



**HAL**  
open science

## **A nonlinear time-series analysis approach to identify thresholds in associations between population antibiotic use and rates of resistance**

José-María López-Lozano, Timothy Lawes, César Nebot, Arielle P. Beyaert, Xavier Bertrand, Didier Hocquet, Mamoon Aldeyab, Michael Scott, Geraldine Conlon-Bingham, David Farren, et al.

### ► To cite this version:

José-María López-Lozano, Timothy Lawes, César Nebot, Arielle P. Beyaert, Xavier Bertrand, et al.. A nonlinear time-series analysis approach to identify thresholds in associations between population antibiotic use and rates of resistance. *Nature Microbiology*, 2019, 4 (7), pp.1160-1172. 10.1038/s41564-019-0410-0 . hal-02315695

**HAL Id: hal-02315695**

**<https://hal.science/hal-02315695v1>**

Submitted on 6 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## A non-linear time-series analysis approach to identify thresholds in associations between population antibiotic use and rates of resistance

THRESHOLDS study group (2019). A non-linear time-series analysis approach to identify thresholds in associations between population antibiotic use and rates of resistance. *Nature Microbiology*, 4(7), 1160-1172. <https://doi.org/10.1038/s41564-019-0410-0>

[Link to publication record in Ulster University Research Portal](#)

**Published in:**  
Nature Microbiology

**Publication Status:**  
Published (in print/issue): 01/07/2019

**DOI:**  
[10.1038/s41564-019-0410-0](https://doi.org/10.1038/s41564-019-0410-0)

**Document Version**  
Author Accepted version

**General rights**  
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**  
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [pure-support@ulster.ac.uk](mailto:pure-support@ulster.ac.uk).

1 **A non-linear time-series analysis approach to identify thresholds in**  
2 **associations between population antibiotic use and rates of resistance**

3  
4 José-María López-Lozano<sup>1</sup>, Timothy Lawes<sup>2\*</sup>, César Nebot<sup>3</sup>, Arielle Beyaert<sup>4</sup>, Xavier  
5 Bertrand<sup>5</sup>, Didier Hocquet<sup>5</sup>, Mamoon Aldeyab<sup>6</sup>, Michael Scott<sup>7</sup>, Geraldine Conlon-Bingham<sup>7</sup>,  
6 David Farren<sup>8</sup>, Gábor Kardos<sup>9</sup>, Adina Fésűs<sup>10</sup>, Jesús Rodríguez Baño<sup>11</sup>, Pilar Retamar<sup>11</sup>,  
7 Nieves Gonzalo-Jiménez<sup>12</sup>, Ian M Gould<sup>13</sup>, the THRESHOLDS study group<sup>14</sup>  
8  
9

10 <sup>1</sup> Medicine Preventive-Infection Control Team, Hospital Vega Baja, Orihuela-Alicante,  
11 Spain.

12 <sup>2</sup> The Wellcome Trust Liverpool-Glasgow Centre for Global Health Research, Liverpool,  
13 UK.

14 <sup>3</sup> Centro Universitario de la Defensa (CUD) de San Javier, Murcia, Spain.

15 <sup>4</sup> Departamento de Métodos Cuantitativos para la Economía y la Empresa, University of  
16 Murcia, Spain

17 <sup>5</sup> Centre hospitalier régional et universitaire (CHRU) and UMR 6249 Chrono-environnement,  
18 Université de Bourgogne-Franche-Comté, Besançon, France

19 <sup>6</sup> School of Pharmacy and Pharmaceutical Science, Ulster University, Northern Ireland, UK

20 <sup>7</sup> Pharmacy Department, Northern Health and Social Care Trust and Regional Medicines  
21 Optimisation Innovation Centre, Antrim, Northern Ireland, UK.

22 <sup>8</sup> Department of Medical Microbiology, Antrim Area Hospital, Antrim, Northern Ireland, UK

23 <sup>9</sup> Department of Medical Microbiology, Faculty of Medicine, University of Debrecen,  
24 Debrecen, Hungary

25 <sup>10</sup> Clinical Pharmacy, University of Debrecen, Debrecen, Hungary

26 <sup>11</sup> Infectious Diseases and Clinical Microbiology Unit, Instituto de Biomedicina de Sevilla  
27 (IBiS), Hospital Universitario Virgen Macarena and University of Seville, Seville, Spain.

28 <sup>12</sup> Medical Microbiology Department, Hospital Vega Baja, Orihuela-Alicante, Spain.

29 <sup>13</sup> Medical Microbiology Department, Aberdeen Royal Infirmary, Aberdeen, Scotland, UK.

30 <sup>14</sup> A full list of authors can be found at the end of the article  
31

32 \*Corresponding author:

33 E-mail: t.lawes@nhs.net (TL)  
34

35 Short title: Population antibiotic use thresholds  
36

37 Word count:

38 Abstract: 200

39 Text: 3157 (Introduction 478; Results 1863, Discussion 816)

40 Figs in manuscript: 5

41 Tables in manuscript 1

42 Extended data file including 3 figs, 6 tables  
43  
44  
45  
46  
47  
48  
49

50 **Abstract**

51

52 Balancing access to antibiotics with control of antibiotic resistance is a global public health  
53 priority. Currently, antibiotic stewardship is informed by a ‘use it and lose it’ principle, in  
54 which population antibiotic use is linearly related to resistance rates. However, theoretical  
55 and mathematical models suggest use-resistance relationships are non-linear. One explanation  
56 is that resistance genes are commonly associated with ‘fitness costs’, impairing pathogen  
57 replication or transmissibility. Therefore, resistant genes and pathogens may only gain a  
58 survival advantage where antibiotic selection pressures exceed critical thresholds. These  
59 thresholds may provide quantitative targets for stewardship: optimising control of resistance  
60 while avoiding over-restriction of antibiotics. We evaluated the generalisability of a non-  
61 linear time-series analysis approach for identifying thresholds using historical prescribing and  
62 microbiological data from five populations in Europe. We identified minimum thresholds in  
63 temporal relationships between use of selected antibiotics and rates of carbapenem-resistant  
64 *Acinetobacter baumannii* (in Hungary), extended spectrum  $\beta$ -lactamase producing  
65 *Escherichia coli* (Spain), cefepime-resistant *Escherichia coli* (Spain), gentamicin-resistant  
66 *Pseudomonas aeruginosa* (France), and methicillin-resistant *Staphylococcus aureus*  
67 (Northern Ireland) in different epidemiological phases. Using routinely generated data, our  
68 approach can identify context-specific quantitative targets for rationalising population  
69 antibiotic use and controlling resistance. Prospective intervention studies restricting antibiotic  
70 consumption are needed to validate thresholds.

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

## Introduction

100 Antimicrobials have facilitated improvements in health and food-security worldwide.  
101 However, excessive use has promoted antimicrobial resistance (AMR), undermining many  
102 aspects of healthcare and generating new disease burdens.<sup>1-4</sup> While decisions around  
103 antimicrobial use are made by individual prescribers and patients, consequences of over-use  
104 are only fully apparent across populations.<sup>5</sup> Competing needs and non-exclusivity lead to a  
105 ‘tragedy of the commons’.<sup>1,2</sup> The Global Action Plan on AMR therefore emphasises cross-  
106 societal efforts to conserve the current stock of effective antimicrobials.<sup>6</sup> Antimicrobial  
107 stewardship aims to control AMR by reducing inappropriate use, while ensuring access to  
108 effective therapy.<sup>2</sup> There is growing evidence that reducing population use of antimicrobials  
109 selecting for resistance, can effectively reduce AMR.<sup>7</sup> However, over-restriction may be  
110 counter-productive: use of alternative antimicrobials may increase in response, leading to  
111 new selection pressures or resistance problems.<sup>8</sup> Balancing the benefits and risks of  
112 antimicrobials in healthcare requires understanding of their effects on AMR at population  
113 scales.<sup>9-11</sup>

114  
115 Most empirical evidence has considered linear use-resistance relationships.<sup>12-16</sup> Although  
116 useful for identifying the most salient selection pressures in a population,<sup>12</sup> theoretical and  
117 mathematical models suggest non-linear relationships are more likely.<sup>9-11,17</sup> In particular,  
118 selection pressures may only impact on resistance rates, where volumes of antibiotic use in  
119 the population are above minimum thresholds. Potential explanations for minimum  
120 thresholds, include: antibiotic substitution with changes in use,<sup>8</sup> associations between the  
121 strength of selection pressure and induction of co-resistance,<sup>18</sup> differential effects on sub-  
122 populations within bacterial species,<sup>19</sup> or ecological niches,<sup>20</sup> and the ‘total use thresholds’  
123 hypothesis.<sup>9</sup> The latter concept arises from the observation that resistance genes are  
124 commonly associated with fitness costs impairing pathogen replication or transmissibility.<sup>21</sup>  
125 Resistant pathogens may therefore only gain a survival advantage over susceptible organisms,  
126 where selection pressures from antibiotic use outweigh fitness costs.<sup>9</sup> If identified, minimum  
127 thresholds may offer important targets for antibiotic stewardship, indicating an upper limit of  
128 safe population antibiotic use which does not increase resistance within populations.

129  
130 Non-linear relationships between fluoroquinolone use and resistant *E.coli* and *Staphylococcus*  
131 *aureus* have been reported.<sup>20,22</sup> We have shown that reducing antibiotic use to below  
132 minimum thresholds predicted shifts in the molecular and clinical epidemiology of *S. aureus*  
133 and *Clostridium difficile*.<sup>23-25</sup> However, early findings suggest antibiotic use-resistance  
134 relationships are likely to depend upon the clinical population,<sup>9,23,25</sup> pathogen genotypes,<sup>24</sup>  
135 intensity of infection control,<sup>23</sup> and transmission dynamics.<sup>11</sup> Moreover, fitness costs  
136 associated with different resistance genes are variable.<sup>21</sup> If fitness costs are low, or attenuated  
137 through epistasis<sup>26</sup> or compensatory mutations,<sup>27</sup> resistance may persist despite removing  
138 previously important antibiotic selection pressures.<sup>28</sup> A generalisable method for finding  
139 minimum thresholds should therefore provide context-specific results, have relatively low-  
140 informational demands, and allow for iteration.

141  
142 Here, we examined the generalisability of a non-linear time-series analysis methodology for  
143 identifying minimum threshold in use-resistance associations. Using routinely generated data  
144 from five European centres, we show this pragmatic approach can provide population-  
145 specific quantitative targets for antimicrobial stewardship for a variety of resistance  
146 outcomes, epidemiological phases, transmission dynamics, and clinical populations.

## Results

### Identifying non-linear temporal relationships: from experiment to application

In a Monte Carlo experiment we compared the ability of linear and non-linear time-series analysis (Multivariate Adaptive Regression Splines, MARS) to identify pre-defined relationships between simulated explanatory and outcome time-series (Supplementary Figure 1). Non-linear time-series analysis (NL-TSA) accurately identified both truly linear and non-linear associations. However, linear time-series analysis provided biased estimations and overall poorer data-fit if relationships were non-linear.

NL-TSA models applied to retrospective time-series data from five European study populations (examples 1-5), frequently identified minimum thresholds in antibiotic use-resistance relationships, (figures 1-5 and Supplementary Table 1). ‘Ceiling effects’, in which further increases in explanatory variables did not affect resistance rates, were found at high-levels of use of some antibiotics and hand hygiene. Non-linearities in autoregression and population interaction terms further indicated the complexity of transmission dynamics within and between clinical populations.

### Example 1: Carbapenem-resistant *Acinetobacter baumannii* (Debrecen, Hungary)

We examined ecological determinants of carbapenem-resistant *A. baumannii* (CRAb) in a tertiary hospital population in Debrecen, Hungary (figure 1). Between Oct 2004 and Aug 2016 (n=143 months), incidence density of CRAb increased from a 12-month average of 0.14 to 9.43 cases per 10 000 OBDs, while that of carbapenem-susceptible *A. baumannii* (CSAb) remained relatively constant. There were no planned antibiotic stewardship interventions in the study period. We observed increasing use of broad-spectrum antibiotics, including a tripling in carbapenem use, and more recent escalation in colistin use.

Carbapenem resistance in this setting was conferred predominantly by *bla*<sub>OXA-23</sub>-like carbapenemases, while *bla*<sub>OXA-24</sub>-like genes occurred sporadically.<sup>29</sup> CRAb were significantly more likely to be resistant to ciprofloxacin, gentamicin, amikacin, piperacillin-tazobactam and ceftazidime than CSAb (Supplementary Table 2), although susceptibility testing for the latter two agents was discontinued in 2013 as recommended by EUCAST. Vector autoregression models found that colistin use followed, rather than predicted, variation in CRAb incidence density. Use of carbapenems, fluoroquinolones, piperacillin-tazobactam, third generation cephalosporins (3GC), and aminoglycosides were considered potential explanatory variables.

Previous CRAb incidence density and recent hospital use of carbapenems, piperacillin-tazobactam, and fluoroquinolones were explanatory variables in the best-fit ( $R^2=0.86$ ) non-linear TSA model (Supplementary Table 1). In an almost identical model with poorer trade-off of data fit and model complexity (higher Modified Generalised Cross Validation statistic), fluoroquinolone use was replaced by the effect of 3GC use above a threshold of 36 (95% CI, 30 to 41) DDDs per 1000 OBDs (coefficient, 95% CI: 0.111, 0.018 to 0.203;  $p=0.019$ ; lag 3).

In this setting increases in CRAb added to, rather than replaced, CSAb, suggesting CRAb occupied new ecological niches. Strong autoregression in the CRAb time-series was consistent with substantial within-hospital transmission.<sup>29</sup> CRAb incidence density increased when population use of carbapenems, piperacillin-tazobactam, fluoroquinolones, and 3GC exceeded minimum thresholds. By the end of the study period, use of fluoroquinolones had reduced to below threshold. However, CRAb could be further controlled by reducing use of

198 carbapenems, 3GC, and piperacillin-tazobactam, from present levels to respective thresholds  
199 (table 1).

200 **Example 2: Extended spectrum  $\beta$ -lactamase producing *E. coli* (Orihuela, Spain)**

201 We examined variables temporal associated with the percentage of *E. coli* producing ESBL  
202 (%Ec-ESBL+) in connected district general hospital and community populations of Orihuela,  
203 Spain, between Jul 1991 and Oct 2016 (n=304 months, figure 2). Limited outbreaks of  
204 ESBL-producing *E. coli* were noted from 1998, but from 2002 the %Ec-ESBL+ increased  
205 rapidly alongside escalating use of fluoroquinolones and co-amoxiclav. While use of these  
206 agents later stabilised or declined, hospital use of third-generation cephalosporins continued  
207 to increase.

208  
209 Over the study period,  $bla_{CTX-M}$  genes were common in *E. coli* across Spain, with  
210 dissemination of the CTX-M-15-producing O25b-ST131 clone and clonally unrelated CTX-  
211 M-14-producing strains identified.<sup>30,31</sup> Acquisition of fluoroquinolone resistance was a key  
212 step in the evolution of dominant  $bla_{CTX-M-15}$  containing sub-clones of O25b-ST131.<sup>32</sup>  
213 Consistent with this, 81% of ESBL-producing *E. coli* in Orihuela were non-susceptible to  
214 ciprofloxacin. They were also significantly more likely to be resistant to co-trimoxazole, co-  
215 amoxiclav and aminoglycosides compared to non-ESBL *E. coli* (Supplemental Table 3).  
216 Vector autoregression models demonstrated bidirectional interactions between community  
217 and hospital %Ec-ESBL+ and that use of piperacillin-tazobactam and carbapenems followed,  
218 rather than predicted, changes in %Ec-ESBL+. In separate hospital and community models,  
219 %Ec-ESBL in the other population was considered as an explanatory variable.

220  
221 In the best-fit model ( $R^2=0.62$ ) hospital %Ec-ESBL+ was predicted by prior %EcESBL in  
222 hospital and community, and hospital use of 3GC, and fluoroquinolones exceeding minimum  
223 thresholds (Supplementary Table 1). A potential ‘ceiling’ effect was noted at high-levels of  
224 fluoroquinolone use, meaning that when use exceeded a second upper threshold, further  
225 increases in %Ec-ESBL+ were small. An initial decrease in %Ec-ESBL+ where  
226 fluoroquinolone use was between 151 and 161 DDDs per 1000 OBDs reflected uncertainty  
227 around this ceiling threshold, which may be resolvable with additional data. In the model for  
228 community %Ec-ESBL+ ( $R^2$  0.767), associations were identified with hospital %Ec-ESBL+,  
229 and community use of fluoroquinolones and co-amoxiclav above minimum thresholds.

230  
231 Autoregressive and population interaction effects suggested the importance of horizontal  
232 transmission of ESBLs, with predominant influence of community on hospital epidemiology.  
233 Population use of broad-spectrum beta-lactams and fluoroquinolones were important  
234 explanatory variables in both settings. In the community, use of fluoroquinolones had fallen  
235 below threshold levels and %Ec-ESBL+ had started to decrease by the end of the study.  
236 However, translating thresholds into antibiotic stewardship targets (table 1) suggested further  
237 restricting community co-amoxiclav use by 31% and hospital use of fluoroquinolones (-41%)  
238 and 3GC (-21%).

239  
240 From best-fit models, we created 24-month projections for %Ec-ESBL+ in the hospital and  
241 community under different antibiotic stewardship options, and compared these to expected  
242 %Ec-ESBL+ under a ‘business as usual’ scenario of antibiotic use (Supplementary Figure 2).  
243 Immediate restriction of hospital use of fluoroquinolones and 3GCs to thresholds was  
244 predicted to cause an abrupt and sustained reduction in hospital %Ec-ESBL+ from 9.89% to  
245 2.35% ( $p<0.0001$ ) and, due to population interactions, a gradual reduction in community  
246 %Ec-ESBL+ from 7.11% to 3.69% ( $p<0.0001$ ). Limiting community co-amoxiclav use to

247 threshold levels was predicted to cause a small reduction in community %Ec-ESBL+ but not  
248 to affect hospital epidemiology.

249

250

### 251 **Example 3: Cefepime-resistant *Escherichia coli* (Seville, Spain)**

252 We examined ecological variables explaining cefepime resistance among urinary or invasive  
253 *E.coli* infections (%Ec-FepR) in a tertiary hospital in Seville, Spain. Between March 2008  
254 and December 2016 (n=108 months) %Ec-FepR fell from 12.6% to 7.9% (figure 3).  
255 Cefepime use was low, declining from 4.4 to 1.6 DDDs per 1000 OBDs with thrice-weekly  
256 (Jan 2012, audit1) and daily (Jan 2014, audit2) prescription audits. By contrast, previously  
257 declining use of 3GC increased from January 2013 when they replaced co-amoxiclav and  
258 ciprofloxacin as first-line empirical therapy for intra-abdominal or urinary infections.

259

260 Resistance to cefepime in *E. coli* is mostly conferred by ESBLs with high affinity for  
261 cefepime (TEM-, SHV- and CTX-M-types). In addition to those agents hypothesised to  
262 predict ESBL-producing *E. coli* (see example 2), we considered the role of piperacillin-  
263 tazobactam.<sup>13</sup> Due to low rates of cefepime prescribing, we grouped this with 3GC use. We  
264 introduced variables for antibiotic auditing interventions and for revised susceptibility  
265 breakpoints (Oct 2014).<sup>33</sup>

266

267 The final non-linear model ( $R^2$  0.30) identified associations with %Ec-FepR 12 months prior  
268 (seasonal effect) and use of third- or fourth-generation cephalosporins and fluoroquinolones  
269 above minimum thresholds (Supplementary Table 1). A significant interaction term between  
270 %Ec-FepR in the previous month (autoregression, lag 1) and the second antibiotic auditing  
271 intervention suggested a gradual effect of the audit in reducing %Ec-FepR.

272

273 Reductions in third- and fourth-generation cephalosporins and fluoroquinolone use to below  
274 minimum thresholds explained modest declines in %Ec-FepR between 2008 and 2012. Partial  
275 reversal in this trend was consistent with increasing use of 3GC towards the end of the study  
276 period. %Ec-FepR could be controlled further by reducing third- and fourth-generation  
277 cephalosporin use by 41% (table 1).

278

### 279 **Example 4: Gentamicin-resistant *Pseudomonas aeruginosa* (Besançon, France)**

280 We examined ecological variables explaining rates of gentamicin-resistant *P. aeruginosa*  
281 (GRPa) among adult and paediatric admissions to a tertiary hospital in Besançon, France  
282 (figure 4). Between Jan 1999 and Dec 2014 (n=192 months), incidence density of GRPa  
283 decreased from 14.0 to 3.4 cases per 1000 OBDs, and the proportion of *P. aeruginosa*  
284 isolates resistant to gentamicin declined from 63% to 16%.

285

286 Aminoglycoside modifying enzymes (AMEs) are the most common mediators of  
287 aminoglycoside resistance in *P. aeruginosa*; with acetyltransferases (e.g. *aac(6')-Ib*) and  
288 nucleotidyltransferases (e.g. *ant(2'')-Ia*) most frequent in Europe.<sup>34</sup> Since related genes in  
289 mobile genetic elements encode AMEs and  $\beta$ -lactamases,  $\beta$ -lactam use may also predict  
290 aminoglycoside resistance.<sup>35</sup> In previous analyses from Besançon, aminoglycosides, cefepime  
291 and fluoroquinolones were predictors of MexXY-OprM overproduction in *P. aeruginosa*.<sup>36</sup>  
292 GRPa isolates were also more likely to overproduce the chromosomally-encoded AmpC  
293 cephalosporinase (56% vs. 20%;  $p<0.001$ ) and be multi-drug resistant (65% vs. 13%;  
294  $p<0.001$ ). We hypothesized that GRPa incidence density may be predicted by use of  
295 aminoglycosides, fluoroquinolones, extended-spectrum penicillins with  $\beta$ -lactamase  
296 inhibitors, carbapenems, monobactams, and third- and fourth-generation cephalosporins.



297 Given potential intra-class differences in promoting resistance, we grouped use of gentamicin  
298 and tobramycin separately from that of amikacin.

299

300 In the best-fit model ( $R^2=0.86$ ), GRPa incidence density was strongly predicted by incidence  
301 density in the previous month, and hospital use of gentamicin/tobramycin and  
302 fluoroquinolones above minimum thresholds (Supplementary Table 1). No independent  
303 association with Amikacin use was identified.

304

305 Declining GRPa was largely explained by falling inpatient use of gentamicin/tobramycin,  
306 and, to a lesser extent, fluoroquinolones. Use of both drug groups was maintained below  
307 minimum thresholds from around 2007. Continuing decreases in GRPa incidence density  
308 were at least partially explained by autocorrelation at lower incidence densities. Reciprocal  
309 increases in gentamicin-susceptible *P. aeruginosa* over the same period, and moderate  
310 inverse correlation ( $r=-0.55$ ), suggest competition for the same niche as GRPa.  
311 Gentamicin/tobramycin and fluoroquinolone use should be maintained below thresholds to  
312 control GRPa.

313

#### 314 **Example 5: Methicillin-resistant *Staphylococcus aureus* (Antrim, Northern Ireland)**

315 We evaluated ecological variables explaining incidence density of MRSA clinical isolates in  
316 adult admissions to a district general hospital in Antrim (Jan 2005 to Sep 2013, n=105  
317 months). Between 2005 and mid-2008, incidence density of MRSA clinical isolates remained  
318 stable at c.3.0 per 1000 OBDs (figure 5). Following restrictiong fluoroquinolones (January  
319 2008) and intensification of hand-hygiene, burdens fell to 1.64 cases per 1000 OBDs by  
320 2013.

321

322 The epidemic hospital MRSA clonal complex CC22, predominated in Northern Ireland  
323 during the study period: its success attributed to an ability to acquire mobile genetic elements  
324 carrying multiple resistance genes, with limited fitness costs.<sup>37</sup> Following prior linear time-  
325 series analyses from the region,<sup>14</sup> we hypothesised use of fluoroquinolones, 3GC, co-  
326 amoxiclav, and macrolides could be important predictors of MRSA epidemiology.

327

328 In the best-fit model ( $R^2=0.53$ ), MRSA incidence density was positively associated with rates  
329 of MRSA in the previous month and use of fluoroquinolones, 3GC, and co-amoxiclav  
330 exceeding minimum thresholds (Supplementary Table 1). An inverse relationship was seen  
331 with increased hospital use of alcohol-based hand rub (ABHR) up to 6.9 Litres per 1000  
332 OBDs, above which further increases in ABHR use was not associated with further declines  
333 in MRSA.

334

335 Declining MRSA incidence density was partly explained by deliberate restriction of  
336 fluoroquinolone use, and concurrent declines in co-amoxiclav and 3GC. Strong  
337 autoregression, and inverse association with ABHR use, were consistent with importance of  
338 infection control measures in interrupting horizontal transmission. Reversal of previous  
339 declines in fluoroquinolone use were seen by the last year of study. Findings suggested  
340 maintaining use of co-amoxiclav and 3GC under thresholds, use of ABHR at threshold levels,  
341 and further restriction of fluoroquinolones (table 1).

342

343

344

345

346 **Discussion**

347 Using a non-linear time-series analysis (NL-TSA) approach, we found empirical evidence of  
348 non-linear relationships between population antibiotic use and resistance rates in five  
349 European settings. The method was generalisable to different clinical populations, resistant  
350 pathogens, definitions of resistance burdens, and epidemiological phases. We demonstrated  
351 that identification of minimum thresholds, and associated confidence intervals, could provide  
352 population-specific quantitative targets for antibiotic stewardship.

353  
354 Our approach builds upon earlier work using linear time-series analysis to explain temporal  
355 relationships between antibiotic use and resistance.<sup>12-16</sup> NL-TSA shares a number of  
356 strengths, including: low-informational demands; ease of reiteration as new data becomes  
357 available; adjustment for the non-independence of serial observations and stochasticity  
358 inherent to time-series of communicable diseases; identification of temporality in associations  
359 and ‘lagged’ effects; and integration of impacts of multiple exposures.<sup>12</sup> Additionally, NL-  
360 TSA reveals non-linear relationships, providing more accurate understanding of how  
361 modifying antibiotic use, infection control or other exposures is likely to affect resistance.  
362 Limitations of NL-TSA include: the need for longer time-series than linear TSA; the potential  
363 for spurious thresholds in areas of limited data or extremes of the exposure variable range;  
364 and difficulty in identifying thresholds in situations of stable resistance and prescribing. We  
365 note the poorer predictive performance of some models (e.g. example 3), may be explained  
366 by absence of data on infection prevention and control activities, and resistance levels in  
367 interacting populations.

368  
369 Our findings may have important implications for antibiotic stewardship. In general, impacts  
370 of changes in antibiotic use on resistance vary dependent upon the antibiotic use level. More  
371 specifically, minimum thresholds may be interpreted as an upper limit for ‘safe’ population  
372 antibiotic use which does not appear to substantially increase resistance rates at the  
373 population level. Alternative theories may suggest the threshold indicates: a maximum level  
374 of selection pressure not conferring a survival advantage to resistant pathogens to spread  
375 within populations or ecological niches,<sup>8,12</sup> or strong enough to induce resistance;<sup>18</sup> or a  
376 minimum level of use, below which antibiotic substitution creates equivalent or greater  
377 selection pressure.<sup>8</sup> Crucially, they may provide quantitative targets for balancing the need to  
378 access therapies with control of resistance, analogous to ‘quotas’ applied to other natural  
379 resources which seek to maximize extracted value while maintain a non-declining stock.<sup>1,2,38</sup>  
380 Moving from qualitative targets of reducing use, to quantitative targets may also aid  
381 operational effectiveness. Targets appear to work best if pragmatic, collaborative and  
382 iterative.<sup>39</sup> Complete restriction of use of agents is rarely feasible: in balancing access to  
383 effective therapies with control of resistance, quantitative targets could align interests of  
384 clinicians and antimicrobial management team.<sup>40</sup>

385  
386 We emphasise the need for caution with interpretations of thresholds. Firstly, thresholds  
387 should offer guidance rather than strict limits. Uncertainty around thresholds is variable, as  
388 reflected in width of associated confidence intervals. Narrower confidence intervals around  
389 threshold locate with reasonable precision the level of antibiotic use at which effects on  
390 resistance are substantially altered. Wider intervals may indicate insufficient data, or the  
391 influence of additional explanatory variables. In rare instances of multiple closely occurring  
392 thresholds in a single functional relationship, the width of confidence intervals may be under-  
393 estimated and should be interpreted with particular caution. We suggest a pragmatic  
394 approach, of interpreting thresholds depending upon the policy scenario. Where the priority is  
395 strict control of resistance a conservative approach of limiting use to the lower limit of the

396 threshold confidence interval is advisable. Where excessive restriction is a concern, the  
397 standard approach of limiting use to the point estimate of the threshold is likely to offer the  
398 best balance between restriction and control of resistance. Secondly, changes in molecular  
399 epidemiology under sustained antibiotic selection pressures, such as compensatory mutations  
400 minimising fitness costs,<sup>26,27</sup> or strain replacement,<sup>24,25</sup> may mean thresholds vary by  
401 epidemiological phase and time. Variation in thresholds across populations can be  
402 anticipated, reflecting host, environment, and organism factors.<sup>23-25</sup> Therefore, models based  
403 on local data and iterative analysis is necessary to ensure time and context-specific guidance.  
404 Thirdly, it is important not to assume that all antibiotic use below thresholds is safe, since  
405 antibiotic exposures may be important for individual patients, or cause unseen change in  
406 reservoirs of resistant pathogens in environment or hosts.

407  
408 Potential foci for further research include: evaluating the consistency of thresholds for  
409 specific antibiotic use-resistance combinations across different settings and identifying  
410 factors affecting thresholds; applying NL-TSA to composite indices of resistance;<sup>41,42</sup> use of  
411 Bayesian approaches in selection of explanatory variables and analysis with short time-series  
412 or rare resistance outcomes;<sup>43</sup> and prospective studies to validate the effectiveness of  
413 quantitative targets in antibiotic stewardship.

414  
415 We have illustrated how non-parametric time-series models based on empirical data can  
416 identify non-linear relationships between population antibiotic use and resistance burdens.  
417 Further we have shown how identification of population-specific minimum thresholds may  
418 guide rational compromises between control of resistance and access to therapeutics. With  
419 the increasing availability of electronic surveillance and healthcare systems, this approach  
420 offers a useful tool for sustaining the effectiveness of current antimicrobials in many areas of  
421 the world.

422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445

446 **Methods**

447

448 **Design and study populations**

449 This was a multi-centre time-series study. We used multivariable non-linear time-series  
450 analysis to quantify associations between ecological exposures, including population use of  
451 antibiotic groups, and rates of antibiotic-resistant infections in five populations from France,  
452 Hungary, Northern Ireland (UK), and Spain (Supplementary Table 5).

453

454 The populations and resistance outcomes were a purposive sample, chosen to reflect varying  
455 epidemiological scenarios, clinical settings, and resistant infections in European centres  
456 participating in the THRESHOLDS (THReshold ESTimation to Help Optimise Local  
457 Decisions on antibiotic Stewardship) study group. This collaborative aims to further the  
458 development of time-series analysis for understanding antibiotic resistance and planning  
459 antibiotic stewardship. Members included centres with prior experience in applying linear  
460 time-series approaches. Investigators were asked to identify an important resistance problem  
461 in a defined clinical population from their region. Minimum data requirements were  
462 consistent microbiological and prescribing data across a minimum of 60 monthly  
463 observations (5 years). Duration of time-series was defined by the longest period of  
464 consistent data for a minimum data set of the outcome and candidate explanatory variables.

465

466 For each population we described the regional scenario for the chosen outcomes, the  
467 theoretical basis for inclusion of candidate explanatory variables, findings from the non-linear  
468 time-series analysis, and how these could inform local antibiotic stewardship policy.

469

470 **Outcome and explanatory time series**

471 The outcome time-series for each population were: carbapenem-resistant *Acinetobacter*  
472 *baumannii* (Debrecen, Hungary); extended spectrum  $\beta$ -lactamase producing *Escherichia coli*  
473 in hospital and community (Orihuela, Spain); cefepime-resistant *Escherichia coli* (Seville,  
474 Spain); gentamicin-resistant *Pseudomonas aeruginosa* (Besançon, France); and methicillin-  
475 resistant *Staphylococcus aureus* – MRSA (Antrim, Northern Ireland). Cases were defined  
476 microbiologically as isolates from all relevant body sites not identified as infection control  
477 specimens and meeting consistent criteria for resistance or resistance mechanism (see  
478 supplemental file for details). Isolates from the same patient identified within 30 days of a  
479 prior isolate with the same organism were considered part of the same infectious episode and  
480 de-duplicated. Outcomes were expressed, where possible, as monthly incidence density of  
481 resistant infections (cases per 1000 or 10,000 occupied bed days, OBDs). Where there were  
482 large changes in testing frequency or organism identification over time, we defined resistance  
483 as a percentage of clinical isolates from the same organism with any susceptibility pattern.

484

485 The primary explanatory variables were monthly population use of antibiotic agents  
486 classified by pharmacological sub-group of antibacterials for systemic use (J01) in the 2016  
487 WHO/ATC index, and expressed as defined daily doses (DDDs) per 1000 OBDs (hospital) or  
488 1000 inhabitant-days (community). Candidate antibiotic sub-groups were identified *a priori*,  
489 on the basis of regional co-resistance profiles, molecular epidemiology in the region, reviews  
490 and prior evidence on individual or population level risk factors for acquisition of the  
491 resistant infection. We included separate time-series for individual antibiotic agents or  
492 chemical sub-groups only where there were strong theoretical grounds for investigating  
493 independent associations, such as prior evidence of variable within-class actions or targeting  
494 within antibiotic stewardship interventions.

495

496 In addition, we incorporated autoregressive terms capturing association between current  
497 incidence density and incidence density in recent months. Where available, we incorporated  
498 infection prevention and control (IPC) variables such as use of alcohol-based hand-rub.  
499 Dummy variables were added to capture immediate and gradual impacts of changes in  
500 laboratory methods or other planned interventions. We considered lags in association of up to  
501 6 months and seasonal autoregressive terms (lag 12).

502

### 503 **Data collection and laboratory procedures**

504 In all centres, microbiological and prescribing and IPC data were extracted from electronic  
505 databases maintained for routine healthcare activities. Data were anonymised and aggregated  
506 before electronic submission to the THRESHOLDS study group. Meta-data on population  
507 characteristics, IPC activities, antibiotic stewardship interventions and resources for control  
508 of antibiotic resistance were captured using a standardised questionnaire.

509

510 Pathogens were identified using standard laboratory methods. Susceptibility testing was by  
511 disc-diffusion (Besançon, Debrecen, Antrim) or broth microdilution (Seville, Orihuela).  
512 Isolates were defined as resistant if not susceptible to an antibiotic agent according to zone  
513 diameter (for disc-diffusion) or minimum inhibitory concentration (MIC, for broth  
514 microdilution) breakpoints as recommended by the European Committee on Antimicrobial  
515 Susceptibility Testing (EUCAST) or Clinical & Laboratory Standards Institute (CLSI).  
516 Details of standards used, and deviations from EUCAST or CLSI criteria are detailed in table  
517 i. Known changes in breakpoints and laboratory methods were adjusted for in time-series  
518 analysis.

519

### 520 **Statistical methods**

521 In the following sections we provide a technical exposition of the statistical methods used.  
522 We offer here, a brief description for the general reader.

523

524 Before applying non-linear time-series analysis (NL-TSA) to real-world datasets from study  
525 populations, we performed a statistical, Monte-Carlo, experiment to illustrate its advantages  
526 over, more familiar, linear TSA. We applied both linear and non-linear TSA to computer-  
527 simulated time-series, where the relationship between the outcome and explanatory time-  
528 series was known. Three types of relationship were explored: a linear relationship; a non-  
529 linear relationship without correlations between successive data points in time-series  
530 (autoregression), and a non-linear relationship with autoregression; By running this  
531 experiment over 10,000 simulated datasets for each type of relationship we evaluated the  
532 typical ability of linear and non-linear TSA to describe the relationships accurately.

533

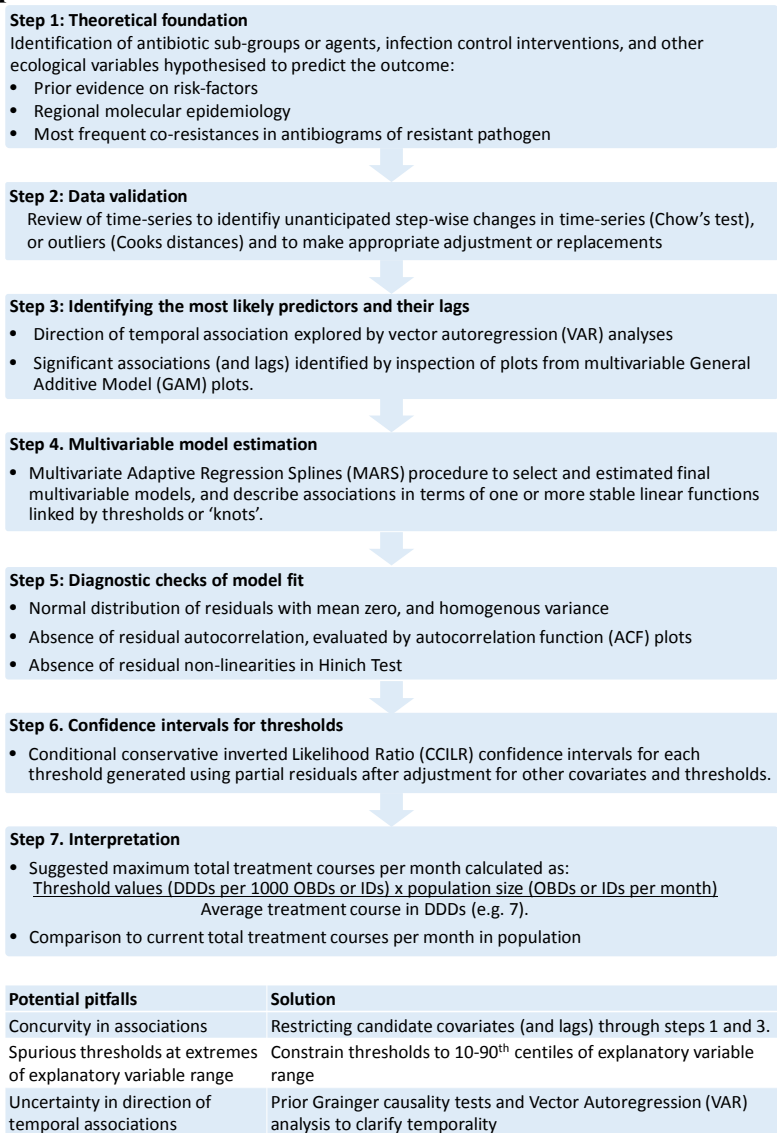
534 We next applied a seven-step non-linear time-series analysis approach to resistance problems  
535 in study populations (figure 6).

536

537 Mirroring more familiar regression techniques, we started by defining a set of explanatory  
538 variables including antibiotic use, infection control, population interactions, alongside terms  
539 for autoregression. This set was defined by (a) expert opinion informed by prior risk-factor  
540 studies, molecular epidemiology in the region; and (b) inspection of resistance profiles of the  
541 pathogen of interest. We consider delays between changes in explanatory time-series and  
542 associated change in outcome time-series (lags) of up to 6 months. Before analysis we  
543 checked time-series and make adjustments for extreme values (outliers) or unexpected shifts  
544 in mean (structural changes). We also used vector autoregression (VAR) models to help  
545 distinguish any reverse causality in relationships between explanatory and outcome series:

546 this might occur, for example, if prescribing behaviour was altered by resistance rates in the  
 547 population in previous months. Next we restrict the set of explanatory variables, and lags, to  
 548 be put into final multivariable models. We use a procedure that fits smooth functions to  
 549 relationships between explanatory and outcome time-series, and allows visual inspection of  
 550 likely significant associations. After identifying the most promising explanatory variables  
 551 (and lags) we enter these into a Multivariate Adaptive Regression Splines (MARS) model  
 552 which both identifies significant predictors, and defines any non-linear relations as a series of  
 553 linear relationships connected by ‘knots’ or thresholds. Model fit for MARS is checked by  
 554 ensuring residuals were normally distributed without unexplained non-linearities. Confidence  
 555 intervals around each threshold were fit by a conditional conservative inverted likelihood  
 556 ratio (CCILR) method, using partial residuals. Finally, we converted thresholds from models  
 557 into suggested maximum total treatment courses per month in the population by multiplying  
 558 model thresholds by the size of the population and dividing by an average treatment course.  
 559

560 **Figure 6: Summary of the 7-step non-linear time-series methodology and potential**  
 561 **pitfalls**



562  
 563

564 To provide an example of how model findings can inform policy, for the Orihuela population,  
 565 we predicted the effects of restricting antibiotic use to threshold levels compared to a  
 566 ‘business as usual’ scenario of prescribing based on the last months of study. We quantify  
 567 impacts of immediate restriction sustained over two years.

568  
 569 A technical explanation follows:

570  
 571 **a) Comparing linear and non-linear time-series model performance**

572 We used a Monte-Carlo experiment to compare the ability of linear (Ordinary Least Squares,  
 573 OLS) and non-linear (Multivariate Adaptive Regression Splines, MARS) time-series models  
 574 in identifying various pre-defined functional relationships between simulated explanatory and  
 575 outcome time-series. We hypothesised that for time-series related by simple linear processes  
 576 MARS and OLS regression methods would perform equally well, but that only MARS would  
 577 accurately identify non-linear associations. We generated 10,000 simulated datasets using  
 578 simple stochastic processes incorporating the following pre-defined functional relationships:

579  
 580 (i) Non-autoregressive without threshold

581 
$$y_t = -4 + 2x_t + u_t \quad u_t \sim N(0, \sigma^2 = 0.02) \quad \forall t = 1, \dots, 200$$

582  
 583 (ii) Non-autoregressive with threshold

584  
 585 
$$\begin{cases} \text{if } x_t \leq 2: & y_t = u_t & u_t \sim N(0, \sigma^2 = 0.5) \\ \text{if } x_t > 2: & y_t = -4 + 2x_t + u_t & u_t \sim N(0, \sigma^2 = 0.5) \end{cases} \quad \forall t = 1, \dots, 200$$

586  
 587 (iii) Autoregressive with threshold

588  
 589 
$$\begin{cases} \text{if } x_t \leq 0: & y_t = 0 + \rho y_{t-1} + u_t & u_t \sim N(0, \sigma^2 = 0.05) \\ \text{if } x_t > 0: & y_t = 0 + \rho y_{t-1} + 1x_t + u_t & u_t \sim N(0, \sigma^2 = 0.05) \end{cases} \quad \forall t = 1, \dots, 200$$

590 Where:

- 591  $x_t$  is the explanatory (independent) time-series variable at time  $t$   
 592  $y_t$  is the outcome (dependent) time-series variable at time  $t$   
 593  $u_t$  is the error term at time  $t$ , with Normal distribution, zero mean and variance  $\sigma^2$   
 594  $\rho y_{t-1}$  is an autoregressive term or order 1 (i.e. dated at  $t-1$ ) with  $\rho = 0.25$

595  
 596 For each dataset we fitted both linear and non-linear time-series analyses, and recorded  
 597 sample parameter estimates ( $a$  constant,  $b$  slope, and  $s^2$  as the estimate of population  
 598 variance) and a measure of goodness of fit ( $R^2$ ). Histograms were created illustrating the  
 599 distributions of  $R^2$  values and parameter estimates from both linear and non-linear models.  
 600 Visual comparison was made to pre-defined parameter values to identify bias in parameter  
 601 estimates. We used a t-test of mean difference for independent samples to compare model  
 602 performance based on  $R^2$  values.

603  
 604 **b) Applications of non-linear time-series analysis to real-world datasets**

605 We applied a seven-step approach to generate non-linear time-series models describing how  
 606 contemporaneous and prior population antibiotic use, and other ecological variables, explain  
 607 monthly variation in clinical burdens from antibiotic-resistant infections in five European  
 608 centres.

609

610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
  
624  
625  
  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658

*Step 1: Theoretical foundation*

Participating centres identified *a priori* a minimum dataset of antibiotic sub-groups or agents they considered most likely to affect the epidemiology of the resistant organism under investigation (target organism). Decisions were based upon: previous empirical evidence of risk factors and molecular epidemiology in the study region or related contexts. Additionally, using antibiogram data from the study period and population, we reviewed co-resistances to other antibiotics among isolates of the target organism with and without the resistance under investigation. We considered antibiotics with the largest absolute rates of co-resistance in the resistant isolates to be most likely to exert significant selection pressures.<sup>44,45</sup> Consensus on the list of potential predictive variables was found through discussion among all THRESHOLDS study group members.

Where data was available we integrated additional explanatory variables on hospital activity or infection control activities, associated with the outcome variable in previous studies.

*Step 2: Data validation*

To ensure consistent time-series we first accounted for known changes in exogenous conditions, such as changes in laboratory method. We captured immediate and gradual effects by entering a dummy variable (0 in months before the change, 1 in months after change) and its interaction with an autoregressive term, as explanatory variables. We then reviewed time-series to detect possible unknown measurement errors as follows. Visual inspection identified potential structural changes (seen as large step-wise change in mean for instance) or outliers (seen as values deviating substantially from surrounding values). Successive Chow tests were applied to automatically detect the most probable dates of structural changes in the time-series and, where necessary, to disaggregate the sample into two or more segments, each with a stable mean. For each segment we applied an outlier detection technique using the following criterion: an observation was considered as an outlier if Cook's distance at this point was greater than five times the mean of Cook's distances of all the observations of the segment. Finally, we replaced outlier values with the mean of the three preceding and three following observations.

*Step 3 Identifying the most likely predictors and their lags*

Given the potentially complex relationships between ecological variables under investigation we sought to refine our understanding of potential associations before applying final multivariable non-linear models.

Firstly, situations of reverse causality could exist when ecological exposures - such as rates of infections with resistant pathogens connected populations, or use of some antibiotic groups in a given population - respond to, rather than predict, rates of resistance. In order to minimise this risk, we tested direction of temporal relationships between explanatory and outcome time-series by applying Granger-causality analysis and Vector Autoregression (VAR) models. Secondly, non-linear models of the type used in this study are potentially complex and difficult to extract form the data if too many predictors are used at the same time. Therefore, we carried out an additional *a priori* data-based selection of candidate explanatory variables and lags (the lag refers to the delay in months between change in exposure and associated change in outcome). This was done through inspection of outputs from fitting a General Additive Model (GAM) to the data. GAM is a very general procedure that can be



659 used for the identification of the most likely predictors, since it runs a non-parametric  
 660 estimation of the functional relationships between explanatory ( $x$ ) and outcome ( $y$ ) time-  
 661 series, based upon iterative data fitting, rather than prior assumptions. It also allows for  
 662 variability in the functional relationships across different values of the explanatory variables  
 663 and can therefore capture non-linear associations between ecological variables and resistance  
 664 outcomes.<sup>49</sup> In particular, we used the GAM procedure in the SCAB34S Splines module  
 665 (available in SCA Workbench, Scientific Computing Associates Corp, Illinois, USA) to  
 666 define the relationship between  $p$  explanatory ( $x$ ) and the outcome ( $y$ ) time-series as a sum of  
 667 smooth, or spline, functions:

$$E(y|x_1, x_2, \dots, x_p) = s_0 + \sum_{j=1}^p s_j(x_j)$$

668 where ( $s_j(x_j)$ ) are the spline functions; they were standardised such that, after removal of free  
 669 constants ( $s_0$ ) their expected contribution to the outcome ( $y$ ) is zero (i.e.  $E s_j(x_j)=0$  for each  $j$ ).  
 670

671  
 672 The splines were derived by a process of splitting the time-series into sections, joined at knot  
 673 points, and fitting simple curves described by cubic functions to the data in each section. The  
 674 GAM methodology identified the optimal combination of spline functions  $s_j(x_j)$ , following  
 675 the iterative procedure suggested by Hastie and Tibshirani(1986).<sup>46</sup> Combining a local  
 676 scoring algorithm and a backfitting procedure, this method converges on a solution balancing  
 677 data fit with smoothness.

678  
 679 To identify the most relevant explanatory time-series, for each centre we started with a  
 680 multivariable GAM model including all theoretically relevant variables at lags of 1 to 6  
 681 months and autoregressive terms at lags 1 and 2. We limited lags to 6-months based on  
 682 widespread evidence of declining relevance of antibiotic exposures by time-since exposure,  
 683 and prior experience that considerations of longer lags lead to problems of concurvity. On  
 684 the basis of the GAM outputs, an explanatory variable with a specific lag was retained in the  
 685 model only if its contribution was significant at a 5% level of probability (identified on  
 686 contribution charts by the zero line of non-association falling outside of 95% confidence  
 687 intervals around the estimate). The process was run iteratively by removing first those  
 688 variables and lag combinations whose contributions were non-significant before re-running  
 689 the GAM model on a reduced subset of variables and lags. The process stopped when the  
 690 model contained only significant contributions of variables and lags. These constituted the  
 691 restricted set of explanatory variables for entry into MARS analysis.

692  
 693 A further objective of applying the GAM procedure was to determine whether consideration  
 694 of non-linear associations is justified in terms of improvement in predictive performance. For  
 695 each explanatory variable (and lag), GAM provides a comparison of the data fit of a non-  
 696 linear spline function ( $nl$ ) with an analysis in which this relationship is restricted to a linear  
 697 function ( $l$ ). Significant improvement in goodness of fit over a linear fit is defined by an  $F$ -  
 698 test, as follows:

$$F_0 = \frac{(SSR_l - SSR_{nl}) / (p_{nl} - p_l)}{SSR_{nl} / (n - p_{nl})} \sim F_{(df_{nl} - df_l), (n - p_{nl})}$$

700  
 701 where;  $SSR$  = Sum of squares of residuals,  $n$  = number of observations,  $p$  = number of  
 702 parameters,  $l$  = linear function, and  $nl$  = non-linear spline function.  
 703

704 The null hypothesis that estimates from an enhanced (non-linear) model do not provide a  
 705 significantly better fit than those from a linear model can be rejected where  $F$  exceeds a  
 706 critical value ( $\alpha = 0.05$ ) from an F-distribution with  $(p_{nl}-p_l, n-p_{nl})$  degrees of freedom.

707

708 *Step 4. Multivariable model estimation*

709 After identifying the most likely explanatory variables (and lags), and whether associations  
 710 with the outcome series were linear or non-linear, we used the MARS procedure (in the  
 711 SCAB34S Splines module) to obtain an easily interpretable characterization of these  
 712 associations.<sup>47</sup> MARS is a non-parametric regression approach suitable for situations of non-  
 713 linear associations that can provide more interpretable and interesting empirical results than  
 714 GAM. Given our research aims, its particular advantages were: (i) the ability to identify  
 715 distinct threshold values ('knots') of the explanatory variables delimiting regions (ranges of  
 716 values) of individual or interacting explanatory variables within which associations with the  
 717 outcome differs substantially from those in other regions; and (ii) a systematic approach for  
 718 model identification and estimation, which automatically selects the combination of  
 719 explanatory variables and threshold values which most efficiently explain variation in  
 720 outcomes. MARS is related to, but more general than, regression tree algorithms, and has  
 721 several important, interrelated advantages: (a) when thresholds may affect several  
 722 explanatory variables, the final MARS model is far more synthetic, and therefore more  
 723 interpretable, in terms of quantifying how each explanatory variable (e.g. antibiotic use)  
 724 affects the response variable (e.g. resistance) and interacts with other explanatory variables;  
 725 (b) the MARS algorithm in general deals much better with numerical data and continuous  
 726 data; (c) it allows for dynamic relationships and thresholds in the dynamic dependences -  
 727 especially important when analysing effects with time series data; and (d) MARS does better  
 728 in extrapolating the results outside the sample data ranges of explanatory and outcome  
 729 variables. These advantages are important when considering policy implications of findings  
 730 for control of AMR.

731

732 MARS approximates the functional relationship between an outcome time-series ( $y_t$ ) and a  
 733 vector of  $p$  explanatory variables  $x_t = (x_t^1 \dots x_t^p)$  as:

734 
$$y_t = \beta_0 + \sum_{m=1}^M \beta_m b_m(x_t) + \varepsilon_t$$

735 where;

736  $\beta_0$  is a constant

737  $\beta_m$  is the coefficient for the  $m^{th}$  basis function,  $m=1, \dots, M$

738  $b_m(x_t)$  is the  $m^{th}$  basis function,  $m=1, \dots, M$

739  $\varepsilon_t$  is an independently distributed error term.

740 The basis functions are products of up to two truncated linear or hinge functions, describing  
 741 the relationship between one or more explanatory variables and the outcome in terms of  
 742 segments of stable association separated by knots or thresholds values. These interacting  
 743 hinge functions allow us to identify possible interactions between variables as in Figure  
 744 1C(ii). Namely, the  $m^{th}$  basis function takes one of the following two forms:

745 No interaction:  $b_m(x_t) = h(x_t^k, t_{k,m})$  for some  $k = 1, \dots, p$

746 With interaction:  $b_m(x_t) = h_m(x_t^k, t_{k,m}) \cdot h_m(x_t^j, t_{j,m})$  for some  $k, j = 1, \dots, p, k \neq j$

747 where  $t_{k,m}$  is the threshold value of  $x_t^k$  in the  $m^{th}$  basis function and where  $h(x_t^k, t_{k,m})$  is a  
 748 hinge function that takes the following form depending on whether the basis function takes  
 749 effect above or below the threshold  $t_{k,m}$

750 a) above the threshold:  $h_m(x_t^k, t_{k,m}) = \max(x_t^k - t_{k,m}, 0)$

751 b) below the threshold:  $h_m(x_t^k, t_{k,m}) = \max(t_{k,m} - x_t^k, 0)$

752

753 If no knot (threshold) is detected, then a simple linear (and therefore constant) association  
754 between explanatory and outcome variable can be specified as a single function applied  
755 across the total range of values of the explanatory variable.

756

757 All potentially significant explanatory variables, and associated lags, identified in previous  
758 steps, were incorporated into models. Model identification and estimation proceed by an  
759 automated, iterative, process:

760

761 *Forward pass:* starting with the simplest model containing only a constant basis function,  
762 MARS generates a matrix of basis functions in a forward stepwise manner. Candidate base  
763 functions are added in order of ability to improve model fit, until the model reaches a  
764 predefined limit of complexity. The candidate basis functions are identified by a nested  
765 exhaustive search looping over the existing set of basis functions, and all other possible  
766 explanatory variables (or interactions) and knot positions.

767

768 *Backwards (pruning) pass:* During the subsequent pruning pass MARS removes basis  
769 functions contributing least to model fit, until no significant improvement is seen in a  
770 modified form of the generalized cross validation (MGCV) criterion:

$$\text{MGCV} = \frac{\frac{1}{N} \sum_{i=0}^n (y - f(x))^2}{1 - ((C(M) + dM)/n)^2}$$

771

772

773 Where;  $N$  is the number of observations,

774  $\sum_{i=0}^n (y - f(x))^2$  is the sum of square of residuals (observed - estimated  $y$ ).

775  $C(M)$  is the number of parameters being fitted,  $M$  the number of non-constant basis  
776 functions and  $d=3$  (conventional value).

777

778 The MGCV incorporates a complexity penalty accounting for the inherent improvement in  
779 explained variance associated with increasing numbers of basis-functions, and its calculation  
780 allows estimates of the relative importance of each basis function. Model selection therefore  
781 converges on a set of basis functions that most efficiently explain variation in antibiotic  
782 resistance before a final model fit by OLS estimation.

783

784 From the output of each MARS model we generated contribution charts illustrating the  
785 change in the outcome time-series across the observed ranges of explanatory variables.

786

787 *Step 5. Diagnostic checks*

788 Adequacy of model fit was defined by three criteria: (i) *Normally distributed residuals* – with  
789 homogenous variance and mean equal to zero, as evaluated by a Normality test; (ii) *absence*  
790 *of significant residual autoregression* – identified in lags 0 to 6 on an autocorrelation function  
791 (ACF) plot; and (iii) *absence of residual non-linearities* – as evaluated by a Hinich test. In  
792 addition to the MGCV we reported more familiar measures of model performance such as  $R^2$   
793 and the mean absolute percentage error (MAPE).

794

795 *Step 6: Confidence intervals (CIs) for thresholds values*

796 In the absence of an existing method for deriving measures of uncertainty around thresholds  
797 derived from non-parametric MARS models, we develop a procedure inspired by Hansen  
798 (2000).<sup>48</sup> His procedure considers a simple threshold model with only one variable affected  
799 by a threshold effect, and obtains a distribution theory for the threshold parameter ( $\tau$ ) from  
800 which asymptotic confidence intervals can be built. He first derives the limiting distribution  
801 of a Likelihood Ratio test (LR) for the null hypothesis that the threshold parameter  $\tau = \tau_0$ . He  
802 then builds confidence intervals through the inversion of LR: the  $(1-\alpha)$  Inverted Likelihood  
803 Ratio (ILR) confidence interval consisting of all the possible values of  $\tau$  for which the null  
804 hypothesis would not be rejected at the  $\alpha$  level. Donayre et al. (2018) examine improvements  
805 of Hansen's ILR confidence interval, increasing its quality in finite samples with large  
806 threshold effects (i.e. when the change in slope from one side of the threshold to the other is  
807 large).<sup>49</sup> They show that a 'conservative modification' enlarging Hansen's ILR confidence  
808 interval is optimal. In this "conservative ILR confidence interval" the lower end of the  
809 interval is enlarged from the first value lower than  $\tau_l$  for which the null hypothesis is  
810 rejected, up to  $\tau_l$ ; at the upper end, it is enlarged from  $\tau_u$  up to the first value greater than  $\tau_u$   
811 for which the null hypothesis is rejected. This modification provides intervals at a confidence  
812 level at least as high as the nominal one that are still informative.

813

814 We adapted this procedure for MARS estimations with more than one explanatory variable  
815 containing thresholds, and one or more thresholds per variable, by using the partial residuals  
816 –i.e. the variation in the outcome not explained by other explanatory variables and their  
817 thresholds. This allowed us to identify conservative ILR confidence intervals for each  
818 explanatory variable, conditional on all the estimated coefficients and thresholds other than  
819 the one for which the confidence interval is computed. Since in MARS all thresholds and  
820 slope coefficients are anyway selected and estimated to optimise overall model fit using  
821 conditional inference, identifying these 'conditional conservative ILR (CCILR) confidence  
822 intervals' does not impose costs to reliability. Computing confidence intervals conditional  
823 only on other thresholds, but with re-estimation of coefficients describing piece-wise  
824 associations (slopes), offers a valid alternative. In Monte Carlo simulations (results available  
825 on request) we found both approaches resulted in adequate coverage (>95% of intervals  
826 including the actual value of the threshold) but CIs were wider with a procedure with slope  
827 re-estimation and less informative for a given coverage rate. Our simulations show that the  
828 superiority of the partial residual approach is independent of the degree of correlation  
829 between antibiotic segments (values of antibiotic use over which association with the  
830 outcome is constant) where correlation is between 0% and 75%. There is some risk of very  
831 slight under coverage with the partial residual approach where correlations are  $\geq 85\%$ , but  
832 such high levels of correlation are unlikely unless several the functional relationship involves  
833 multiple, closely occurring, thresholds. In such situations, re-estimation of slope co-efficients  
834 is anyway not feasible. As a result, we recommend the partial residual approach, with a  
835 caution that the CI could be somewhat wider in specific and infrequent situations where  
836 multiple close thresholds for the same independent variable are detected (e.g. fluoroquinolone  
837 association with %Ec-ESBL in Orihuela hospital population).

#### 838 *Step 7. Interpretation*

839 The minimum thresholds identified for each significantly associated antibiotic group were  
840 translated into suggested maximum numbers of patient treatment courses per month not  
841 expected to adversely affect resistance at population levels. We multiplied the threshold,  
842 expressed in DDDs per 1000 OBDs (or IDs), by the size of the monthly patient population (in  
843 thousands of OBDs or IDs), and then divided by an average patient treatment course of 7  
844 DDDs (except for aminoglycosides which were considered as 3 DDDs). These maximums  
845 were further compared to contemporary levels of antibiotic use in the last year of study, to

846 provide indications of how current use of antibiotics should be modified to avoid resistance  
847 spread.

848

849

### 850 **Projections for alternative antibiotic stewardship policy options**

851 To illustrate the potential effects of restricting antibiotics associated with %Ec-ESBL+ to  
852 threshold levels in populations of Orihuela, we compared the expected evolution of %Ec-  
853 ESBL+ under a ‘business as usual’ scenario in which antibiotic use continued as in last year  
854 of study, to projected time-series with antibiotics restricted to threshold levels. We used a  
855 breakpoint analysis to identify the last stationary segment in %Ec-ESBL+ time-series from  
856 the study period. We recursively estimated MARS models using means from these stationary  
857 segments as starting points to derive steady states for %Ec-ESBL+ in community and  
858 hospital populations. Based on steady state values and MARS models for the study period  
859 (baseline) we simulated 1000 samples of 24-month projections, incorporating random error  
860 term with variance as derived in the baseline MARS model. For each sample projection we  
861 entered mean antibiotic levels in the last year of the study period (‘business as usual’) and  
862 alternative levels set at identified thresholds. We calculated the mean difference between  
863 business as usual and each policy option for every month along with 95% confidence  
864 intervals. Finally, we illustrated alternative projections and differences using medians of  
865 distributions from the 1,000 sample projections.

866

867

### 868 **Data availability**

869 The data that support the findings of this study are available from the corresponding  
870 author upon reasonable request.

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943

## References

1. Laxminarayan, R. Antibiotic effectiveness: Balancing conservation against innovation. *Science* **345**, 1299–301 (2014)
2. Laxminarayan, R., et al. Antibiotic resistance—the need for global solutions. *Lancet Infect Dis.* **13**, 1057–98 (2013)
3. World Bank. Drug-Resistant Infections: A Threat to Our Economic Future (discussion draft). World Bank, Washington DC. Available from: <http://pubdocs.worldbank.org/en/689381474641399486/1701381-AMR-Lab-Report-Web.pdf> (2016).
4. O’Neill, J. (chair). Review on Antimicrobial Resistance. Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations. Available from: <https://amr-review.org/> (2014)
5. Millar, M. Constraining the use of antibiotics: applying Scanlon’s contractualism. *J Med Ethics.* **38**, 465-9 (2012)
6. WHO. Global action plan on antimicrobial resistance. World Health Organization, Geneva, Switzerland. Available from: <http://www.who.int/antimicrobial-resistance/publications/global-action-plan/en/> (2015).
7. Davey, P., et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database of Systematic Reviews* **2**, CD003543. (2017)
8. Peterson, L.R., Squeezing the antibiotic balloon: the impact of antimicrobial classes on emerging resistance. *Clin Microbiol Infect.* 2005 Oct;11 Suppl 5:4-16.
9. Levy, S.B. Balancing the drug-resistance equation. *Trends Microbiol.* **2**, 341–2 (1994)
10. Levin, B.R., et al. The Population Genetics of Antibiotic Resistance. *Clin Infect Dis.* **24**(Supplement 1), S9–16 (1997)
11. Austin, D.J., Kristinsson, K.G. & Anderson, R.M. The relationship between the volume of antimicrobial consumption in human communities and the frequency of resistance. *Proc Natl Acad Sci.* **96**, 1152–6 (1999).

- 944 12. López-Lozano, J.M., et al. Modelling and forecasting antimicrobial resistance and its  
945 dynamic relationship to antimicrobial use: a time series analysis. *Int J Antimicrob Agents*. **14**,  
946 21–31 (2000).  
947
- 948 13. Vernaz, N., et al. Modelling the impact of antibiotic use on antibiotic-resistant  
949 *Escherichia coli* using population-based data from a large hospital and its surrounding  
950 community. *J Antimicrob Chemother*. **66(4)**, 928-35 (2011).  
951
- 952 14. Aldeyab, M.A., et al. Modelling the impact of antibiotic use and infection control  
953 practices on the incidence of hospital-acquired methicillin-resistant *Staphylococcus aureus*: a  
954 time-series analysis. *J Antimicrob Chemother*. **62**, 593-600 (2008).  
955
- 956 15. Monnet, D.L., et al. Antimicrobial drug use and methicillin-resistant *Staphylococcus*  
957 *aureus*, Aberdeen, 1996-2000. *Emerg Infect Dis*. **10**, 1432-41 (2004).  
958
- 959 16. Aldeyab, M.A., et al. The impact of antibiotic use on the incidence and resistance pattern  
960 of extended-spectrum beta-lactamase-producing bacteria in primary and secondary healthcare  
961 settings. *Br J Clin Pharmacol*. **74**, 171-9 (2012).  
962
- 963 17. Haber, M., Levin, B.R. & Kramarz, P. Antibiotic control of antibiotic resistance in  
964 hospitals: a simulation study. *BMC Infect Dis*. **10**, 254 (2010)  
965
- 966 18. Oz T, Guvenek A, Yildiz S, et al. Strength of Selection Pressure Is an Important  
967 Parameter Contributing to the Complexity of Antibiotic Resistance Evolution. *Molecular*  
968 *Biology and Evolution*. **31(9)**, 2387-2401 (2014).  
969
- 970 19. Arepyeva, M.A., et al. A mathematical model for predicting the development of bacterial  
971 resistance based on the relationship between the level of antimicrobial resistance and the  
972 volume of antibiotic consumption. *J Glob Antimicrob Resist*. **8**, 148-156 (2017).  
973
- 974 20. Kiffer, C.R., et al. A spatial approach for the epidemiology of antibiotic use and  
975 resistance in community-based studies: the emergence of urban clusters of *Escherichia coli*  
976 quinolone resistance in Sao Paulo, Brasil. *Int J Health Geogr*. **10**, 17 (2011).  
977
- 978 21. Vogwill, T. & MacLean, R.C. The genetic basis of the fitness costs of antimicrobial  
979 resistance: a meta-analysis approach. *Evol Appl*. **8**, 284–95 (2015).  
980
- 981 22. Berger, P., et al. Generalized Additive Model Demonstrates Fluoroquinolone  
982 Use/Resistance Relationships for *Staphylococcus aureus*. *Eur J Epidemiol*. **19**, 453–60  
983 (2004).  
984
- 985 23. Lawes, T., et al. Effects of national antibiotic stewardship and infection control strategies  
986 on hospital-associated and community-associated methicillin-resistant *Staphylococcus aureus*  
987 infections across a region of Scotland: a non-linear time-series study. *Lancet Infect Dis*.  
988 **15**, 1438–49 (2015).  
989
- 990 24. Lawes, T., et al. Turning the tide or riding the waves? Impacts of antibiotic stewardship  
991 and infection control on MRSA strain dynamics in a Scottish region over 16 years: non-linear  
992 time series analysis. *BMJ Open*, **5**, e006596 (2015).  
993

- 994 25. Lawes, T., et al. Effect of a national 4C antibiotic stewardship intervention on the clinical  
995 and molecular epidemiology of *Clostridium difficile* infections in a region of Scotland: a non-  
996 linear time-series analysis. *Lancet Infect Dis.* **17**, 194–206 (2017).  
997
- 998 26. Wong A. Epistasis and the Evolution of Antimicrobial Resistance. *Front Microbiol.* **8**,  
999 246 (2017).  
1000
- 1001 27. Levin BR, Perrot V, Walker N. Compensatory mutations, antibiotic resistance and the  
1002 population genetics of adaptive evolution in bacteria. *Genetics* **154**, 985–97 (2000).  
1003
- 1004 28. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse  
1005 resistance? *Nat Rev Microbiol.* **8(4)**, 260–71 (2010).  
1006
- 1007 29. Mózes J, Ebrahimi F, Gorácz O, Miszti C, Kardos G. Effect of carbapenem consumption  
1008 patterns on the molecular epidemiology and carbapenem resistance of *Acinetobacter*  
1009 *baumannii*. *J Med Microbiol.* **63**, 1654–62 (2014).  
1010
- 1011 30. Díaz, M.A., et al. Diversity of *Escherichia coli* Strains Producing Extended-Spectrum  $\beta$ -  
1012 Lactamases in Spain: Second Nationwide Study. *J Clin Microbiol.* **48(8)**, 2840–5 (2010).  
1013
- 1014 31. Rodríguez-Baño J, Alcalá JC, Cisneros JM, Grill F, Oliver A, Horcajada JP, Tórtola T,  
1015 Mirelis B, Navarro G, Cuenca M, Esteve M, Peña C, Llanos AC, Cantón R, Pascual A.  
1016 Community infections caused by extended-spectrum beta-lactamase-producing *Escherichia*  
1017 *coli*. *Arch Intern Med.* **168(17)**, 1897-902 (2008).  
1018
- 1019 32. Branger, C., et al. Genetic Background of *Escherichia coli* and Extended-spectrum  $\beta$ -  
1020 Lactamase Type. *Emerg Infect Dis.* **11(1)**, 54–61 (2005).  
1021
- 1022 33. Rodríguez-Baño, J., et al. Impact of changes in CLSI and EUCAST breakpoints for  
1023 susceptibility in bloodstream infections due to extended-spectrum  $\beta$ -lactamase-producing  
1024 *Escherichia coli*. *Clin Microbiol Infect Dis.* **18(9)**, 894–900 (2012).  
1025
- 1026 34. Poole, K. *Pseudomonas aeruginosa*: Resistance to the Max. *Frontiers in Microbiology* **2**,  
1027 65. doi:10.3389/fmicb.2011.00065 (2011).  
1028
- 1029 35. Dubois, V., et al. Beta-lactam and aminoglycoside resistance rates and mechanisms  
1030 among *Pseudomonas aeruginosa* in French general practice (community and private  
1031 healthcare centres). *J Antimicrob Chemother.* **62**:316–23 (2008).  
1032
- 1033 36. Hocquet, D., et al. Relationship between antibiotic use and incidence of MexXY-OprM  
1034 overproducers among clinical isolates of *Pseudomonas aeruginosa*. *Antimicrob Agents*  
1035 *Chemother.* **52**, 1173–5 (2008).  
1036
- 1037 37. Horváth, A., et al. Varying fitness cost associated with resistance to fluoroquinolones  
1038 governs clonal dynamic of methicillin-resistant *Staphylococcus aureus*. *Eur J Clin Microbiol*  
1039 *Infect Dis* **31**:2029–3 (2012)  
1040
- 1041 38. Laxminarayan, R., & Brown, G.M. Economics of Antibiotic Resistance: A Theory of  
1042 Optimal Use. *Journal of Environmental Economics and Management* **42**, 183-206 (2001).  
1043



1044 39. The Health Foundation. On targets: How targets can be most effective in the English  
1045 NHS. the Health Foundation, London (2015).  
1046  
1047 40. Hulscher, M.E., Grol, R.P., van der Meer J.W. Antibiotic prescribing in hospitals: a social  
1048 and behavioural scientific approach. *Lancet Infect Dis.*, **10(3)**167-75. (2010)  
1049  
1050 41. Laxminarayan, R., & Klugman, K.P. Communicating trends in resistance using a drug  
1051 resistance index. *BMJ Open* **1**, e000135 (2011).  
1052  
1053 42. Hughes, J.S., Hurford, A., Finley, R.L., et al. How to measure the impacts of antibiotic  
1054 resistance and antibiotic development on empiric therapy: new composite indices. *BMJ Open*  
1055 **6**, e012040. (2016)  
1056  
1057 43. Murphy, T.E., Van Ness, P.H., Araujo, K.L.B., Pisani, M.A. Bayesian Time-Series  
1058 Analysis of a Repeated-Measures Poisson Outcome With Excess Zeroes. *American Journal*  
1059 *of Epidemiology* **174(11)**, 1230-1237 (2011)  
1060  
1061 44. Søgaard P. The epidemiology of antibiotic resistance in three species of the  
1062 *Enterobacteriaceae* and the relation to consumption of antimicrobial agents in Odense  
1063 University Hospital. *Dan Med Bull.***36(1)**:65-84 (1989).  
1064  
1065 45. Møller JK. Antimicrobial usage and microbial resistance in a university hospital during a  
1066 seven-year period. *J Antimicrob Chemother.***24(6)**:983-92 (1989).  
1067  
1068 46. Hastie, T. & Tibshirani, R. Generalized Additive Models. *Statist. Sci.* **1**, 297-310 (1986)  
1069  
1070 47. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Statist* **19**, 1-67 (1991)  
1071  
1072 48. Hansen, B.E. Sample splitting and threshold estimation. *Econometrica* **68(3)**, 575-603  
1073 (2000).  
1074  
1075 49. Donayre, L., Eo, Y., Morley, J. Improving likelihood-ratio-based confidence intervals for  
1076 threshold parameters in finite samples. *Studies in Nonlinear Dynamics and Econometrics*,  
1077 **22(1)**, Article number 20160084 (2018).  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091

1092 **Additional authors in THRESHOLDS study group**  
1093 María Núñez-Núñez<sup>14</sup>, Ana I. Suárez<sup>14</sup>, Michelle Thouverez<sup>5</sup>, Navarro Cots<sup>15</sup>, Emilio  
1094 Borrajo<sup>15</sup>, Carlos Devesa<sup>15</sup>, Joan Gregori<sup>15</sup>, Inmaculada García Cuello<sup>15</sup>, Isabel Pacheco<sup>15</sup>,  
1095 María Cerón<sup>15</sup>, Hajnalka Tóth<sup>16</sup>.

1096  
1097 14. Hospital Virgen de la Macarena, Sevilla, Spain  
1098 15. Hospital Vega Baja, Orihuela, Spain  
1099 16. Department of Medical Microbiology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary  
1100

### 1101 **Supplementary information**

1102 Is available in the online version of the article  
1103 Correspondence and requests for data, R-scripts or other study materials can be addressed to  
1104 Dr Timothy Lawes, [t.lawes@nhs.net](mailto:t.lawes@nhs.net)  
1105

### 1106 **Acknowledgements**

1107 All authors acknowledge Dominique L. Monnet, from the European Centre for Disease  
1108 Prevention and Control (ECDC) for his continuous support, guidance and intellectual  
1109 contributions to the THRESHOLDS project. J-M.L-L., acknowledges the continuous support  
1110 of the Management Team of the Hospital Vega Baja, and the technical support of Carlos  
1111 Quiles (Cabosoft SL), in the development of time-series analysis techniques in the field of  
1112 antimicrobial resistance, and in particular the WebResist project ([www.webresist.org](http://www.webresist.org)) which  
1113 provided a framework for the development of this study. G.K., was supported by a Bolyai  
1114 Research Scholarship of the Hungarian Academy of Sciences. T.L., was supported by the  
1115 Wellcome Trust. J.R-B., and P.R., received funding for research from Plan Nacional de  
1116 I+D+i 2013-2016 and Instituto de Salud Carlos III, Subdirección General de Redes y Centros  
1117 de Investigación Cooperativa, Ministerio de Economía, Industria y Competitividad, Spanish  
1118 Network for Research in Infectious Diseases (REIPI RD16/0016/0001) - co-financed by  
1119 European Development Regional Fund “A way to achieve Europe”, Operative Programme  
1120 Intelligent Growth 2014-2020.  
1121

### 1122 **Author contributions**

1123 J-M.L-L., T.L., C.N., A.B., and I.M.G., proposed the original idea and designed the study.  
1124 X.B., D.H., M.A., G.C-B., M.S., D.F., G.K., J.R-B., P.R., and N.G., collated centre-specific  
1125 data, situational analysis, and hypotheses. J-M.L-L., T.L., C.N., A.B., contributed to  
1126 statistical analysis, with C.N. and A.B, the principal analysts. All authors discussed the  
1127 results and commented on the manuscript.  
1128

### 1129 **Competing interests**

1130 IMG is in receipt of payments for consultancies and lectures from numerous Pharmaceutical  
1131 firms developing new antimicrobials. Otherwise, the authors declare no competing interests.  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139

1140 **Figure legends**

1141

1142 **Figure 1 | Carbapenem-resistant *A. baumannii* and antibiotic use**

1143 **(a)** Time series for observed and predicted incidence density of CRAb, with observed  
1144 incidence density of CSAb. **(b)** Time series for use of potential explanatory antibiotic groups  
1145 (5-month moving averages). **(c)** Contribution charts illustrating the relationship between  
1146 explanatory variables and CRAb incidence density (sample size, n=143 months of  
1147 observation).

1148 <sup>a</sup> Change relative to median monthly CRAb incidence density for study period.

1149 CRAb, carbapenem-resistant *A. baumannii*. CSAb, carbapenem-susceptible *A. baumannii*.

1150 DDDs, defined daily doses. IQR, Interquartile range. MARS, Multivariate Adaptive

1151 Regression Splines. NL-TSA, non-linear time series analysis. OBDs, occupied bed days.

1152

1153 **Figure 2 | Extended-spectrum  $\beta$ -lactamase (ESBL)-producing *Escherichia coli* and  
1154 antibiotic use in hospital and community.**

1155 **(a)** Time series for observed and predicted %Ec-ESBL+ in hospital population. **(b)** Time  
1156 series of potential explanatory antibiotic groups in hospital population (5-month moving  
1157 averages). **(c)** Time series for observed and predicted %Ec-ESBL+ in community population.  
1158 **(d)** Time series of potential explanatory antibiotic groups in community population (5-month  
1159 moving averages). **(e)** Contribution charts illustrating the relationship between explanatory  
1160 variables and hospital or community %Ec-ESBL+. (sample size, n=304 months of  
1161 observation).

1162 <sup>a</sup> compared to median monthly %Ec-ESBL+ for study period.

1163 \* confidence interval around lower threshold in association between fluoroquinolones and  
1164 %Ec-ESBL+ in hospital may be wider due to risk of slight under-coverage of confidence  
1165 interval estimation arising from a problem of multiple closely occurring thresholds.

1166 %Ec-ESBL+, percentage of *E. coli* isolates producing extended-spectrum beta-lactamases.

1167 DDDs, defined daily doses. IQR, Interquartile range. MARS, Multivariate Adaptive

1168 Regression Splines. NL-TSA, non-linear time series analysis. OBDs, occupied bed days.

1169

1170 **Figure 3 | Cefepime-resistant *Escherichia coli* and antibiotic use**

1171 **(a)** Time series for observed and predicted incidence density of %Ec-FepR. **(b)** Time series  
1172 of potential explanatory antibiotic groups (5-month moving averages). **(c)** Contribution charts  
1173 illustrating the relationship between explanatory variables and %Ec-FepR (sample size,  
1174 n=105 months of observation).

1175 <sup>a</sup> compared to median monthly %Ec-FepR for study period.

1176 DDDs, Defined Daily Doses. IQR, interquartile range. %Ec-FepR, percentage of *E. coli*

1177 isolates resistant to cefepime. OBDs, Occupied Bed Days.

1178

1179 **Figure 4: Gentamicin-resistant *Pseudomonas aeruginosa* and antibiotic use**

1180 **(a)** Time series for observed and model predicted incidence density of GRPa. **(b)** Time series  
1181 of potential explanatory antibiotic groups (5-month moving averages) **(c)** Contribution charts  
1182 illustrating the relationship between explanatory variables and GRPa incidence density

1183 <sup>a</sup> compared to median monthly GRPa incidence density for study period (sample size, n=192  
1184 months of observation).

1185 DDDs, Defined Daily Doses. GRPa, gentamicin-resistant *P. aeruginosa*; GSPa, gentamicin-

1186 susceptible *P. aeruginosa* isolates. IQR, Interquartile range. MARS, multivariate adaptive

1187 regression splines OBDs, Occupied Bed Days.

1188

1189

1190 **Figure 5: Methicillin-resistant *Staphylococcus aureus*, hand hygiene, and antibiotic use**  
1191 **(a)** Time series for observed and predicted incidence density of MRSA. **(b)** Time series of  
1192 potential explanatory antibiotic groups (5-month moving averages). **(c)** Time series for  
1193 alcohol-based hand rub (ABHR) use. **(d)** Contribution charts illustrating the relationship  
1194 between explanatory variables and MRSA incidence density (sample size, n=105 months of  
1195 observation).  
1196 <sup>a</sup> compared to median monthly MRSA incidence density for study period.  
1197 ABHR, alcohol-based hand rub. DDDs, Defined Daily Doses. ECDC, European Centre for  
1198 Disease Prevention and Control. MRSA, Methicillin-resistant *Staphylococcus aureus*. OBDs,  
1199 Occupied Bed Days.

1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239

1240 **Tables**

1241

1242 **Table 1: Translation of thresholds identified in non-linear models into population-**  
 1243 **specific antibiotic stewardship policy recommendations**

Antibiotic	Patient treatments per month <sup>a</sup>			
	Maximum suggested by threshold (95% CI)	Average use in last 12 months of study	Suggested reduction in use (%)	
			Standard (using point estimate for threshold)	Conservative (using lower limit of 95% CI for threshold)
<b>1. Carbapenem-resistant <i>Acinetobacter baumannii</i> (CRAB)‡ Debrecen, Hungary</b>				
Carbapenems	86 (45 to 134)	226	140 (62%)	181 (80%)
3 <sup>rd</sup> generation cephalosporins	203 (169-229)	299	96 (32%)	130 (57%)
Fluoroquinolones	478 (367-576)	375	Maintain below threshold	9 (2%)
Piperacillin-tazobactam	39 (28-78)	41	2 (5%)	13 (32%)
<b>2. Extended-spectrum <math>\beta</math>-lactamase producing <i>Escherichia coli</i> (%Ec-ESBL+) Orihuela, Spain</b>				
<i>a) Hospital population</i>				
Fluoroquinolones	78 (67 to 96) <sup>b</sup>	165	68 (41%)	87 (53%) <sup>b</sup>
Third-generation cephalosporins	62 (35 to 91)	78	16 (21%)	43 (55%)
<i>b) Community population</i>				
Fluoroquinolones	80 (40 to 148)	66	Maintain below threshold	26 (39%)
Co-amoxiclav	191 (142 to 284)	277	86 (31%)	135 (49%)
<b>3. Cefepime-resistant <i>Escherichia coli</i> (%EcFepR) Seville, Spain</b>				
Fluoroquinolones	392 (225 to 426)	351	Maintain below threshold	126 (36%)
3 <sup>rd</sup> /4 <sup>th</sup> generation cephalosporins	130 (83 to 158)	211	87 (41%)	128 (61%)
<b>4. Gentamicin-resistant <i>Pseudomonas aeruginosa</i> (GRPa), France</b>				
Gentamicin and tobramycin	75 (36-99) <sup>b</sup>	68 <sup>c</sup>	Maintain below threshold	32 (47%)
Fluoroquinolones	324 (316 to 330)	223	Maintain below threshold	Maintain below threshold
<b>5. Methicillin-resistant <i>Staphylococcus aureus</i> (MRSA) Antrim, Northern Ireland</b>				
Fluoroquinolones	24 (20 to 29)	39	15 (39%)	19 (49%)
3 <sup>rd</sup> generation cephalosporins	7 (6 to 8)	6	Maintain below threshold	Maintain below threshold
Co-amoxiclav	320 (189 to 422)	320	Maintain below threshold	131 (41%)

1244

1245

1246

1247

1248

1249

a Derived by multiplying the threshold in table 1, expressed in DDDs per 1000 OBDs, by the size of the population (in 1000 OBDs or IDs), and then dividing by an average patient treatment (considered as 7 DDDs unless otherwise specified)

b Confidence intervals may be wider, and the recommended restriction greater in the conservative approach, due to risk of slight undercoverage of estimated confidence intervals due to multiple closely occurring thresholds in association between fluoroquinolone and %Ec-ESBL+ in hospital.

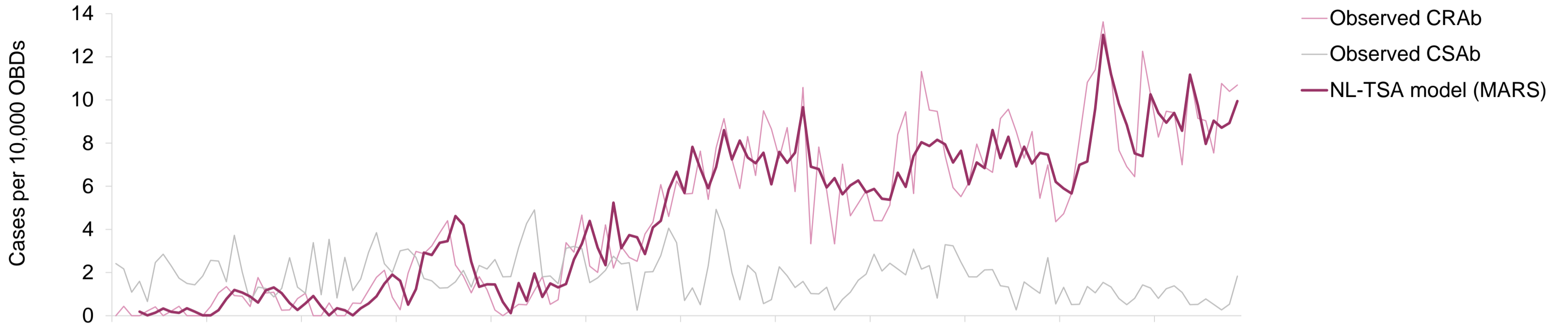
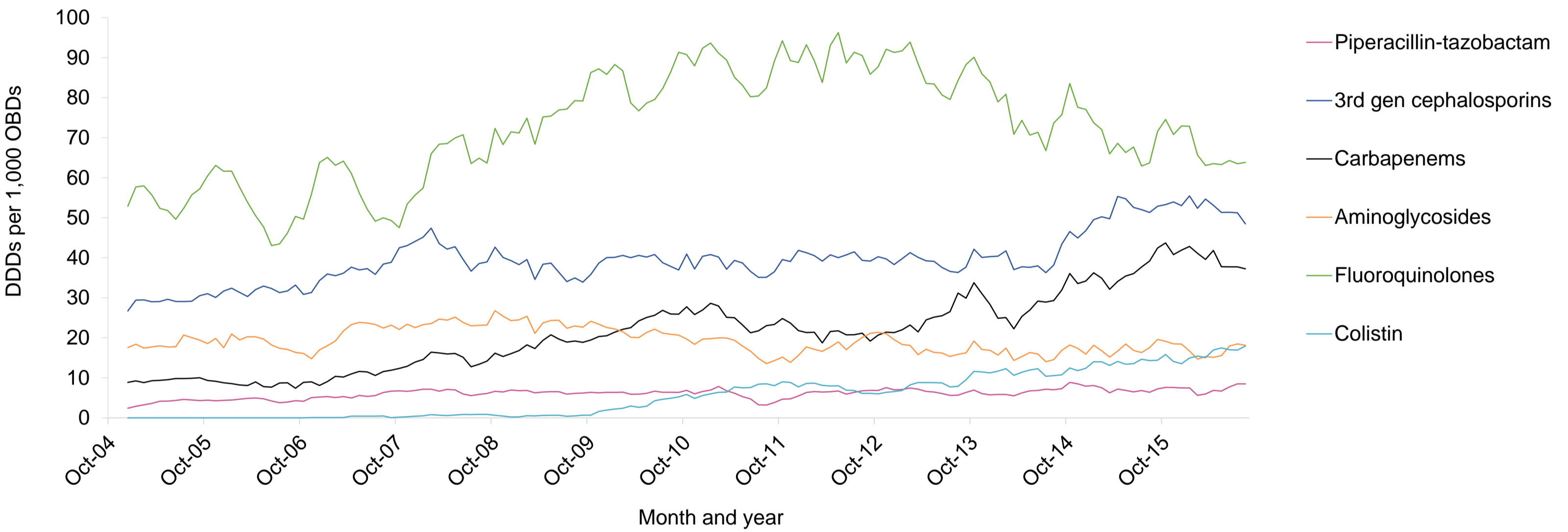
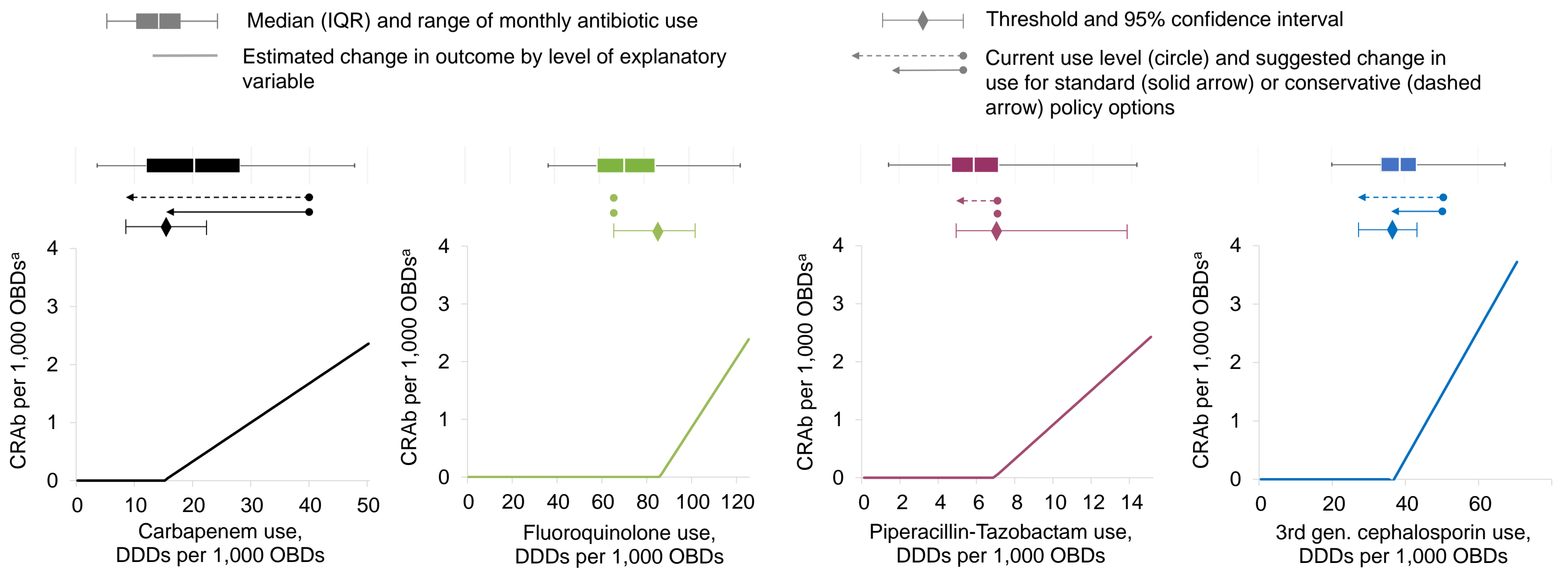
c Average treatment course considered as 3 DDDs.

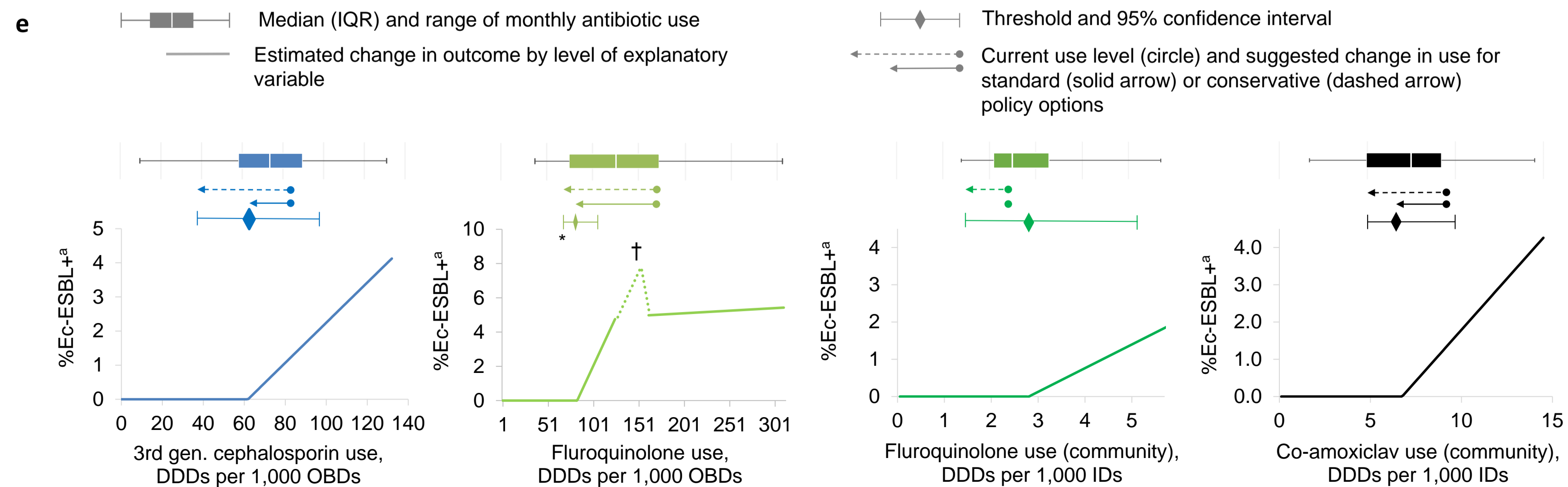
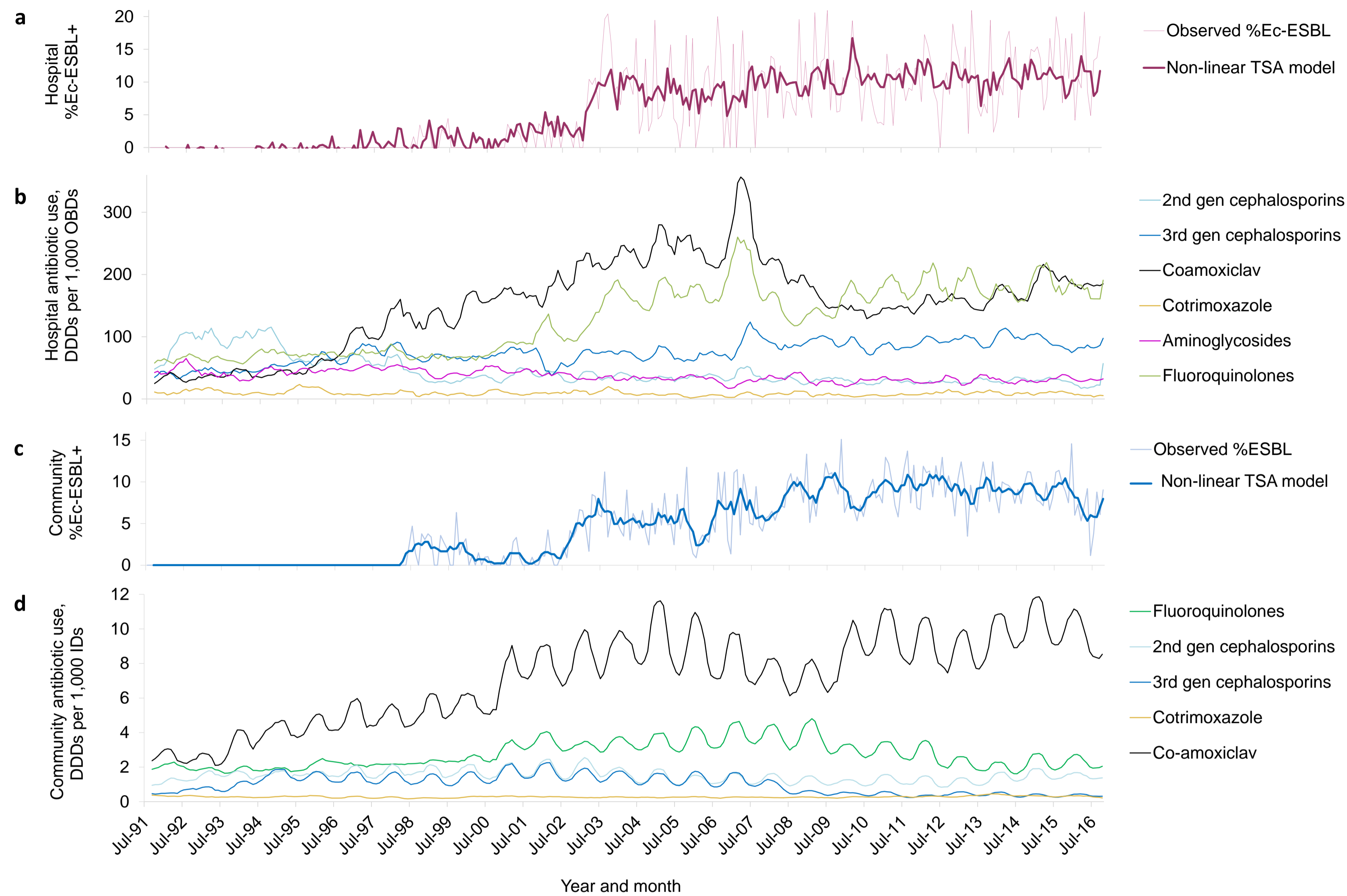
1250

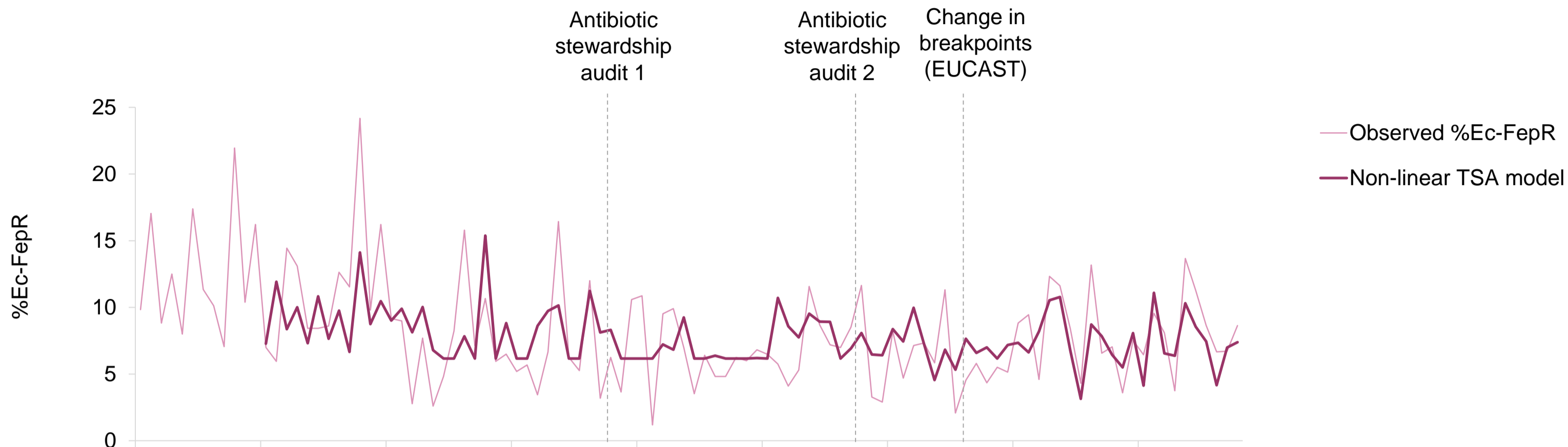
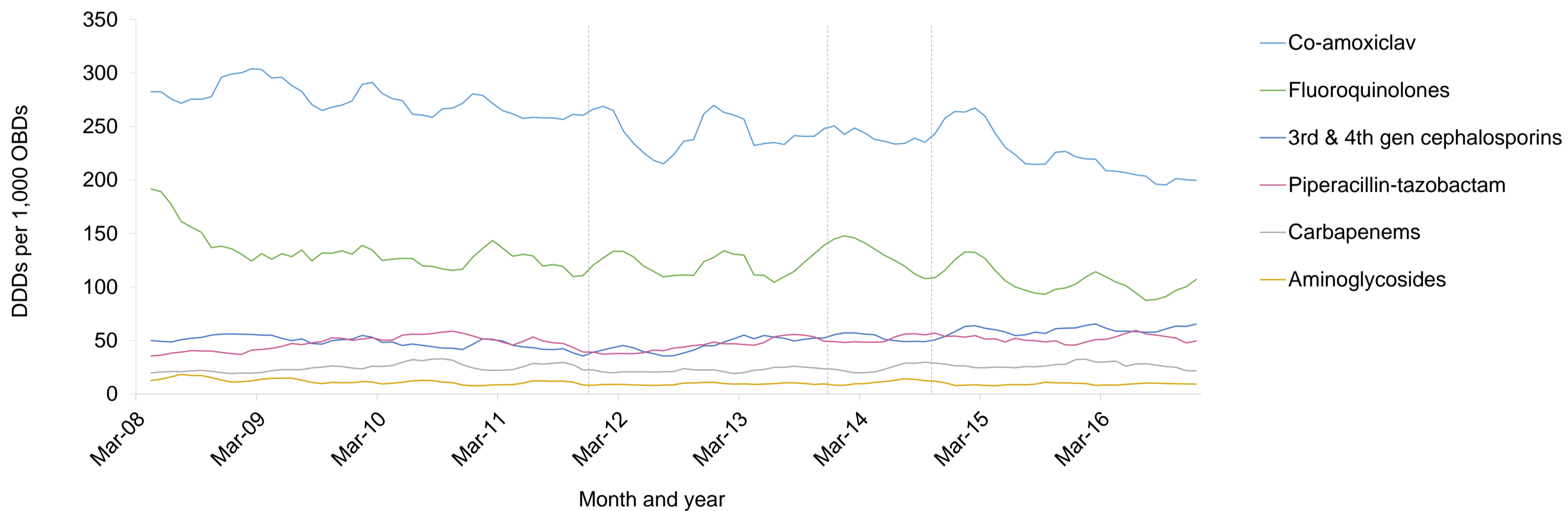
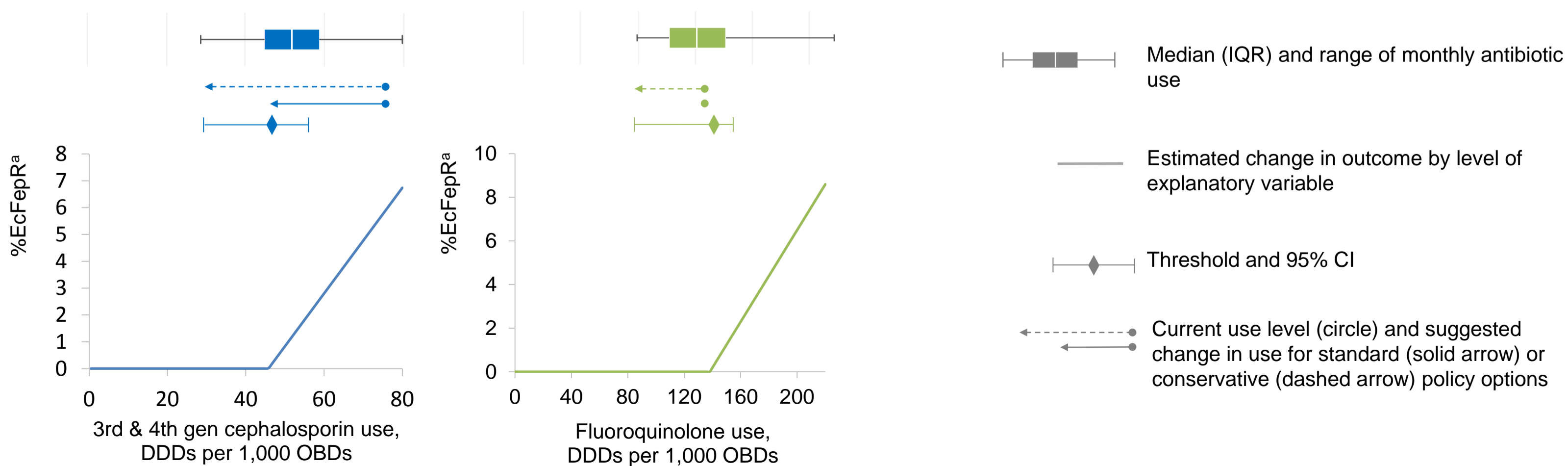
1251

1252

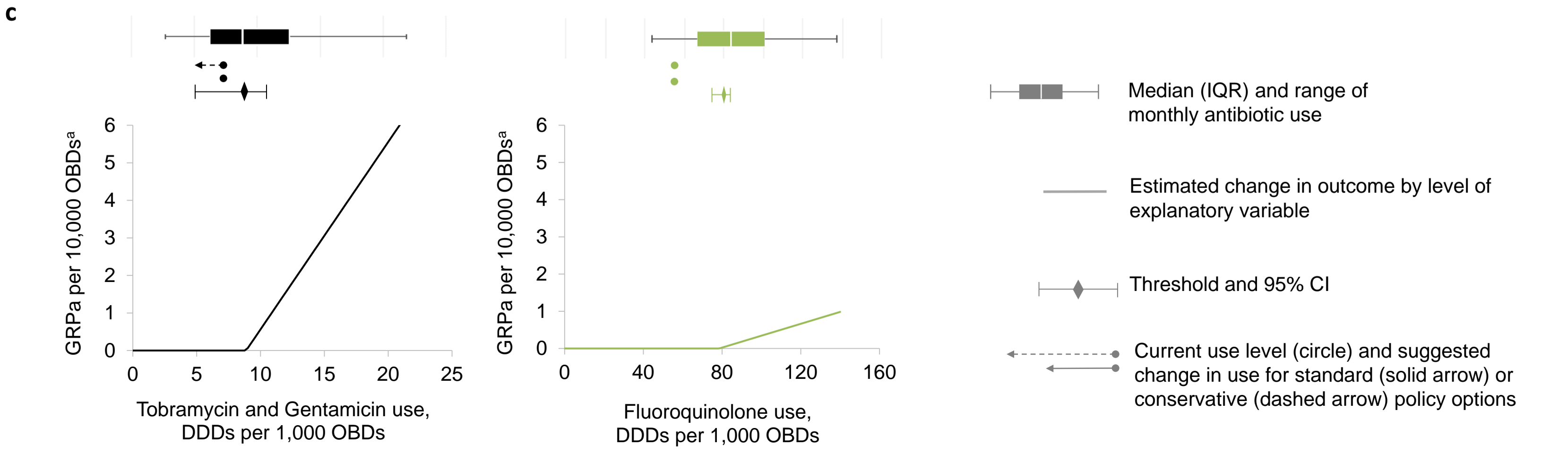
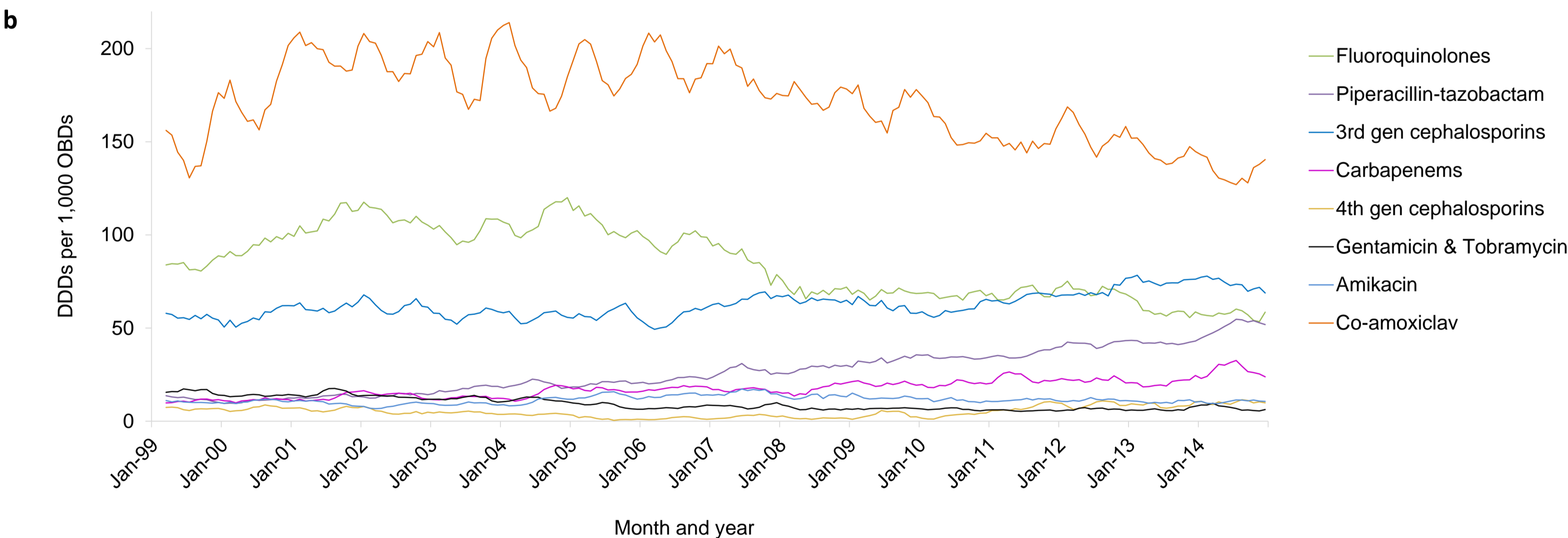
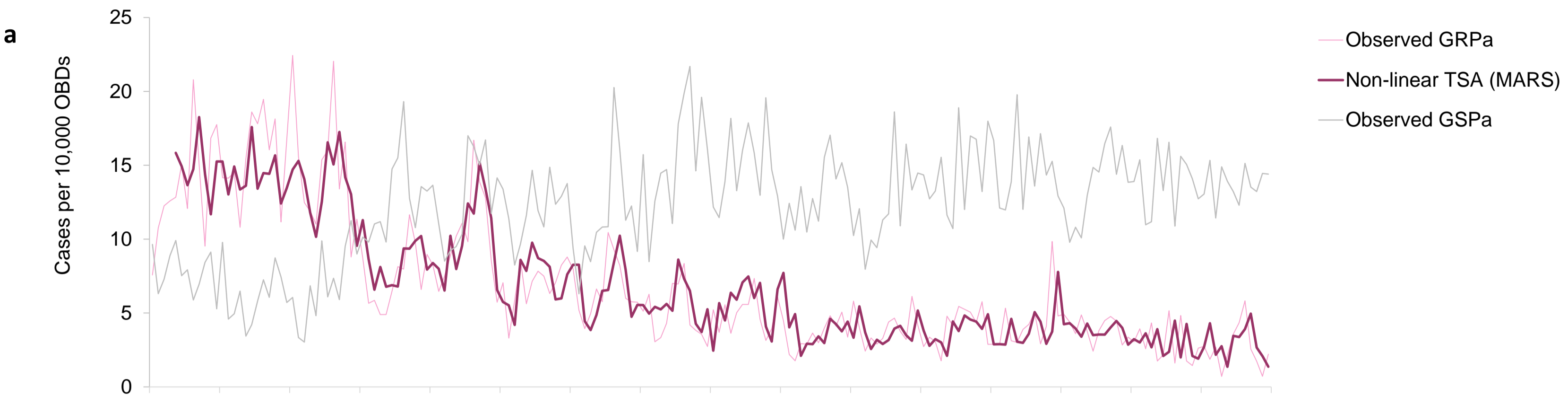
1253

**a****b****c**



**a****b****c**





Restriction of Fluoroquinolones

