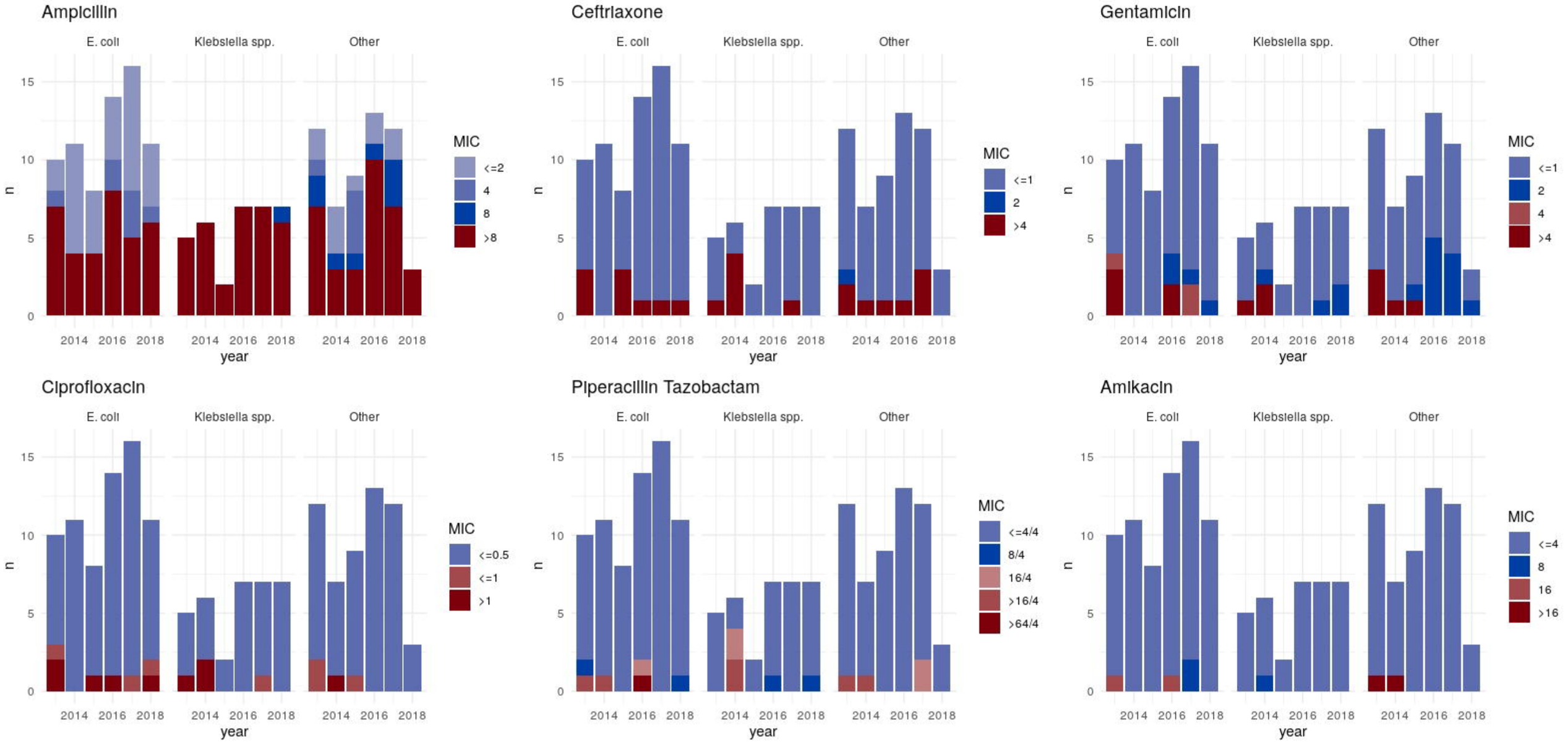
456 37. Huttner A, Gambillara V. The development and early clinical testing of the ExPEC4V conjugate
457 vaccine against uropathogenic Escherichia coli. *Clin Microbiol Infect*. 2018;24(10):1046-1050.

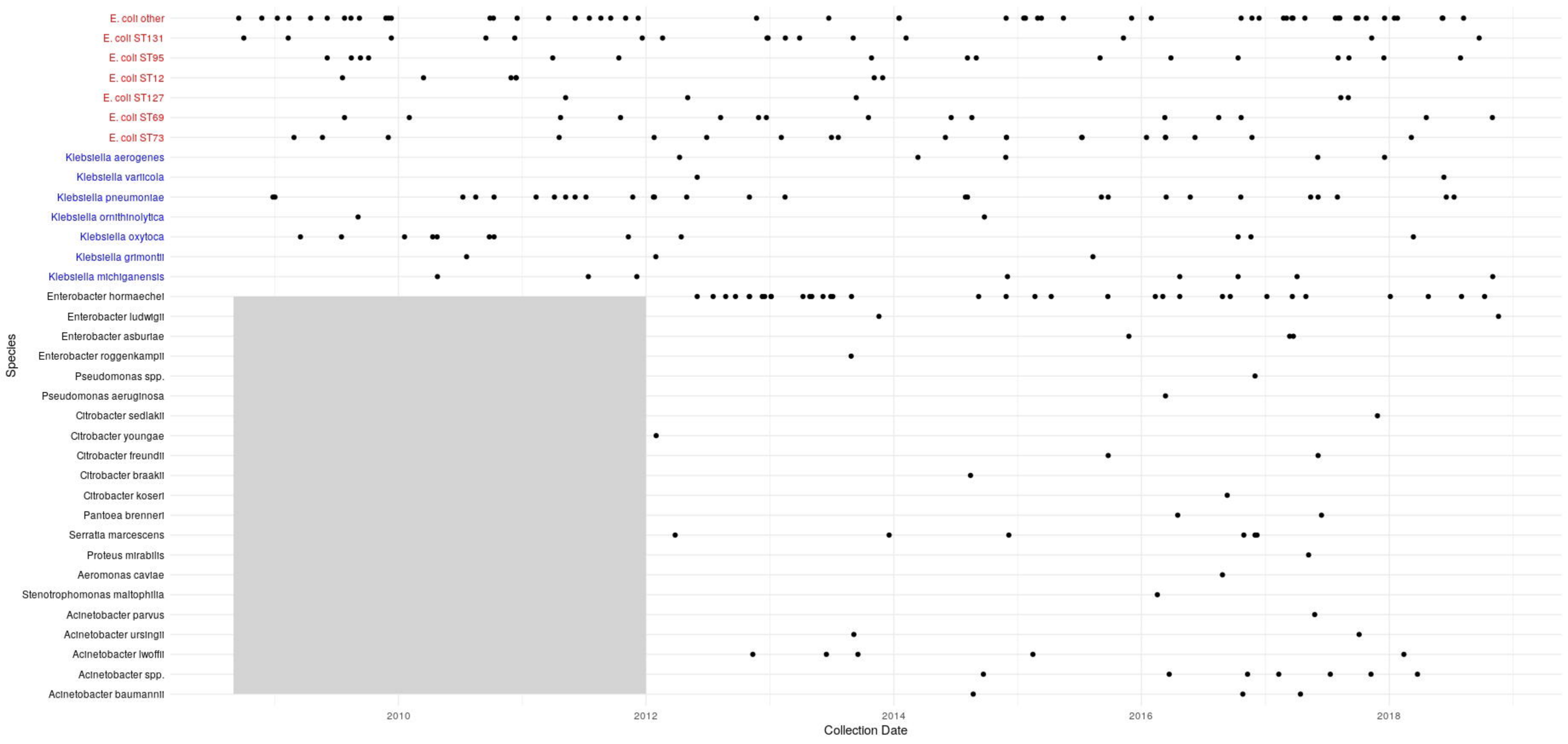
458 38. Tamburini FB, Andermann TM, Tkachenko E, Senchyna F, Banaei N, Bhatt AS. Precision
459 identification of diverse bloodstream pathogens in the gut microbiome. *Nat Med*.
460 2018;24(12):1809-1814.

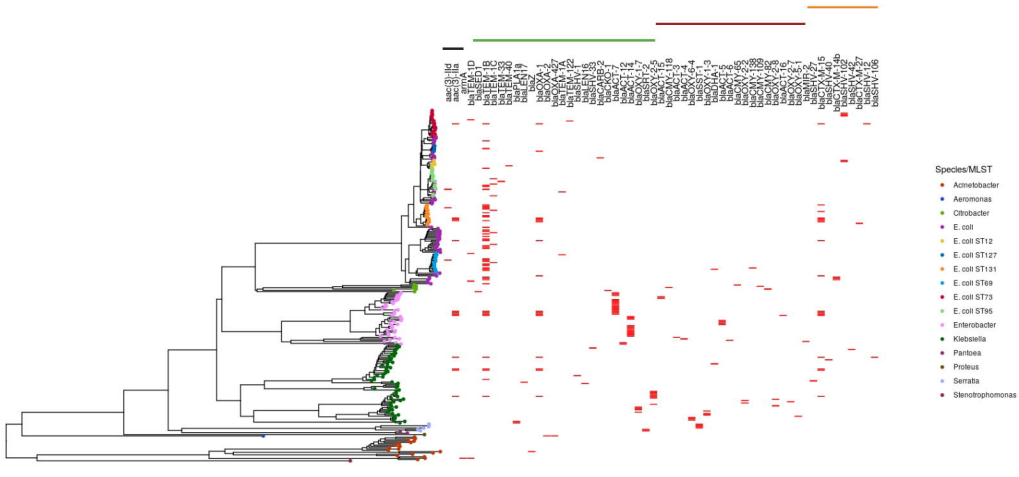
461 39. Giammanco GM, Aleo A, Guida I, Mammina C. Molecular epidemiological survey of Citrobacter
462 freundii misidentified as Cronobacter spp. (Enterobacter sakazakii) and Enterobacter
463 hormaechei isolated from powdered infant milk formula. *Foodborne Pathog Dis*. 2011;8(4):517-
464 525.

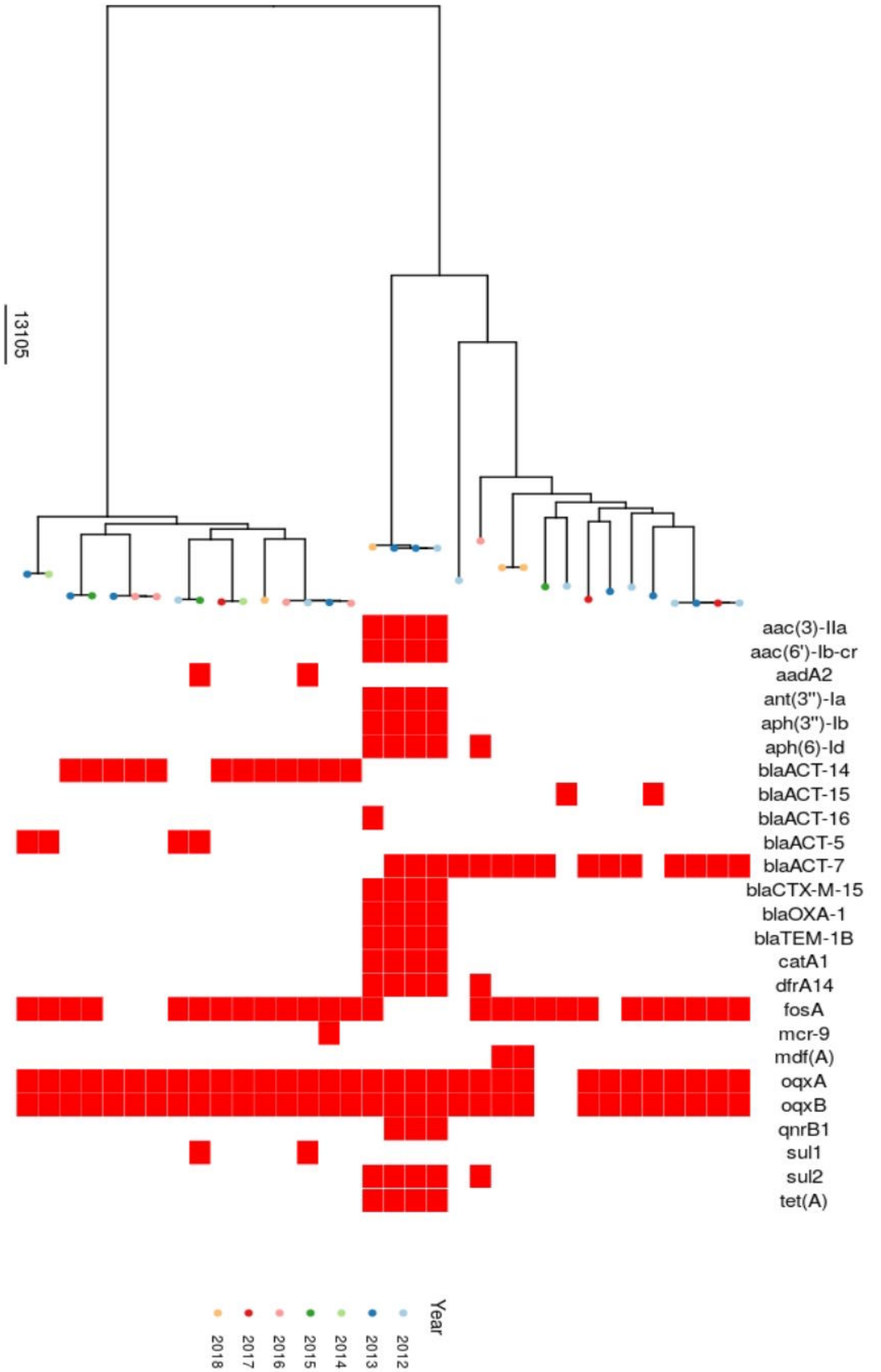
465 40. El-Zamkan MA, Mohamed HMA. Molecular detection ofEnterobacterspp. and other related
466 species in powdered milk infant formula and milk powder. *J Food Saf*. 2018;38(6):e12538.

467 41. da Silva CLP, Miranda LEV, Moreira BM, et al. Enterobacter hormaechei bloodstream infection
468 at three neonatal intensive care units in Brazil. *Pediatr Infect Dis J*. 2002;21(2):175-177.


Ampicillin
E. coli
Klebsiella spp.
Other
n
year
MIC
<=2
4
8
>8
Ceftriaxone
MIC
<=1
2
>4
Gentamicin
MIC
<=1
2
4
>4
Ciprofloxacin
MIC
<=0.5
<=1
>1
Piperacillin Tazobactam
MIC
<=4/4
8/4
16/4
>16/4
>64/4
Amikacin
MIC
<=4
8
16
>16


E. coli other
E. coli ST131
E. coli ST95
E. coli ST12
E. coli ST127
E. coli ST69
E. coli ST73
Klebsiella aerogenes
Klebsiella varticola
Klebsiella pneumoniae
Klebsiella ornithinolytica
Klebsiella oxytoca
Klebsiella grimontii
Klebsiella michiganensis
Enterobacter hormaechei
Enterobacter ludwigii
Enterobacter asburiae
Enterobacter roggenkampii
Pseudomonas spp.
Pseudomonas aeruginosa
Citrobacter sedlakii
Citrobacter youngae
Citrobacter freundii
Citrobacter braakii
Citrobacter koseri
Pantoea brenneri
Serratia marcescens
Proteus mirabilis
Aeromonas caviae
Stenotrophomonas maltophilia
Acinetobacter parvus
Acinetobacter ursingii
Acinetobacter lwoffii
Acinetobacter spp.
Acinetobacter baumannii
Species
2010
2012
2014
2016
2018
Collection Date


aac(3)-Id
aac(3)-Ia
aadA1
baTEM-1D
baSHV-1
baTEM-1B
baTEM-1C
baTEM-30
baTEM-70
baPER-1
baOXA-1
baOXA-427
baTEM-122
baTEM-116
baSHV-12
baSHV-69
baCMY-2
baOXA-2
baACT-1
baACT-12
baOXA-17
baOXA-10
baOXA-25
baOXA-18
baACT-4
baOXA-64-4
baOXA-2
baOXA-13
baACT-5
baACT-15
baCMY-55
baCMY-2-like
baCMY-89
baOXA-28
baOXA-29
baOXA-27
baOXA-15
baTMB-2
baCTX-M-15
baCTX-M-14b
baSHV-12
baCTX-M-2
baCTX-M-27
baSHV-106
Species/MLST
Acinetobacter
Aeromonas
Citrobacter
E. coli
E. coli ST12
E. coli ST127
E. coli ST131
E. coli ST69
E. coli ST73
E. coli ST95
Enterobacter
Klebsiella
Pantoea
Proteus
Serratia
Stenotrophomonas


13105
aac(3)-IIa
aac(6')-Ib-cr
aadA2
ant(3'')-Ia
aph(3'')-Ib
aph(6)-Id
blaACT-14
blaACT-15
blaACT-16
blaACT-5
blaACT-7
blaCTX-M-15
blaOXA-1
blaTEM-1B
catA1
dfrA14
fosA
mcr-9
mdf(A)
oqxA
oqxB
qnrB1
sul1
sul2
tet(A)
Year
2012
2013
2014
2015
2016
2017
2018

		Neonatal Early Onset	**Neonatal Late Onset**	**<12 months**	**1-4 years**	**5-9 years**	**10-17 years**	**p**
Species (%)	*Acinetobacter* spp.	1 (12.5)	3 (2.6)	6 (9.1)	9 (14.5)	4 (14.3)	4 (8.5)	<0.001
	E. coli	6 (75.0)	51 (44.0)	42 (63.6)	17 (27.4)	6 (21.4)	27 (57.4)	
	Enterobacter spp.	1 (12.5)	24 (20.7)	8 (12.1)	13 (21.0)	5 (17.9)	4 (8.5)	
	Citrobacter spp.		2 (1.7)		2 (3.2)	3 (10.7)	1 (2.1)	
	Klebsiella spp.		28 (24.1)	10 (15.2)	16 (25.8)	6 (21.4)	9 (19.1)	
	Serratia spp.		8 (6.9)				1 (2.1)	
	Aeromonas spp.				3 (4.8)	1 (3.6)		
	Proteus spp.				2 (3.2)			
	Pantoea spp.					3 (10.7)	1 (2.1)	
Age (days)	Median (IQR)	0.0 (0.0 to 1.2)	20.0 (11.8 to 49.2)	91.0 (57.5 to 177.5)	689.0 (512.2 to 969.8)	1893.0 (1770.0 to 2346.2)	5391.0 (3992.5 to 6168.5)	<0.001
Sex (%)	Female	4 (50.0)	38 (32.8)	28 (42.4)	28 (45.2)	9 (32.1)	19 (40.4)	0.51
	Male	4 (50.0)	78 (67.2)	38 (57.6)	34 (54.8)	19 (67.9)	28 (59.6)	

		Neonatal Early Onset	**Neonatal Late Onset**	**<12 months**	**1-4 years**	**5-9 years**	**10-17 years**	**p**
Specimen (%) Category	Nosocomial	0 (0.0)	86 (74.1)	7 (10.6)	18 (29.0)	0 (0.0)	15 (31.9)	<0.001
	Healthcare associated 0-30d	2 (25.0)	15 (12.9)	8 (12.1)	21 (33.9)	13 (46.4)	16 (34.0)	
	Healthcare associated 31-365d	0 (0.0)	3 (2.6)	29 (43.9)	11 (17.7)	4 (14.3)	5 (10.6)	
	Community	6 (75.0)	12 (10.3)	22 (33.3)	12 (19.4)	11 (39.3)	11 (23.4)	
Admission Length	Median (IQR)	14.2 (9.9 to 27.5)	51.7 (20.8 to 129.7)	4.1 (2.5 to 7.9)	8.8 (2.5 to 30.6)	3.1 (2.1 to 7.8)	5.5 (2.9 to 28.0)	<0.001
Death Within 30 Days (%)	No	6 (75.0)	86 (76.1)	44 (67.7)	52 (83.9)	21 (75.0)	39 (83.0)	0.30
	Yes	2 (25.0)	27 (23.9)	21 (32.3)	10 (16.1)	7 (25.0)	8 (17.0)	
CRP	Median (IQR)	8.2 (0.6 to 37.9)	64.1 (16.1 to 150.3)	99.2 (45.5 to 125.0)	50.1 (10.1 to 122.2)	16.7 (5.0 to 56.8)	59.2 (13.9 to 118.4)	0.08
Amoxicillin	S	5 (71.4)	27 (24.1)	25 (42.4)	11 (22.0)	5 (20.8)	5 (12.5)	0.002
(missing=35)	R	2 (28.6)	85 (75.9)	34 (57.6)	39 (78.0)	19 (79.2)	35 (87.5)	

		Neonatal Early Onset	Neonatal Late Onset	<12 months	1-4 years	5-9 years	10-17 years	p
Ceftriaxone	S	6 (85.7)	96 (85.0)	57 (93.4)	43 (81.1)	23 (85.2)	36 (85.7)	0.46
(missing=24)	R	1 (14.3)	17 (15.0)	4 (6.6)	10 (18.9)	4 (14.8)	6 (14.3)	
Meropenem	S	8 (100.0)	115 (100.0)	62 (100.0)	61 (100.0)	28 (100.0)	43 (97.7)	0.26
(missing=9)	R	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.3)	
Ciprofloxacin	S	7 (87.5)	110 (95.7)	62 (100.0)	55 (91.7)	27 (96.4)	39 (88.6)	0.05
(missing=10)	R	1 (12.5)	5 (4.3)	0 (0.0)	5 (8.3)	1 (3.6)	5 (11.4)	
Amikacin	S	3 (100.0)	47 (90.4)	15 (100.0)	21 (87.5)	17 (100.0)	21 (91.3)	0.63
(missing=193)	R	0 (0.0)	5 (9.6)	0 (0.0)	3 (12.5)	0 (0.0)	2 (8.7)	
Gentamicin	S	7 (87.5)	105 (92.1)	60 (96.8)	53 (86.9)	25 (92.6)	38 (86.4)	0.27
(missing=11)	R	1 (12.5)	9 (7.9)	2 (3.2)	8 (13.1)	2 (7.4)	6 (13.6)	
Piperacillin-Tazobactam	S	8 (100.0)	105 (92.1)	56 (93.3)	48 (85.7)	23 (88.5)	37 (86.0)	0.57
(missing=20)	R		9 (7.9)	4 (6.7)	8 (14.3)	3 (11.5)	6 (14.0)	

		Neonatal Early Onset	Neonatal Late Onset	<12 months	1-4 years	5-9 years	10-17 years	p
Fosfomycin	S	1 (100.0)	45 (93.8)	10 (90.9)	14 (87.5)	11 (84.6)	18 (94.7)	0.72
(missing=219)	R		3 (6.2)	1 (9.1)	2 (12.5)	2 (15.4)	1 (5.3)	