Epidemic potential of *Escherichia coli* ST131 and *Klebsiella pneumoniae* ST258: a systematic review and meta-analysis

M J D Dautzenberg,1,2 M R Haverkate,1 M J M Bonten,1,2 M C J Bootsma1,3

**ABSTRACT**

**Objectives:** Observational studies have suggested that *Escherichia coli* sequence type (ST) 131 and *Klebsiella pneumoniae* ST258 have hyperendemic properties. This would be obvious from continuously high incidence and/or prevalence of carriage or infection with these bacteria in specific patient populations. Hyperendemicity could result from increased transmissibility, longer duration of infectiousness, and/or higher pathogenic potential as compared with other lineages of the same species. The aim of our research is to quantitatively estimate these critical parameters for *E. coli* ST131 and *K. pneumoniae* ST258, in order to investigate whether *E. coli* ST131 and *K. pneumoniae* ST258 are truly hyperendemic clones.

**Primary outcome measures:** A systematic literature search was performed to assess the evidence of transmissibility, duration of infectiousness, and pathogenicity for *E. coli* ST131 and *K. pneumoniae* ST258. Meta-regression was performed to quantify these characteristics.

**Results:** The systematic literature search yielded 639 articles, of which 19 data sources provided information on transmissibility (*E. coli* ST131 n=9; *K. pneumoniae* ST258 n=10), 2 on duration of infectiousness (*E. coli* ST131 n=2), and 324 on pathogenicity (*E. coli* ST131 n=285; *K. pneumoniae* ST258 n=39). Available data on duration of carriage and on transmissibility were insufficient for quantitative assessment. In multivariable meta-regression *E. coli* isolates causing infection were associated with ST131, compared to isolates only causing colonisation, suggesting that *E. coli* ST131 can be considered more pathogenic than non-ST131 isolates. Date of isolation, location and resistance mechanism also influenced the prevalence of ST131. *E. coli* ST131 was 3.2 (95% CI 2.0 to 5.0) times more pathogenic than non-ST131. For *K. pneumoniae* ST258 there were not enough data for meta-regression assessing the influence of colonisation versus infection on ST258 prevalence.

**Conclusions:** With the currently available data, it cannot be confirmed nor rejected, that *E. coli* ST131 or *K. pneumoniae* ST258 are hyperendemic clones.

**INTRODUCTION**

Infections caused by *Escherichia coli* and *Klebsiella pneumoniae* producing extended-spectrum β-lactamases (ESBL) or carbapenemases, are increasing worldwide. There is growing evidence that certain clonal lineages of these species, such as *E. coli* sequence type (ST) 131 and *K. pneumoniae* ST258, have more epidemic potential than other lineages within their species group. *E. coli* ST131 was first described in 20081 and *K. pneumoniae* ST258 in 2009.2 *E. coli* ST131 is reported from around the globe, both in healthcare settings and in the community,3 4 and is mostly associated with ESBL production and fluoroquinolone resistance.3 5 6 8 9 *K. pneumoniae* ST258 is mainly associated with *K. pneumoniae* carbapenemase (KPC) production, and other resistance mechanisms,6 and is widespread in the USA, and expanding in Europe.6 8 In the scientific literature, *E. coli* ST131 and *K. pneumoniae* ST258 are widely considered hyperendemic clones.3 5 6 8 9 But the evidence underlying these assumptions is not that obvious.3 5 If *E. coli* ST131 or *K. pneumoniae* ST258 are truly hyperendemic clones, interventions may be targeted to these specific clones.

From a simple model in which patients can be susceptible, colonised or infected (figure 1), the characteristics of hyperendemicity follow as explained below. Susceptible hosts can acquire colonisation through
transmission, either directly (from another colonised or infected person) or indirectly (from the environment or via the hands of healthcare workers). Both colonised and infected patients contribute to transmission, as long as they are infectious, which can be expressed with the duration of colonisation. Duration of colonisation can be influenced by fitness cost associated with resistance or by antibiotic use. Colonisation can proceed to infection, which typically occurs in a fraction of colonised patients, and the rate of this progression can be expressed as the pathogenicity level. Decolonisation can occur in both colonised and infected persons.

To be hyperendemic, a clone has to have advantages over other clones in at least one of the traits: transmissibility, duration of colonisation or pathogenicity. Therefore, we define a hyperendemic clone as a clone that is more transmissible, has a longer duration of colonisation, and/or is more pathogenic than other clones of the same species. The presence of any or more of these traits will then lead to a continuously high incidence and/or prevalence of carriage or disease in a specific patient population. We performed a systematic review to quantitatively estimate these critical parameters for E. coli ST131 and K. pneumoniae ST258, in order to investigate whether E. coli ST131 and K. pneumoniae ST258 are truly hyperendemic clones.

**METHODS**

**Search strategy**

A PubMed and EMBASE search was performed to retrieve relevant articles published until 1 January 2015. The complete search string can be found in online supplementary text 1. A cross-reference check was performed to include relevant articles not found during the search. Only English, full-text articles were included. Articles unavailable online were requested from the authors. The Meta-analysis Of Observational Studies in Epidemiology statement was followed for reporting in this paper.

**Study selection**

Titles and abstracts were independently reviewed by two reviewers (MRH and MJDD) and selected for further review if they met the inclusion criteria. Selections were compared between the two reviewers, and if consensus was not reached, a third reviewer (MCJB or MJMB) was consulted.

The inclusion criteria for articles on transmissibility were that possible transmissions should be described, and the number of cases should be reported. Outbreak reports were included. Articles focusing on duration of colonisation should include at least two cultures per patient taken at two different time points. Pathogenicity was defined as the difference in the prevalence of ST131 or ST258 in infections (clinical isolates) compared to colonisation. We considered a clone to be more pathogenic when the relative abundance of this clone in isolates causing infections is higher compared to isolates associated with colonisation. Therefore, articles on pathogenicity of E. coli ST131 or K. pneumoniae ST258 should report the prevalence or incidence of infections among patients colonised with E. coli ST131 or K. pneumoniae ST258, the prevalence of E. coli ST131 or K. pneumoniae ST258 among patients colonised with E. coli or K. pneumoniae, respectively, or the prevalence of E. coli ST131 or K. pneumoniae ST258 among at least 10 clinical isolates of E. coli or K. pneumoniae, respectively.

Articles were excluded if they did not contain original data (such as reviews, commentaries, or articles reusing existing data sets), if they considered E. coli or K. pneumoniae only in non-human sources, or if there was no clear information on the isolate collection or selection.

**Data extraction**

Data were extracted by the same two reviewers independently, and crosschecked using a standard form developed by the researchers. Data were collected on population and setting, recording if participants were inpatients, outpatients/community residents, travellers or from another/unknown group. The area/region where the study took place was recorded and categorised into (mainly) from Africa, Asia, Australia, Europe, North America and South America. It was recorded whether data collection took place during an outbreak period, and if a selection on antibiotic susceptibility or resistance was made, divided into selection on ESBL/AmpC-producing isolates (including third-generation cephalosporin-resistant isolates), carbapenem-resistant or carbapenemase-producing Enterobacteriaceae (CRE/CPE, eg, KPC, OXA-48), other resistance profiles (eg, ciprofloxacin-resistant, fluoroquinolone-susceptible or multidrug resistant), or no selection on resistance. Furthermore, the method to detect sequence types was documented, split up into multilocus sequence typing...
(MLST, when all isolates were typed by MLST), extrapolation based on pulsed-field gel electrophoresis (PFGE, when only selected isolates were typed with MLST and the sequence types were inferred based on PFGE type), (PCR, when all isolated underwent PCR-screening for ST-specific alleles), extrapolation based on PCR (mainly MLST for E. coli isolates that were positive for O25b-ST131 by PCR), or other/unknown (such as fumC/fimH typing). Also, the sample site of the included isolates (percentage of isolates isolated from blood, urine, gastrointestinal, respiratory, wound/abscess or other sites), and time period of the study were recorded. For the time period, the middle date was used in the model if the study covered a longer time period.

For transmissibility, if available, information was gathered on admission prevalence, number of cases, number of uncolonised patients and transmission measure given. For duration of colonisation, the number of cases and duration of colonisation was recorded. For pathogenicity, information was collected on the prevalence or incidence of infections in patients colonised with E. coli ST131 or K. pneumoniae ST258, the prevalence of E. coli ST131 or K. pneumoniae ST258 in patients colonised with E. coli or K. pneumoniae, respectively, and/or the prevalence of E. coli ST131 or K. pneumoniae ST258 in patients infected with E. coli or K. pneumoniae, respectively.

Quality of the included articles was assured by only including papers with a proper selection of isolates. Furthermore, quality was implicitly incorporated in the data that were collected on the detection method used, the sample sites, whether data were collected during an outbreak and the setting and time period in which data were collected.

Several studies allowed splitting the data into multiple ‘data sources’. For example, if data was available per year or per country, these were recorded separately. Figure 2 shows a flow diagram with the included and excluded articles. Since only 19 data sources were available on transmissibility (9 on E. coli ST131 and 10 on K. pneumoniae ST258), and two on duration of colonisation (both on E. coli ST131), we could only describe these without quantifying summary measures. For pathogenicity, enough data was available on E. coli to do a meta-regression analysis and calculate summary measures.

Meta-regression pathogenicity

In order to evaluate the pathogenicity of E. coli ST131 and K. pneumoniae ST258, and to assess which factors influence this, meta-regression was performed using all reported data on the prevalence of E. coli ST131 in clinical (infection) or screening (colonisation) isolates of E. coli, and for all reported data on the prevalence of K. pneumoniae ST258 in clinical (infection) isolates of K. pneumoniae. The prevalence estimates (calculated as the number of ST131-positive or ST258-positive isolates divided by the total number of E. coli or K. pneumoniae isolates, respectively) and standard errors (SEs) were logit transformed in the analysis. Heterogeneity between studies was evaluated with Cochrans’s Q and the I² statistic. Because of high heterogeneity (I² >75%), a meta-analysis using a generalised linear

---

**Figure 2** Flow chart of article selection.
mixed-effect model with random effects per data source was used to assess sources of variability in the overall prevalence estimates. Univariate analyses were performed to identify covariates associated with the overall prevalence estimates. All covariates with a p value <0.20 were included in the multivariate model, and backward selection was performed using the likelihood ratio test. There, as we are performing an exploratory analysis, a cut-off of p<0.10 was used to determine statistical significance. The variable describing sample site was not included in the models, because of great dependency on the type of isolate (clinical or screening isolate, eg, blood isolates representing infection), and the effect of culture site, might not be comparable for isolates representing colonisation or infection. The estimated between-study variance ($\tau^2$) was evaluated for the model with and without explanatory parameters. The exponent of the coefficient for colonisation/infection found in the metaregression model is an OR, which can be interpreted as a risk ratio. This was taken as a measure of how much more pathogenic E. coli ST131 was compared to non-ST131, that is, a value of 2 would indicate that per colonised day colonisation with ST131 leads two times more often to an infection as compared to colonisation with non-ST131. All analyses were performed in R V3.0.3 (http://CRAN.R-project.org) using the ‘metafor’ package.

**RESULTS**

In all, 345 useful data sources were identified (see figure 2 for the consecutive steps followed for identification). For transmissibility, 19 data sources were identified; for duration of carriage, 2; and for pathogenicity, 324. Most studies (n=206, 72%) were performed in Europe and North America, and 266 (93%) were performed in a non-outbreak setting (table 1). E. coli isolates were most selected on ESBL production or resistance against third-generation cephalosporins, and K. pneumoniae isolates on being CRE/CPE. Colonisation isolates were most often from gastrointestinal origin (85.2%), and infection isolates from urine (54.8%) or blood (24.5%).

**Transmissibility**

There were 19 studies reporting transmissibility of E. coli ST131 (n=9) and K. pneumoniae ST258 (n=10), some being case reports or describing single possible transmission events (table 2). Transmission events for E. coli ST131 have been described or suggested in household (n=4), day care (n=1), nursing home (n=1) and hospital settings (n=4). For K. pneumoniae ST258 all sources reported on transmission events in hospital settings, and all included CRE/CPE.

Transmissibility can be quantified by the number of transmissions per patient, or patient-days at risk, which requires information on the number of index cases, number of transmissions, and number of days or patients at risk. Yet, one or more of these aspects, especially time at risk, is missing in all studies but one. Most studies are cross-sectional studies, in which transmission cannot be proven.

Differences in transmission capacity between E. coli ST131 and non-ST131, or between K. pneumoniae ST258 and non-ST258, have not been quantified, precluding any conclusion on the relative transmissibility of E. coli ST131 and K. pneumoniae ST258 compared to other clonal lineages.

**Duration of carriage**

The duration of carriage of E. coli ST131 was investigated in two studies. In one study, colonisation with E. coli was still apparent after 12 months in 64% (n=9), and 40% (n=14) of those carrying E. coli ST131 or other STs, respectively (p=0.12).32 In another study, of two patients acquiring colonisation with E. coli ST131 during travel, one was a prolonged carrier with this strain. However, the definition of prolonged carriage was not given.33

The duration of carriage of K. pneumoniae ST258 has not been determined.

**Pathogenicity**

E. coli

From 285 data sources, we retrieved data from 34 253 E. coli isolates (2041 associated with colonisation and 32 212 with infection). Prevalence of E. coli ST131 in these studies ranged from 0% to 100% (see online supplementary figure S1), with high statistical heterogeneity between studies (I²=96.9%).

In univariable meta-regression the E. coli ST131 prevalence in individual studies increased in time, and appeared to be influenced by whether isolates were associated with infection or colonisation, resistance patterns used for isolate selection and location, where the study was performed (p value <0.20; table 3). These variables were included in the multivariable meta-regression model, and time, location and selection remained significantly associated with E. coli ST131 prevalence (table 4). No significant effects were present for study population, microbiological methods used to detect ST131, or whether the study was performed in an outbreak situation or not.

The prevalence of ST131 was highest if E. coli isolates were selected upon the presence of ESBL production, or third-generation cephalosporin resistance, and lowest if derived from non-selective media. Prevalence of E. coli ST131 was highest in North America, and lowest in South America. The estimated prevalence of ST131 in E. coli, given particular values of the covariates, can be derived from the regression equation (table 4). For example, the estimated logit (prevalence ST131) for isolates causing infection, selected on presence of ESBL, in North America in January 2010 is given by 2.9668 +12×0.0140+1.1545+1.3826+0.4436=0.1819, which corresponds to a prevalence of ST131 of exp(0.1819)/(1+exp(0.1819))=54.5%. The estimated prevalence in the
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>EC transmissibility (n=9)</th>
<th>KP transmissibility (n=10)</th>
<th>EC duration (n=2)</th>
<th>EC pathogenicity colonisation (n=35)</th>
<th>EC pathogenicity infection (n=249)</th>
<th>KP pathogenicity colonisation (n=3)</th>
<th>KP pathogenicity infection (n=35)</th>
<th>KP pathogenicity colonisation and infection (n=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of isolates (mean, SD)</td>
<td></td>
<td></td>
<td></td>
<td>58 (67)</td>
<td>129 (357)</td>
<td>59 (69)</td>
<td>40 (64)</td>
<td></td>
</tr>
<tr>
<td>Number of isolates (median, IQR)</td>
<td></td>
<td></td>
<td></td>
<td>36 (21–62)</td>
<td>53 (20–115)</td>
<td>36 (20–87)</td>
<td>20 (14–41)</td>
<td></td>
</tr>
<tr>
<td>Population—outpatients</td>
<td>2 (22.2%)</td>
<td>8 (80.0%)</td>
<td>1 (50.0%)</td>
<td>11 (31.4%)</td>
<td>128 (51.4%)</td>
<td>3 (100.0%)</td>
<td>24 (68.6%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—outpatients/community</td>
<td>6 (66.7%)</td>
<td>2 (20.0%)</td>
<td>0 (0.0%)</td>
<td>18 (51.4%)</td>
<td>25 (10.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—inpatients</td>
<td></td>
<td></td>
<td></td>
<td>2 (5.7%)</td>
<td>63 (25.3%)</td>
<td>0 (0.0%)</td>
<td>2 (5.7%)</td>
<td>1 (100.0%)</td>
</tr>
<tr>
<td>Population—inpatients travellers</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>1 (50.0%)</td>
<td>3 (8.6%)</td>
<td>3 (1.2%)</td>
<td>0 (0.0%)</td>
<td>1 (2.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—other/unknown</td>
<td>2 (22.2%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>9 (25.7%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—Asia</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>2 (5.7%)</td>
<td>16 (6.4%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—Europe</td>
<td>4 (44.4%)</td>
<td>7 (70.0%)</td>
<td>1 (50.0%)</td>
<td>23 (65.7%)</td>
<td>182 (73.1%)</td>
<td>4 (23.5%)</td>
<td>1 (66.7%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—North America</td>
<td>3 (33.3%)</td>
<td>7 (70.0%)</td>
<td>1 (50.0%)</td>
<td>14 (40.0%)</td>
<td>96 (38.6%)</td>
<td>1 (33.3%)</td>
<td>14 (40.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Population—South America</td>
<td>0 (0.0%)</td>
<td>2 (20.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>9 (10.0%)</td>
<td>6 (4.0%)</td>
<td>1 (33.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Outbreak setting</td>
<td>3 (33.3%)</td>
<td>10 (100.0%)</td>
<td>0 (0.0%)</td>
<td>1 (2.9%)</td>
<td>16 (5.0%)</td>
<td>1 (33.3%)</td>
<td>2 (22.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Selection—ESBL/3GC-R</td>
<td>8 (88.9%)</td>
<td>0 (0.0%)</td>
<td>1 (50.0%)</td>
<td>23 (65.7%)</td>
<td>182 (73.1%)</td>
<td>2 (66.7%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Selection—CRE/CPE</td>
<td>0 (0.0%)</td>
<td>9 (90.0%)</td>
<td>0 (0.0%)</td>
<td>8 (32.5%)</td>
<td>8 (32.5%)</td>
<td>1 (33.3%)</td>
<td>29 (82.9%)</td>
<td>1 (100.0%)</td>
</tr>
<tr>
<td>Selection—other</td>
<td>1 (11.1%)</td>
<td>0 (0.0%)</td>
<td>1 (50.0%)</td>
<td>5 (14.3%)</td>
<td>31 (12.4%)</td>
<td>0 (0.0%)</td>
<td>5 (14.3%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Selection—none</td>
<td>0 (0.0%)</td>
<td>1 (10.0%)</td>
<td>0 (0.0%)</td>
<td>7 (20.0%)</td>
<td>28 (11.2%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Selection—MLST</td>
<td>6 (66.7%)</td>
<td>4 (40.0%)</td>
<td>0 (0.0%)</td>
<td>10 (28.6%)</td>
<td>134 (53.8%)</td>
<td>1 (33.3%)</td>
<td>25 (71.4%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Detection—extrapolation based on PFGE</td>
<td>1 (11.1%)</td>
<td>3 (30.0%)</td>
<td>0 (0.0%)</td>
<td>3 (8.6%)</td>
<td>15 (6.0%)</td>
<td>1 (33.3%)</td>
<td>9 (25.7%)</td>
<td>1 (100.0%)</td>
</tr>
<tr>
<td>Detection—extrapolation based on PCR</td>
<td>2 (22.2%)</td>
<td>0 (0.0%)</td>
<td>2 (100.0%)</td>
<td>21 (60.0%)</td>
<td>83 (33.3%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Detection—CH</td>
<td>0 (0.0%)</td>
<td>1 (10.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>13 (5.2%)</td>
<td>1 (33.3%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Detection—other/unknown</td>
<td>0 (0.0%)</td>
<td>2 (20.0%)</td>
<td>0 (0.0%)</td>
<td>1 (2.9%)</td>
<td>4 (1.6%)</td>
<td>0 (0.0%)</td>
<td>1 (2.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—blood</td>
<td>1 (11.1%)</td>
<td>3 (30.0%)</td>
<td>0 (0.0%)</td>
<td>64 (25.7%)</td>
<td>0 (0.0%)</td>
<td>7 (20.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—urine</td>
<td>2 (22.2%)</td>
<td>3 (30.0%)</td>
<td>1 (50.0%)</td>
<td>143 (57.4%)</td>
<td>1 (33.3%)</td>
<td>12 (34.3%)</td>
<td>1 (100.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—gastrointestinal tract</td>
<td>6 (66.7%)</td>
<td>3 (30.0%)</td>
<td>1 (50.0%)</td>
<td>32 (91.4%)</td>
<td>5 (2.0%)</td>
<td>1 (33.3%)</td>
<td>7 (20.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—respiratory tract</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>3 (12.5%)</td>
<td>3 (12.5%)</td>
<td>1 (33.3%)</td>
<td>3 (8.6%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—wound</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>1 (4.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Site—other/unknown</td>
<td>0 (0.0%)</td>
<td>1 (10.0%)</td>
<td>0 (0.0%)</td>
<td>33 (13.3%)</td>
<td>0 (0.0%)</td>
<td>6 (17.1%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>

CH, fumC/fimH typing; CPE, carbapenemase-producing Enterobacteriaceae; CRE, carbapenem-resistant Enterobacteriaceae; EC, *Escherichia coli*; ESBL, extended-spectrum beta-lactamase; KP, *Klebsiella pneumoniae*; KPC, *Klebsiella pneumoniae* carbapenemase; MLST, multilocus sequence typing; PFGE, pulsed-field gel electrophoresis; Site, site from which most isolates were identified.
Table 2  Summary of articles describing transmissibility of *Escherichia coli* ST131 and *Klebsiella pneumoniae* ST258

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Country</th>
<th>Year</th>
<th>Setting</th>
<th>Organism</th>
<th>Resistance mechanism</th>
<th>Index cases (n)</th>
<th>Secondary cases (n)</th>
<th>Uncolonised</th>
<th>Exposure time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veenemans (2014)</td>
<td>The Netherlands</td>
<td>2013</td>
<td>Nursing homes</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>5 and 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kojima (2014)</td>
<td>Japan</td>
<td>2009–2010</td>
<td>Household</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blanc (2014)</td>
<td>France</td>
<td>2012</td>
<td>Day care centers</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giuffrè (2013)</td>
<td>Italy</td>
<td>2012</td>
<td>Neonatal intensive care unit</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>15</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Veenemans (2014)</td>
<td>Switzerland</td>
<td>2008–2010</td>
<td>University hospital</td>
<td><em>E. coli</em> non-ST131</td>
<td>ESBL</td>
<td>31</td>
<td>36</td>
<td>367</td>
<td>48 index inpatients for a total of 400 000 patient-days</td>
</tr>
<tr>
<td>Owens (2011)</td>
<td>USA</td>
<td>Before 2011</td>
<td>Household</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>15</td>
<td>7</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Johnson (2010)</td>
<td>USA</td>
<td>Before 2010</td>
<td>Household</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>42</td>
<td>13</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Ender (2009)</td>
<td>USA</td>
<td>Before 2009</td>
<td>Hospital Intensive care unit</td>
<td><em>E. coli</em> ST131</td>
<td>Fluoro-quinolone resistance</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Marquez (2014)</td>
<td>Uruguay</td>
<td>2011</td>
<td>Hospital Intensive care unit</td>
<td><em>E. coli</em> ST131</td>
<td>ESBL</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Gaibani (2014)</td>
<td>Italy</td>
<td>2010</td>
<td>Hospital</td>
<td><em>K. pneumoniae</em> ST258</td>
<td>KPC</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giuffrè (2013)</td>
<td>Italy</td>
<td>2012</td>
<td>Neonatal intensive care unit</td>
<td><em>K. pneumoniae</em> ST258</td>
<td>KPC</td>
<td>10</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agodi (2011)</td>
<td>Italy</td>
<td>2009</td>
<td>Hospital Acute care hospitals Intensive care unit</td>
<td><em>K. pneumoniae</em> ST258</td>
<td>KPC</td>
<td>16</td>
<td></td>
<td>16 (+7 presumed cases)</td>
<td></td>
</tr>
<tr>
<td>Won (2011)</td>
<td>USA</td>
<td>2008</td>
<td>Hospital Acute care hospitals and long-term acute care hospitals</td>
<td><em>K. pneumoniae</em> ST258</td>
<td>KPC</td>
<td>33</td>
<td></td>
<td>33 (+7 presumed cases)</td>
<td></td>
</tr>
</tbody>
</table>
The estimated between-study variance ($\tau^2$) reduced from 1.68 in the model without parameters to 1.1 in the final model, implying that a high level of heterogeneity remained.

### Table 3: Effect of covariates on prevalence of ST131 in *Escherichia coli* (univariable random effects meta-regression models)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Estimate (SE)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study period (per month*)</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>Infection or colonisation</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Colonisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outbreak setting</td>
<td>0.9112</td>
<td></td>
</tr>
<tr>
<td>Selection of isolates based on resistance pattern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No selection on resistance profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESBL/3GC-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE/CPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.6219</td>
<td></td>
</tr>
<tr>
<td>Study population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inpatients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatients/community</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travellers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other/unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method used to detect ST131</td>
<td>0.3598</td>
<td></td>
</tr>
<tr>
<td>MLST</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extrapolation based on PFGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extrapolation based on PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other/unknown</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Parameter estimates (SEs) are presented on a logit scale. Reference date: 1 January 2009.

### Table 4: Effect of covariates on prevalence of ST131 in *Escherichia coli* (multivariable random effects meta-regression model)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Estimate (SE)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>$-2.9668$ (0.2959)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Study period (per month*)</td>
<td>0.0140 (0.0023)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Infection or colonisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colonisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infection</td>
<td>1.1545 (0.2281)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Selection of isolates based on resistance pattern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No selection on resistance profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESBL/3GC-R</td>
<td>1.3826 (0.2207)</td>
<td></td>
</tr>
<tr>
<td>CRE/CPE</td>
<td>0.5994 (0.4879)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.9058 (0.2709)</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>0.4436 (0.1675)</td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td>$-2.2868$ (0.6101)</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>$-0.4209$ (0.3407)</td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>$-0.3657$ (0.1927)</td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>$-0.2246$ (0.3154)</td>
<td></td>
</tr>
</tbody>
</table>

*Parameter estimates (SEs) are presented on a logit scale. Reference date: 1 January 2009.

### K. pneumoniae

There were 35 and three data sources providing information on the prevalence of ST258 *K. pneumoniae* in clinical and colonising isolates, respectively (see online supplementary figure S3). Because of limited data on colonisation, quantitative analyses were performed for clinical isolates only.

In the univariable meta-regression model, outbreak setting yes/no, selection of isolates based on resistance pattern, study population and geographic location were all associated with a higher prevalence of ST258 with a p value <0.20 and were, thus, included in the multivariable model (table 5). If data were collected during an outbreak of *K. pneumoniae*, this was associated with a higher prevalence of ST258 (table 6). Furthermore, the model yielded a significant effect of resistance patterns on the prevalence of ST258 in *K. pneumoniae*. ST258 prevalence was associated with selection of isolates on CRE-positivity, but the number of data sources describing isolates that are not CRE/CPE is low and varied (n=5). Furthermore, study population characteristics also appeared to influence ST258 prevalence in *K. pneumoniae*, with higher prevalence of ST258 in inpatients, compared to ‘other’ populations. Yet, the ‘other’ group is not defined accurately, precluding firm conclusions. Only one data source was available for outpatients or persons residing in the community. Finally, the reported ST258 prevalence was lower in Asia and Australia than in other continents.
The estimated prevalence of ST258 in K. pneumoniae, given particular values of the covariates, can be derived from the regression equation. For example, the estimated logit (prevalence of ST258) for isolates selected on presence of CRE in hospital inpatients in North America during an outbreak is given by $-0.0.0320 + 2.8038 + 0.3332 = 3.1050$, which corresponds to a prevalence of ST258 of $\exp(3.1050)/(1+\exp(3.1050))=95.7\%$. The estimated prevalence in the reference category (during an outbreak, non CRE/CPE, hospital inpatients, Europe) is $\exp(-0.0320)/(1+\exp(-0.0320))=50.8\%$.

The estimated between-study variance ($\tau^2$) reduced from 6.43 in the model without parameters to 2.25 in the final model, indicating a considerable improvement, but still a high level of heterogeneity.

ST258 was not detected in two studies reporting on colonisation with K. pneumoniae, that included 36 and 4 isolates, respectively. Only from the study of van Duin et al. can we deduce a prevalence of ST258 in K. pneumoniae of 31% in colonising isolates. This precludes any quantification of the pathogenicity of K. pneumoniae ST258.

The only study in which both colonisation and infection with K. pneumoniae ST258 were investigated included a set of seven KPC-producing K. pneumoniae ST258 isolates collected from a long-term acute-care facility in South Florida. Three patients were colonised, and four had both colonisation and infection. Again, the sample size is too small for drawing conclusions.

**DISCUSSION**

Based on published information, we conclude that there is evidence that E. coli ST131 is more pathogenic than E. coli non-ST131, but not for increased transmissibility or prolonged duration of carriage. Because of the heterogeneity in the data, it cannot be concluded (nor rejected) that E. coli ST131 is a hyperendemic clone. For K. pneumoniae ST258, the published data precluded any conclusion on increased transmissibility, longer duration of carriage or increased pathogenicity.

Several limitations in our study should be acknowledged. Because of our search strategy, the prevalence of E. coli ST131 and K. pneumoniae ST258 that were retrieved are likely overestimations of the real prevalence. We required the articles to report ST131/ST258 in their title and/or abstract, and therefore, articles that did not report this, or that did not detect ST131/ST258 in their study, may have been missed. Since the prevalence is dependent on factors including time, location, resistance pattern, population studied and possibly variables not included in this review (e.g. patient-specific details like age, gender), we deemed it not meaningful to estimate an overall prevalence of ST131 in E. coli or ST258 in K. pneumoniae.

We also did not create a funnel plot to assess publication bias, as such an analysis also assumes that there is one overall effect or prevalence. Thus, publication bias cannot be excluded. It is possible that identification of E. coli ST131 or K. pneumoniae ST258 stimulates publication because of the current interest in these clones. However, this will most likely equally influence studies reporting infection and colonisation isolates, which...
would not influence our conclusions. Also, the finding of ESBL or KPC might instigate investigation of sequence types. As 70% of the included studies on E. coli selected isolates based on the presence of ESBL or 3GC-R, our findings might be more applicable to ESBL-producing E. coli ST131 than all E. coli ST131 in general. The same holds for K. pneumoniae, for which around 90% of included studies selected isolates based on the presence of carbapenemase production of carbapenem resistance, mainly corresponding to KPC production. In our analysis, we used grouped variables (eg, continent instead of country), as there are limitations to the number of variables that can be studied.

There could also be differences in detecting infection and colonisation-associated isolates. Infection isolates are mainly collected retrospectively, when a pattern or outbreak is recognised, whereas, colonisation isolates are more often collected prospectively. Yet, since determination of sequence types is unambiguous, it is unlikely that such differences have affected our conclusions.

Our analysis clearly demonstrates that more—and better designed—studies are needed to determine whether E. coli ST131 and K. pneumoniae ST258 are truly hyperendemic clones. This would be possible with a prospective cohort study of a population (eg, the general population or hospitalised patients) with a certain contact structure, in which carriage with E. coli or K. pneumoniae is regularly (eg, weekly or monthly) determined. As K. pneumoniae ST258 is mainly a healthcare-associated pathogen, choice of study population might be different than for E. coli ST131, that is also a community-associated pathogen. For determination of transmissibility, genotyping should be performed, preferably with highly discriminatory methods, and preferably with inclusion of multiple isolates per patient. The duration of exposure to persons colonised or infected with E. coli ST131/K. pneumoniae ST258 should be determined to calculate the number of acquisitions per unit of time. Carriers could be studied in more detail to determine the duration of carriage and the infection rate (and duration until infection), preferably with inclusion of the effects of antibiotic use on these parameters. There should be a sufficient duration of follow-up, and isolates should be characterised to determine whether multiple isolates represent persistent carriage or re-colonisation with different strains.

In conclusion, current evidence does not allow the conclusion that E. coli ST131 and K. pneumoniae ST258 are hyperendemic clones.

Contributors MJDD and MRH performed the systematic literature search, reviewed and summarised data from each selected article, performed the analyses and wrote the first draft of the manuscript. MJDD, MRH, MJMB and MCJB all revised the manuscript.

Funding European Community (RGNOSIS Integrated project [FP7/2007-2013] under grant agreement no. 282512 to MRH, MJMB and MCJB; Netherlands Organization of Scientific Research (VICI NWO Grant 918.76.611 to MJMB and Priority Medicines Antimicrobial Resistance grant 205100013 to MRH and MCJB).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES


Epidemic potential of *Escherichia coli* ST131 and *Klebsiella pneumoniae* ST258: a systematic review and meta-analysis

M J D Dautzenberg, M R Haverkate, M J M Bonten and M C J Bootsma

*BMJ Open* 2016 6:
doi: 10.1136/bmjopen-2015-009971

Updated information and services can be found at:
http://bmjopen.bmj.com/content/6/3/e009971

These include:

**References**
This article cites 246 articles, 113 of which you can access for free at:
http://bmjopen.bmj.com/content/6/3/e009971#BIBL

**Open Access**
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

*Epidemiology* (2054)
*Infectious diseases* (551)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/