

1 Title:  
2 Understanding the Effects of Sub-Inhibitory Antibiotic Concentrations on the Development of  $\beta$ -  
3 Lactamase Resistance Based on Quantile Regression Analysis  
4

5 Authors:  
6 Portia Mira<sup>1\*</sup>, Candace Guzman-Cole<sup>2</sup>, Juan C. Meza<sup>2</sup>  
7

8 Affiliations:  
9 <sup>1</sup> Department of Microbiology, Immunology and Molecular Genetics, University of California,  
10 Los Angeles  
11 <sup>2</sup>University of California, Merced  
12

13 \*Corresponding author: portiamira20@gmail.com  
14

15 Abstract:

16 Aims:

17 Quantile regression is an alternate type of regression analysis that has been shown to have  
18 numerous advantages over standard linear regression. Unlike linear regression, which uses the  
19 mean to fit a linear model, quantile regression uses a data set's quantiles (or percentiles), which  
20 leads to a more comprehensive analysis of the data. However, while relatively common in other  
21 scientific fields such as economic and environmental modeling, it is infrequently used to  
22 understand biological and microbiological systems.  
23

24 Methods and Results:

25 We analyzed a set of bacterial growth rates using quantile regression analysis to better understand  
26 the effects of antibiotics on bacterial fitness. Using a bacterial model system containing 16 variant  
27 genotypes of the TEM  $\beta$ -lactamase enzyme, we compared our quantile regression analysis to a  
28 previously published study that uses the Tukey's range test, or Tukey Honestly Significantly  
29 Different (HSD) test. We find that trends in the distribution of bacterial growth rate data, as viewed  
30 through the lens of quantile regression, can distinguish between novel genotypes and ones that  
31 have been clinically isolated from patients. Quantile regression also identified certain  
32 combinations of genotypes and antibiotics that resulted in bacterial populations growing faster as  
33 the antibiotic concentration increased- the opposite of what is expected. These analyses can  
34 provide new insights into relationships between enzymatic efficacy and antibiotic concentration.  
35

36 Conclusions:

37 Quantile regression analysis enhances our understanding of the impacts of sublethal antibiotic  
38 concentrations on enzymatic (TEM  $\beta$ -lactamase) efficacy and bacterial fitness. We illustrate that  
39 quantile regression analysis can link patterns in growth rates with clinically relevant mutations and  
40 provides an understanding of how increasing sub-lethal antibiotic concentrations, like those found

41 in our modern environment, can affect bacterial growth rates and provide insight into the genetic  
42 basis for varied resistance.

43  
44

45 **Significance:**

46 Quantile regression analysis provides a distinctive way of analyzing bacterial growth rate data that  
47 has not yet been done and offers unique advantages to microbiologists. Identifying the mutations  
48 that are most likely to appear in the clinic can help scientists and epidemiologists better predict the  
49 directionality of antibiotic resistance and develop novel pharmaceuticals to combat this worldwide  
50 crisis.

51  
52

### 53 **Introduction**

54

55 Antibiotic resistance is a serious public health threat and responsible for over thirty-five-  
56 thousand deaths worldwide in 2019 (Arnold, Thom et al. 2011, Frieri, Kumar et al. 2017, Centers  
57 for Disease Control and Prevention 2019, Iwu, Korsten et al. 2020, Kaur, Pham et al. 2021). Due  
58 to the abuse of antibiotics, multi-drug-resistant pathogens have been rising across the globe  
59 (Zhuang, Achmon et al. 2021). Bacteria are regularly exposed to sub-inhibitory concentrations of  
60 antibiotics found in wastewater from agriculture (Leiva, Pina et al. 2021, Sanz, Casado et al.  
61 2021) and from clinics (Bojar, Sheridan et al. 2021). These sub-inhibitory concentrations of  
62 antibiotics accelerate the diversification of resistance genes (Li, Phulpoto et al. 2021, Li, Xia et  
63 al. 2021, Byun, Ha Han et al. 2022, Xu, Tan et al. 2022) and increase the prevalence of  
64 antibiotic-resistant pathogens more than lethal concentrations do (Mira, Meza et al. 2015, Mira,  
65 Østman et al. 2021, Byun, Ha Han et al. 2022).

66

67 B-lactam antibiotics are the most widely prescribed antibiotics worldwide because of their  
68 efficacy and low toxicity to human and animal organisms (Demain and Elander 1999). However,  
69 due to their abuse,  $\beta$ -lactam resistance is prominent. As a result, the most studied group of  
70 resistance genes are those that encode for  $\beta$ -lactamases, enzymes that cleave the active site of  $\beta$ -  
71 lactam antibiotics and render them inactive (Poole 2004). To combat increasing resistance,  
72 different types of  $\beta$ -lactam antibiotics have been developed that differ structurally but have the  
73 same mechanism of action, ranging from early penicillins to fourth generation cephalosporins  
74 and  $\beta$ -lactam with the  $\beta$ -lactamase inhibitors. Typically, the newer the antibiotic generation, the  
75 more complex the chemical structure and/or the more effective against bacteria that produce  $\beta$ -  
76 lactamases.

77

78 In an earlier study (Mira, Østman et al. 2021), we investigated the effects of sub-inhibitory  
79 concentrations of  $\beta$ -lactam antibiotics on two representatives from the penicillin group,  
80 representatives from all four generations of cephalosporins, and one penicillin with beta-

81 lactamase inhibitor on the selection of different combinations of the four point mutations within  
82 a TEM  $\beta$ -lactamase resistance gene, *bla*<sub>TEM-85</sub>. Briefly, (Mira, Østman et al. 2021) measured the  
83 growth rate of 16 TEM genotypes that contained different combinations of the four substitutions  
84 found in *bla*<sub>TEM-85</sub> in the presence of three sub-inhibitory concentrations of  $\beta$ -lactam antibiotics.  
85 Using the Tukey honestly significant difference (HSD) test, we calculated all pairwise p-values  
86 between growth rates across antibiotic concentrations. Surprisingly, we were able to identify  
87 significant increases in growth rates as antibiotic concentration increased for 10  $\beta$ -lactam  
88 antibiotics.

89

90 The Tukey HSD, like standard linear regression, relies on the assumption that variance is equal  
91 across the groups associated with each mean—also referred to as homogeneity of variance.  
92 However, errors are often non-normally distributed, which can happen when response variables  
93 behave differently at varying levels of the independent variable. For example, growth rates could  
94 differ markedly at low versus high concentrations of antibiotics due to the genetic background of  
95 the bacterial population, mutation rates, or limiting effects of the antibiotic. For this reason, we  
96 re-examine our data from this earlier study using quantile regression analysis.

97

98 Quantile regression analysis is an alternative approach to studying the relationships between  
99 response and independent variables first proposed by Koenker (Koenker 1978). As the name  
100 implies, quantile regression uses quantiles (also known as percentiles) within a dataset, rather  
101 than just the mean, and can provide a more comprehensive analysis (De Oliveira, Fischmeister et  
102 al. 2013, Chen, Bertke et al. 2021). In clinical studies, quantile regression can identify trends not  
103 otherwise seen in standard linear regression, such as differences between high- and low-risk  
104 patients (Staffa, Kohane et al. 2019). Quantile regression has been particularly useful in ecology  
105 (Thomson, Weiblen et al. 1996), where limiting factors (uncontrolled factors that are not  
106 measured but nonetheless influence the response variable) arise in dealing with ecological and/or  
107 biological datasets. In our case, limiting factors include the mutations, and different  
108 combinations of mutations, that are present in the TEM  $\beta$ -lactamase gene. In spite of its many  
109 advantages, only a few studies have applied quantile regression to biological datasets (Villain,  
110 Lozano et al. 2014). In addition, we are unaware of any study that uses quantile regression  
111 analysis on microbial growth rates over time or considers limiting factors on microbial growth  
112 rate, even though limiting factors are prevalent in microbial fitness datasets.

113

114 In this study, we analyze previously published bacterial growth rate data (Mira, Østman et al.  
115 2021) to investigate whether quantile regression can enhance our insight into the effects of sub-  
116 inhibitory drug concentrations on the selection of point mutations within a TEM  $\beta$ -lactamase  
117 resistance gene, *bla*<sub>TEM-85</sub>. Specifically, we use quantile regression analysis to investigate the  
118 effects of sub-inhibitory antibiotic concentrations on the evolution of TEM resistance genes,  
119 which other approaches that assume normally distributed errors may have missed. The  
120 differences across quantiles provide a more complete picture of the association between

121 antibiotic concentration and growth rates and how point mutations affect enzymatic efficacy,  
122 which may be linked to bacterial fitness. Overall, these analyses can lead to an improved  
123 understanding of the impact of sublethal concentrations of antibiotics in the environment on  
124 bacterial fitness.

125

126

## 127 **Methods**

128

### 129 **Quantile Regression Analysis**

130 To investigate whether quantile regression analysis provides more insight into bacterial growth  
131 rates than standard linear regression, we used a dataset containing growth rates of variant TEM  
132 genotypes on the pBR322 plasmid expressed in *E. coli* strain DH5 $\alpha$ E (Table 1) (Goulart,  
133 Mahmudi et al. 2013). Specifically, we used all 16 combinations of the four amino acid  
134 substitutions present on the  $\beta$ -lactamase gene *bla*<sub>TEM-85</sub> (Mira, Østman et al. 2021), only 10 of  
135 which have been clinically identified (Jacoby 2020)(Table 1). These genotypes range from the  
136 wild-type TEM-1, that confers resistance to only penicillins, four individual single substitutions  
137 that each carry their own resistance phenotype based on the location of the substitution in the  
138 TEM enzyme, the six possible combinations of double mutations, four possible triple  
139 combinations and finally TEM-85 that contains all four substitutions and has evolved to confer  
140 resistance to penicillins as well as higher generations of cephalosporins (Mira, Østman et al.  
141 2021) (Table 1). It is important to note that these genotypes all carry individualized efficacy to  
142 different generations of  $\beta$ -lactam antibiotics depending on the location of the point mutation  
143 within the TEM enzyme and any interactions the amino acid residues have with one another. We  
144 exposed each of the 16 variants to various sublethal concentrations of 12  $\beta$ -lactam antibiotics for  
145 22 hours (Table 2). The  $\beta$ -lactam antibiotics represented all generations, from early penicillins to  
146 fourth generation cephalosporins and  $\beta$ -lactam + inhibitor combinations. The growth rates were  
147 calculated using the *growthrates* package (Hall, Acar et al. 2014).

148

### 149 **Creating quantile regression plots**

150 We first normalized and plotted the growth rate data to the highest concentration of each  
151 antibiotic, scaling the growth rate values by a factor of 1000 to make the analysis easier and to  
152 control for numerical ill-conditioning (top panel, Figure 1). We then calculated the slopes of the  
153 lines of best fit through each quantile (20<sup>th</sup>, 40<sup>th</sup>, 60<sup>th</sup>, and 80<sup>th</sup>). Using the *quantreg* package in  
154 R, we then generated quantile regression plots that represent these slopes at each quantile  
155 (bottom panel, Figure 1).

156

### 157 **Identifying genotypes with non-normal distributions of growth rate**

158 Using the quantile regression plots for each antibiotic, we identified genotypes that had  
159 differences in distribution—i.e., those in which quantile regression analysis would be most  
160 useful. For the purposes of this study, we focus on non-normally distributed data to better

161 understand effects across quantiles. We visualized quantile regression plots using the *rq* function  
162 from Koenker's *quantreg* package in R and identified four types of trends, representing how the  
163 distribution of the data changes across quantiles (Figure 2). A constant trend (Figure 2, Panel A)  
164 is normally distributed data and represents genotypes that grew at the same rate- indicating stable  
165 enzymatic activity- across the concentrations of antibiotic. An increasing trend (Figure 2, Panel  
166 B) indicates non-normally distributed data at the higher quantiles and represents genotypes that  
167 led to an increase in enzymatic activity the faster the bacteria grew. Contrarily, a decreasing  
168 trend (Figure 2, Panel C) indicates non-normally distributed data at the higher quantiles and  
169 represents genotypes that led to a decrease in enzymatic activity the faster the bacteria grew  
170 Finally, a U-shaped trend (Figure 2, Panel D) indicates non-normally distributed data at both the  
171 higher and/or lower quantiles with a dip in the middle quantiles. This type of trend represents  
172 genotypes that led to fluctuations of enzymatic activity as the antibiotic concentrations increased.  
173 It is important to note the sign of the slopes (y-axis). In some cases, for any trend, the slopes at  
174 quantiles are positive, indicating that as the antibiotic concentration increased, so did the growth  
175 rates. These distribution patterns in the quantile regression plots can give us insight into the  
176 efficacy of the TEM enzyme in the presence of increasing concentrations of certain antibiotics.  
177

## 178 **Results**

179 In order to test the hypotheses that analyzing different quantiles of bacterial growth might  
180 provide a more complete picture of the relationship between antibiotic concentrations and  
181 bacterial fitness, we examined distribution patterns in the quantile regression plots across  
182 genotypes (Figure 3). Across all genotypes (gray bars, Figure 3), we found ~32% constant  
183 trends, meaning that growth rate data was normally distributed across increasing antibiotic  
184 concentrations in 32% of antibiotic treatments. Genotypes that have been clinically identified  
185 account for 71% of the normally distributed data. However, most of the data (~68%) showed  
186 non-normal distributions (i.e., increasing, decreasing, or U-shape trends) shown by the green,  
187 red, and yellow bars, respectively (Figure 3).  
188

189 To investigate whether certain antibiotics or  $\beta$ -lactam generations produce more non-normal  
190 distributions in growth rate data, we visualized the trend for each genotype-antibiotic pair  
191 (Figure 4). We found that ampicillin (AMP) and ampicillin + sulbactam and piperacillin +  
192 tazobactam (SAM, TZP) had largely decreasing trends. AMP, SAM, and TZP had more of an  
193 impact on growth rates in higher quantiles than lower quantiles, which means that the faster the  
194 bacteria grew, the more the antibiotic treatment inhibited their growth.  
195

196 Second-generation cephalosporins cefaclor, cefotetan and cefuroxime (CEC, CTT, and CXM)  
197 resulted in mostly increasing trends. These antibiotics had less of an impact on growth rates at  
198 higher quantiles compared to lower quantiles. In other words, the slower the bacteria grew, the  
199 more the antibiotic inhibited their growth. Cefaclor (CEC) showed mostly increasing trends  
200 across all genotypes, specifically in genotypes that encode L21F and/or R164S substitutions

201 (L21F, R164S, LR, RE, RT, LRT, RET, and TEM-85). Cefotetan (CTT) provided similar results,  
202 with largely increasing trends for amino acid substitutions L21F, E240K, and T265M (LE, LT,  
203 RT, LRE, LET, TEM-85).

204

205 Neither third- nor fourth-generation cephalosporins showed a prevalent pattern in their quantile  
206 regression trends. We noticed a similar number of increasing and decreasing trends, with slightly  
207 more decreasing trends overall. However, the genotype that carries the single amino acid  
208 substitution R164S exhibited increasing trends across 80% of third generation cephalosporins.  
209 We also noticed a shift in trends in the fourth-generation cephalosporin cefepime (FEP) as the  
210 number of substitutions increased (Figure 4). The pattern changed from consistent decreasing  
211 trends in TEM-1 wildtype, single-substitution, and double-substitution genotype variations to  
212 increasing trends in two triple-substitution genotypes and the four-substitution TEM-85. This  
213 indicates that the  $\beta$ -lactamase enzyme became more efficient at hydrolyzing FEP at higher  
214 concentrations as the number of amino acid substitutions increased.

215

216 Next, we compared the clinically isolated genotypes to those that have not yet been clinically  
217 isolated to see if there was any significant difference in quantile regression patterns that  
218 correlated to the TEM enzyme's ability to hydrolyze  $\beta$ -lactam antibiotics. We found that the  
219 genotypes that have not been clinically isolated had more decreasing trends across all antibiotics  
220 (46%) than clinically identified genotypes (26%; compare red boxes for genotypes without and  
221 with an asterisk in Figure 4). Decreasing trends indicate that the faster the bacterial populations  
222 grew, the more inefficient the TEM enzyme was at hydrolyzing the antibiotics or the more  
223 effective the antibiotics were at inhibiting growth. However, accumulated mutations in  
224 populations with decreasing trends may have resulted in faster growth unrelated to the TEM  
225 enzyme's ability to hydrolyze antibiotics. Clinically and non-clinically isolated genotypes had  
226 similar frequencies of increasing trends (22% and 26%). In these cases, when the bacterial  
227 populations grew faster, the TEM enzyme was able to sufficiently hydrolyze increasing  
228 antibiotic concentrations.

229

230 Finally, we identified patterns in genotypes or antibiotics that resulted in a positive slope across  
231 some or all quantiles. A positive slope indicates that growth rates increased along with antibiotic  
232 concentration - the opposite of what is expected when exposing bacterial populations to  
233 antibiotics. Apart from cefotetan and ceftazidime (CTT, CAZ), all other antibiotics had at least  
234 one genotype that exhibited a positive slope (Figure 4). Approximately 27% of all treatments (13  
235 antibiotic treatments X 16 genotypes) had positive slopes at some or all quantiles. Half of these  
236 positive slopes occurred across all quantiles, 29% at higher quantiles (50<sup>th</sup> to 90<sup>th</sup>) and 21% at  
237 lower quantiles (10<sup>th</sup> to 40<sup>th</sup>) (Figure 4). There was no direct correlation between the quantile  
238 regression trends and positive slopes. However, nearly all genotypes (12/16) had a positive slope  
239 in part or all the quantiles in the presence of the only penicillin we tested, ampicillin (AMP). We

240 also noticed genotypes that have been clinically identified had positive slopes across all quantiles  
241 more often than non-clinically identified genotypes.

242

## 243 **Discussion**

244 In contrast to our earlier study (Mira, Ostman et al. 2021), which assumed homogeneity of  
245 variance and used the Tukey HSD test, our quantile regression analysis of the same dataset has  
246 provided more robust findings related to amino acid substitutions and enzyme efficacy, antibiotic  
247 class, and whether genotypes have been clinically identified.

248 Typically, as antibiotic concentration increases, bacterial growth rate decreases. However, some  
249 bacteria can hydrolyze antibiotics so quickly that they can use the degraded antibiotics as a  
250 carbon source for continued growth and proliferation (Dantas, Sommer et al. 2008). This is the  
251 case for some TEM genotypes (Mira, Meza et al. 2015, Mira, Østman et al. 2021). Authors  
252 previously showed that growth rate increased as antibiotic concentration increased in response to  
253 at least one antibiotic for all but two TEM genotypes: T265M and LET. However, our quantile  
254 regression analysis revealed that all genotypes had positive slopes across some or all quantiles,  
255 including T265M and LET- meaning the presence of these amino acid substitutions allowed the  
256 bacterial populations to grow faster as antibiotic concentration increased. LET only had positive  
257 slopes at the lower quantiles for the penicillin/inhibitor treatment piperacillin + tazobactam  
258 (TZP), which suggests that the slower the bacteria grew, the better the TEM enzyme with the  
259 L21F/ E240K/T265M substitutions was at hydrolyzing TZP and using the degraded antibiotic  
260 product as a carbon source for faster growth.

261 The location of amino acid substitutions within the TEM enzyme can influence the enzyme's  
262 efficacy, even if not located in the binding site. For example, the genotype with the single  
263 substitution T265M had positive slopes across all quantiles in the presence of ampicillin (AMP),  
264 in upper quantiles in the presence of cefuroxime (CXM), and in lower quantiles in the presence  
265 of cefprozil (CPR). CXM and CPR are second- and third generation cephalosporins, so their  
266 structures are more complex than AMP, a penicillin class antibiotic. The T265M substitution is  
267 distant from the binding site of the enzyme, positioned on the outer face of the  $\beta$ -sheet but buried  
268 by surrounding amino acids (Knox 1995). Recently, this single mutation has been identified  
269 clinically and has shown resistance to  $\beta$ -lactam inhibitor treatments like piperacillin +  
270 tazobactam (TZP) and resistance to some third-generation cephalosporins (Mulvey and Boyd  
271 2009). This clinical isolate also had a G162T mutation in promoter region *P4*, resulting in  
272 hyperproduction of the TEM gene; this could explain the resistance to  $\beta$ -lactamase inhibitor  
273 treatments (Mulvey and Boyd 2009). In this study we used *E. coli* strain K12, which lacks the *P4*  
274 region, potentially explaining why we did not see positive slopes for T265M in any  $\beta$ -lactam  
275 inhibitor treatments.

276 Our analysis reveals that TEM genotypes that have been identified in the clinic have more  
277 positive slopes across all antibiotic treatments than non-clinically identified genotypes,  
278 particularly in the presence of AMP. Genotypes are typically clinically identified when they are  
279 highly selected for in natural environments (Shawky, Suleiman et al. 2021), which can correlate  
280 with a more stable and effective enzyme. In fact, we find that clinical isolates had normal  
281 distributions across all quantiles in 71% of treatments, compared to only 32% among non-  
282 clinically isolated genotypes. This suggests that normal distributions across all quantiles  
283 correlates with more-stable enzymatic activity throughout populations.

284 R164S is the most common amino acid substitution observed within the more than 220 clinically  
285 identified TEM  $\beta$ -lactamases and is highly selected for because it increases enzyme efficacy in  
286 response to higher-generation cephalosporins. The 164<sup>th</sup> amino acid position is just below the  
287 binding site that contains the catalytic glutamate at position 166 (Herzberg 1991, Knox and  
288 Moews 1991). In R164S, serine (S) replaces arginine (R), resulting in a hydrogen-bonding amino  
289 acid with one less hydrogen bond donor (Palzkill, Le et al. 1994). This reduction in the number  
290 of hydrogen bonds eradicates electrostatic attraction and allows for more flexibility in the  
291 binding site, creating more space for the bulky sidechains of third- and fourth-generation  
292 cephalosporins (Escobar, Tan et al. 1994). We found increasing trends in R164S populations for  
293 five cephalosporin treatments—the highest number of increasing trends across all 16 genotypes  
294 tested. Increasing trends mean that the faster a bacterial population grows, the better it is at  
295 hydrolyzing the applied antibiotic. R164S showed no increasing trends in penicillin or penicillin  
296 + inhibitor treatments, supporting the idea that R164S is selected for cephalosporin resistance.

297 In comparison to clinical isolates, non-clinically identified genotypes are suspected to have less  
298 stable or effective enzymes. This likely has to do with the amino acid substitution locations and  
299 their impact on the three-dimensional structure of the enzyme and is reflected in the distribution  
300 of growth rate data and quantile regression trends. For example, genotypes LE, LRE, and LET  
301 have the double substitution LE (L21F and E240K) in common and have not been clinically  
302 identified. The first amino acid substitution, L21F, is a recently identified point mutation in  
303 TEM-117, a TEM  $\beta$ -lactamase first clinically isolated in 2003; no current research demonstrates  
304 L21F's role in TEM  $\beta$ -lactamase (Box, Paauw et al. 2002, Zeil, Widmann et al. 2016). Although  
305 no work has been done on L21F as a single mutation, it has been clinically identified with other  
306 amino acid substitutions within TEM  $\beta$ -lactamase, though not solely with E240K (Goussard and  
307 Courvalin 1999, Baraniak, Fiett et al. 2005, Morris, Whelan et al. 2006, Drissi, Ahmed et al.  
308 2008). E240K is positioned at the end of  $\beta$ -strand B3 such that the hydrophilic glutamic acid (E)  
309 residue is exposed, allowing it to interact with the acylamide substituents of cephalosporins  
310 (Knox 1995). However, B3 is too far in proximity to interact with smaller ligands like the  
311 inhibitors (clavulanic acid or sulbactam). Therefore, substitutions at this position typically do not  
312 appear in inhibitor-resistant TEM variants (Knox 1995). Overall, the distribution of growth rate  
313 data and trends in the quantile regression analysis reflect likely enzyme instability in LE  
314 genotypes. The genotypes that include LE have the highest numbers of decreasing trends in their



315 quantile regression plots (LE has eight, and LRE and LET both have seven) across all antibiotics.  
316 This means that the higher the antibiotic concentration, the slower these genotypes grew,  
317 suggesting an enzyme less effective at hydrolyzing  $\beta$ -lactam antibiotics. Likewise, genotypes  
318 with LE combination also had positive slopes at lower quantiles in the presence of penicillin +  
319 inhibitor (TZP) and a second-generation cephalosporin (CEC). This indicates the slower the  
320 populations grew, the TEM  $\beta$ -lactamase enzyme was more effective at hydrolyzing TZP and  
321 CEC. Altogether, these results suggests that the combination of L21F and E104K are selected for  
322 at lower concentrations of antibiotics, highlighting the influence of sub-lethal concentrations of  
323 antibiotic in the environment on the evolution of antibiotic resistance.

324

## 325 **Conclusion**

326 While infrequently used for biological datasets and not previously used for bacterial growth rate  
327 data, our study demonstrates that quantile regression analysis offers unique advantages to  
328 microbiologists. Our results show that quantile regression analysis provides a more detailed  
329 understanding of how increasing sub-lethal antibiotic concentrations, similar to those found in  
330 the environment, affect bacterial growth rates and provides insight into the genetic basis for  
331 varied responses. We illustrate how quantile regression analysis can link patterns in growth rates  
332 with certain types of  $\beta$ -lactam antibiotics and with clinically relevant mutations that can either  
333 hinder or enhance enzymatic activity on antibiotics. Ultimately, identifying the mutations that are  
334 most likely to appear in the clinic can help scientists and epidemiologists better predict the  
335 directionality of antibiotic resistance and develop novel pharmaceuticals to combat this  
336 worldwide crisis.

337 **References**

338

339 Abdi, H. and L. J. Williams (2010). "Tukey's honestly significant difference (HSD) test."  
340 Encyclopedia of research design **3**(1): 1-5.

341 Angrist, J., V. Chernozhukov and I. Fernandez-Val (2006). "Quantile regression under  
342 misspecification, with an application to the US wage structure." Econometrica **74**(2): 539-563.

343 Arnold, R. S., K. A. Thom, S. Sharma, M. Phillips, J. Kristie Johnson and D. J. Morgan (2011).  
344 "Emergence of *Klebsiella pneumoniae* carbapenemase-producing bacteria." South Med J **104**(1):  
345 40-45.

346 Baraniak, A., J. Fiett, A. Mrowka, J. Walory, W. Hryniewicz and M. Gniadkowski (2005).  
347 "Evolution of TEM-type extended-spectrum  $\beta$ -lactamases in clinical Enterobacteriaceae strains  
348 in Poland." Antimicrob Agents Chemother **49**(5): 1872-1880.

349 Bojar, B., J. Sheridan, R. Beattie, C. Cahak, E. Liedhegner, L. S. Munoz-Price, K. R. Hristova  
350 and T. Skwor (2021). "Antibiotic resistance patterns of *Escherichia coli* isolates from the clinic  
351 through the wastewater pathway." Int J Hyg Environ Health **238**: 113863.

352 Box, A. T. A., A. Paauw and M. A. Leverstein-vanHall (2002). *Escherichia coli* isolate 340  
353 TEM-117  $\beta$ -lactamase gene, partial cds. GenBank.

354 Byun, K. H., S. Ha Han, M. Woo Choi, S. Hong Park and S. D. Ha (2022). "Effect of sublethal  
355 concentrations of bactericidal antibiotics on mutation frequency and stress response of *Listeria*  
356 *monocytogenes*." Food Res Int **151**: 110903.

357 Cade, B. S. and B. R. Noon (2003). "A gentle introduction to quantile regression for ecologists."  
358 Frontiers in Ecology and the Environment **1**(8): 412-420.

359 Centers for Disease Control and Prevention, C. (2019). "Antibiotic Resistance Threats in the  
360 United States, 2019."

361 Cheah, A. L., T. Spelman, D. Liew, T. Peel, B. P. Howden, D. Spelman, M. L. Grayson, R. L.  
362 Nation and D. C. Kong (2013). "Enterococcal bacteraemia: factors influencing mortality, length  
363 of stay and costs of hospitalization." Clin Microbiol Infect **19**(4): E181-189.

364 Chen, I. C., S. J. Bertke and B. D. Curwin (2021). "Quantile regression for exposure data with  
365 repeated measures in the presence of non-detects." J Expo Sci Environ Epidemiol **31**(6): 1057-  
366 1066.

367 Dantas, G., M. O. Sommer, R. D. Oluwasegun and G. M. Church (2008). "Bacteria subsisting on  
368 antibiotics." Science **320**(5872): 100-103.

369 De Oliveira, A. B., S. Fischmeister, A. Diwan, M. Hauswirth and P. F. Sweeney (2013). "Why  
370 you should care about quantile regression." ACM SIGARCH Computer Architecture News  
371 **41**(1): 207-218.

372 Demain, A. L. and R. P. Elander (1999). "The  $\beta$ -lactam antibiotics: past, present, and future."  
373 Antonie Van Leeuwenhoek **75**(1-2): 5-19.

374 Do, Y. K., K. Foo, Y. Y. Ng and M. E. Ong (2013). "A quantile regression analysis of  
375 ambulance response time." Prehosp Emerg Care **17**(2): 170-176.

376 Drissi, M., Z. B. Ahmed, B. Dehecq, R. Bakour, P. Plesiat and D. Hocquet (2008). "Antibiotic  
377 susceptibility and mechanisms of  $\beta$ -lactam resistance among clinical strains of *Pseudomonas*  
378 *aeruginosa*: first report in Algeria." Med Mal Infect **38**(4): 187-191.

379 Escobar, W. A., A. K. Tan, E. R. Lewis and A. L. Fink (1994). "Site-directed mutagenesis of  
380 glutamate-166 in  $\beta$ -lactamase leads to a branched path mechanism." Biochemistry **33**(24): 7619-  
381 7626.

382 Frieri, M., K. Kumar and A. Boutin (2017). "Antibiotic resistance." J Infect Public Health **10**(4):  
383 369-378.

384 Goulart, C. P., M. Mahmudi, K. A. Crona, S. D. Jacobs, M. Kallmann, B. G. Hall, D. C. Greene  
385 and M. Barlow (2013). "Designing antibiotic cycling strategies by determining and  
386 understanding local adaptive landscapes." PLoS One **8**(2): e56040.

387 Goussard, S. and P. Courvalin (1999). "Updated sequence information for TEM  $\beta$ -lactamase  
388 genes." Antimicrob Agents Chemother **43**(2): 367-370.

389 Hall, B. G., H. Acar, A. Nandipati and M. Barlow (2014). "Growth rates made easy." Mol Biol  
390 Evol **31**(1): 232-238.

391 Herzberg, O. (1991). "Refined crystal structure of  $\beta$ -lactamase from *Staphylococcus aureus* PC1  
392 at 2.0 Å resolution." J Mol Biol **217**(4): 701-719.

393 Huletsky, A., J. R. Knox and R. C. Levesque (1993). "Role of Ser-238 and Lys-240 in the  
394 hydrolysis of third-generation cephalosporins by SHV-type  $\beta$ -lactamases probed by site-directed  
395 mutagenesis and three-dimensional modeling." J Biol Chem **268**(5): 3690-3697.

396 Iwu, C. D., L. Korsten and A. I. Okoh (2020). "The incidence of antibiotic resistance within and  
397 beyond the agricultural ecosystem: A concern for public health." Microbiologyopen **9**(9): e1035.

398 Jacoby, G. (2020, March 15, 2019). "B-Lactamase Classification and Amino Acid Sequences for  
399 TEM, SHV, and OXA Extended-Spectrum and Inhibitor Resistant Enzymes." Retrieved  
400 October 23, 2020, from [http://www.lahey.org/Studies/webt.asp#Ref\\_TEM\\_152](http://www.lahey.org/Studies/webt.asp#Ref_TEM_152).

401 Kaur, R., M. Pham, K. O. A. Yu and M. E. Pichichero (2021). "Rising Pneumococcal Antibiotic  
402 Resistance in the Post-13-Valent Pneumococcal Conjugate Vaccine Era in Pediatric Isolates  
403 From a Primary Care Setting." Clin Infect Dis **72**(5): 797-805.

404 Knox, J. R. (1995). "Extended-spectrum and inhibitor-resistant TEM-type  $\beta$ -lactamases:  
405 mutations, specificity, and three-dimensional structure." Antimicrob Agents Chemother **39**(12):  
406 2593-2601.

407 Knox, J. R. and P. C. Moews (1991). "B-lactamase of *Bacillus licheniformis* 749/C. Refinement  
408 at 2 Å resolution and analysis of hydration." J Mol Biol **220**(2): 435-455.

409 Koenker, R., Bassett Jr., G. (1978). "Regression Quantiles." Econometrica **46**(1): 33-50.

410 Labia, R., A. Morand, K. Tiwari, J. Sirot, D. Sirot and A. Petit (1988). "Interactions of new  
411 plasmid-mediated β-lactamases with third-generation cephalosporins." Rev Infect Dis **10**(4):  
412 885-891.

413 Leiva, A. M., B. Pina and G. Vidal (2021). "Antibiotic resistance dissemination in wastewater  
414 treatment plants: a challenge for the reuse of treated wastewater in agriculture." Reviews in  
415 Environmental Science and Bio/Technology **20**: 1043-1072.

416 Li, J., I. A. Phulpoto, G. Zhang and Z. Yu (2021). "Acceleration of emergence of *E. coli*  
417 antibiotic resistance in a simulated sublethal concentration of copper and tetracycline co-  
418 contaminated environment." AMB Express **11**(1): 14.

419 Li, Y., L. Xia, J. Chen, Y. Lian, A. A. Dandekar, F. Xu and M. Wang (2021). "Resistance  
420 elicited by sub-lethal concentrations of ampicillin is partially mediated by quorum sensing in  
421 *Pseudomonas aeruginosa*." Environ Int **156**: 106619.

422 Mira, P., B. Østman, C. Guzman-Cole, S. Sindi and M. Barlow (2021). "Adaptive Processes  
423 Change as Multiple Functions Evolve." Antimicrob Agents Chemother.

424 Mira, P. M., J. C. Meza, A. Nandipati and M. Barlow (2015). "Adaptive Landscapes of  
425 Resistance Genes Change as Antibiotic Concentrations Change." Mol Biol Evol.

426 Morris, D., M. Whelan, G. Corbett-Feeney, M. Cormican, P. Hawkey, X. Li and G. Doran  
427 (2006). "First report of extended-spectrum-β-lactamase-producing *Salmonella enterica* isolates in  
428 Ireland." Antimicrob Agents Chemother **50**(4): 1608-1609.

429 Mulvey, M. R. and D. A. Boyd (2009). "TEM-168, a heretofore laboratory-derived TEM β-  
430 lactamase variant found in an *Escherichia coli* clinical isolate." Antimicrob Agents Chemother  
431 **53**(11): 4955-4956.

432 Palzkill, T., Q. Q. Le, K. V. Venkatachalam, M. LaRocco and H. Ocera (1994). "Evolution of  
433 antibiotic resistance: several different amino acid substitutions in an active site loop alter the  
434 substrate profile of β-lactamase." Mol Microbiol **12**(2): 217-229.

435 Poole, K. (2004). "Resistance to β-lactam antibiotics." Cell Mol Life Sci **61**(17): 2200-2223.

436 Sanz, C., M. Casado, L. Navarro-Martin, D. Tadic, J. Parera, J. Tugues, J. M. Bayona and B.  
437 Pina (2021). "Antibiotic and antibiotic-resistant gene loads in swine slurries and their digestates:  
438 Implications for their use as fertilizers in agriculture." Environ Res **194**: 110513.

439 Shawky, M., W. B. Suleiman and A. A. Farrag (2021). "Antibacterial Resistance Pattern in  
440 Clinical and Non-clinical Bacteria by Phenotypic and Genotypic Assessment." Journal of Pure  
441 and Applied Microbiology **15**(4): 2270-2279.

442 Staffa, S. J., D. S. Kohane and D. Zurakowski (2019). "Quantile Regression and Its Applications:  
443 A Primer for Anesthesiologists." Anesth Analg **128**(4): 820-830.

444 Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro and P. Legendre (1996). "Untagling  
445 multiple factors in spatial distributions: Lilies, gophers, and rocks " Ecology **77**(6): 1698-1715.

446 Venkatachalam, K. V., W. Huang, M. LaRocco and T. Palzkill (1994). "Characterization of  
447 TEM-1  $\beta$ -lactamase mutants from positions 238 to 241 with increased catalytic efficiency for  
448 ceftazidime." J Biol Chem **269**(38): 23444-23450.

449 Villain, J., S. Lozano, M. P. Halm-Lemeille, G. Durrieu and R. Bureau (2014). "Quantile  
450 regression model for a diverse set of chemicals: application to acute toxicity for green algae."  
451 Journal of Molecular Modeling **20**(12).

452 Xu, Y., L. Tan, Q. Li, X. Zheng and W. Liu (2022). "Sublethal concentrations of heavy metals  
453 Cu and Zn can induce the emergence of bacterial multidrug resistance." Environmental  
454 Technology & Innovation **27**.

455 Zeil, C., M. Widmann, S. Fademrecht, C. Vogel and J. Pleiss (2016). "Network Analysis of  
456 Sequence-Function Relationships and Exploration of Sequence Space of TEM  $\beta$ -Lactamases."  
457 Antimicrob Agents Chemother **60**(5): 2709-2717.

458 Zhuang, M., Y. Achmon, Y. Cao, X. Liang, L. Chen, H. Wang, B. A. Siame and K. Y. Leung  
459 (2021). "Distribution of antibiotic resistance genes in the environment." Environ Pollut **285**:  
460 117402.  
461  
462

463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491

**Acknowledgements and Funding**

We are grateful for the University of California, Office of the President Presidential Postdoctoral Fellowship Award and Ruth L. Kirschstein National Research Service Award AI007323 (PM).

**Authors' contributions**

Portia Mira analyzed the data and lead writing.  
Candace C. Cole collected additional data that was needed for the analysis and revised the manuscript.  
Juan C. Meza suggested the use of quantile regression, implemented the R code, analyzed the data, and contributed to writing the manuscript.

**Availability of data**

All data and code used in this work can be found on GitHub upon acceptance.

**Conflicts of interest/Competing interests**

There are no conflicts of interest.

**Ethics approval**

Not applicable

**Consent to participate**

All authors consent to the participation of this work.

**Consent for publication**

All authors consent to the publication of the work.

492 **Figure Legends**

493 **Figure 1: Representative data showing how quantile regression plots are made.**

494 **Top:** Growth rate (cells/minute  $\times 10^{-3}$ ) plotted against normalized antibiotic concentration. X-  
495 axis ranges from 0 (no antibiotic) to 1 (highest concentration of antibiotic). Four quantiles are  
496 plotted (80<sup>th</sup> quantile – red, 60<sup>th</sup> quantile – orange, 40<sup>th</sup> quantile – blue, 20<sup>th</sup> quantile - black)  
497 across the growth rate data. **Bottom:** Quantile regression plots show the slope of each quantile  
498 line plotted across all quantiles. The colored dots represent the slope of the best fit model of each  
499 line in the above graph which corresponds to each of the quantiles. The red solid horizontal line  
500 represents data that would be normally distributed with the 95% confidence interval of normally  
501 distributed data shown by the red dotted lines. The gray shaded area represents the variation of  
502 the data calculated using the quantile regression package.

503  
504  
505 **Figure 2: Four common trends of quantile regression plots. A) Constant:** a trend remaining  
506 within the confidence intervals (red dotted lines) of the quantile regression plot. This type of  
507 trend means that the slope of the line of best fit at each quantile remains within the confidence  
508 intervals of a standard linear regression. Constant trends signify that the data is normally  
509 distributed and that a standard linear regression would work well. Variation that deviates beyond  
510 the red dotted lines signifies that the data is not normally distributed and that there are different  
511 effects at different quantiles. This example is given with the genotype R164S/E240K in the  
512 presence of cefepime (FEP) **B) Increasing:** a trend from bottom left to upper right. In this trend  
513 type, lower quantiles have more-negative slopes than higher quantiles. In our study, this trend  
514 means that higher antibiotic concentrations have a more detrimental effect on bacterial  
515 populations that grow slower (lower quantiles) compared to populations that grow faster (higher  
516 quantiles). This example is given with the genotype L21F/R164S in the presence of cefprozil  
517 (CPR) **C) Decreasing:** a trend from upper left to bottom right; the opposite of the increasing  
518 trend described above. Here, lower quantiles have less-negative slopes than higher quantiles,  
519 meaning that higher antibiotic concentrations have a more detrimental effect on bacterial  
520 populations that grow faster (higher quantiles) than on those that grow slower (lower quantiles).  
521 This example is given with the genotype L21F in the presence of cefotaxime (CTX) **D) U-**  
522 **shaped:** the extreme ends of this trend (80<sup>th</sup> and 20<sup>th</sup> quantiles) have less-negative slopes and the  
523 middle quantiles (25<sup>th</sup> to 75<sup>th</sup> quantiles) have more-negative slopes. This means that higher  
524 antibiotic concentrations are less detrimental to the bacterial populations that grow fastest and  
525 slowest compared to those in the middle. This example is given with the genotype  
526 R164S/E240K/T265M in the presence of ceftriaxone.

527  
528 **Figure 3: Frequency of quantile regression trend types for TEM-85 genotypes across all**  
529 **antibiotic treatments.** The number of occurrences of each trend type across the 16 TEM  
530 genotypes. Total number of U-shaped trends are yellow, constant trends are gray, decreasing

531 trends are red, and increasing trends are green. The genotypes that have been clinically identified  
532 are marked with an asterisk (\*).

533

534 **Figure 4: Quantile regression trends and positive slopes for TEM-85 genotypes across**  
535 **antibiotic treatments.** Antibiotics are listed by increasing chemical complexity and class (first  
536 column) and identified by abbreviation (second column). Genotypes that have been clinically  
537 identified are marked with an asterisk (\*). Genotypes are listed in subsequent columns in order of  
538 increasing number of substitutions, from wild-type (TEM-1) to TEM-85, which contains all four  
539 substitutions. The count of each trend type by antibiotic is listed in the table to the right.



540 **Tables**541 **Table 1: Variant TEM genotypes.**

542 Genotypes are listed by increasing number of substitutions, from TEM-1 (no substitutions,  
 543 wildtype) to TEM-85 (four amino acid substitutions). The number of substitutions is listed in the  
 544 first column; full amino acid substitutions are listed in the second column. If the genotypes have  
 545 been clinically isolated, their TEM name (third column) and year of first isolation (last column)  
 546 are denoted

547

No. of Substitutions	Substitution	Isolated	Year
0	TEM-1	TEM-1	1965
1	L21F	TEM-117	2003
1	R164S	TEM-12	1999
1	E240K	TEM-191	2011
1	T265M	TEM-168	2009
2	L21F/R164S	TEM-53	1999
2	L21F/E240K	-	-
2	L21F/T265M	TEM-110	2002
2	R164S/E240K	TEM-10	1989
2	R164S/T265M	-	-
2	E240K/T265M	-	-
3	L21F/R164S/E240K	-	-
3	L21F/R165S/T265M	TEM-102	2003
3	L21F/E240K/T265M	-	-
3	R164S/E240K/T265M	-	-
4	L21F/R164S/E240K/T265M	TEM-85	2005

548

549

550

551  
552  
553  
554  
555  
556

**Table 2: Antibiotics and concentrations used to measure bacterial growth rates**

Antibiotics are listed in order of increasing complexity ( $\beta$ -lactam generation) in the first column, followed by full antibiotic name and abbreviation. The three concentrations used in (Mira, Østman et al. 2021) are listed in the last column.

Antibiotic Name	Abbreviation	Concentrations ( $\mu\text{g/mL}$ )	$\beta$ -lactam group
Amoxicillin	AMX	256, 512	Penicillin
Ampicillin	AMP	1024, 2048, 3072	Penicillin
Ampicillin + Sulbactam	SAM	8, 16, 32	Penicillin + $\beta$ -lactamase Inhibitor
Piperacillin + Tazobactam	TZP	32, 64, 128	Penicillin + $\beta$ -lactamase Inhibitor
Cefaclor	CEC	2, 4, 8	2 <sup>nd</sup> generation cephalosporin
Cefotetan	CTT	0.063, 0.125, 0.25	2 <sup>nd</sup> generation cephalosporin
Cefuroxime	CXM	2.25, 3, 4	2 <sup>nd</sup> generation cephalosporin
Ceftazidime	CAZ	0.125, 0.25, 0.5	3 <sup>rd</sup> generation cephalosporin
Cefprozil	CPR	8, 12, 16	3 <sup>rd</sup> generation cephalosporin
Ceftriaxone	CRO	0.025, 0.05, 0.1	3 <sup>rd</sup> generation cephalosporin
Cefotaxime	CTX	0.03, 0.06, 0.123	3 <sup>rd</sup> generation cephalosporin
Ceftizoxime	ZOX	0.0078, 0.0156, 0.03	3 <sup>rd</sup> generation cephalosporin
Cefepime	FEP	0.0312, 0.0625, 0.125	4 <sup>th</sup> generation cephalosporin

557  
558