Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Construction and Application of a Qualitative Risk Model for β-Lactam Antimicrobial Resistance in Escherichia coli from Livestock and Poultry

Mao Yicheng

199794325640163.com

Chen Yu
Liu Fangqin
Zhang Chunping
Cui Mingquan
Li Ting
Wang Hejia

Research Article

Keywords: Escherichia coli, antimicrobial resistance, β-lactam antimicrobials, risk assessment model

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11504909/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Construction and Application of a Qualitative Risk Model for β-Lactam Antimicrobial Resistance in Escherichia coli from Livestock and Poultry

Mao Yicheng[1,2※] Chen Yu[1,2], Liu Fangqin[1,2], Zhang Chunping[1,2], Cui Mingquan, Li Ting[1,2※※], Wang Hejia[1,2※※]

(1. China Institute of Veterinary Drug Control, Beijing 100081; 2. Key Laboratory of Animal-origin Bacterial Antimicrobial Resistance Surveillance, Ministry of Agriculture and Rural Affairs/FAO Reference Centre for Antimicrobial Resistance, Beijing 100081)

Abstract

The resistance spectrum of Escherichia coli from livestock and poultry to β-lactam antimicrobials has continued to expand, and has spread from penicillins to third-generation cephalosporins, posing a severe challenge to public health security. This study systematically collected 1058 Escherichia coli strains from pigs and chickens covering six provinces and municipalities in China from 2011 to 2021, conducted antimicrobial susceptibility testing and whole-genome sequencing, and constructed a risk assessment model combining strain phenotype and genotype. The study found that the resistance rates of the strains to ampicillin, cefotaxime, and cefoxitin were 81.0%, 33.4%, and 11.2%, respectively; the major prevalent genes were *bla*$_{\text{TEM}}$ (51.0%) and *bla*$_{\text{CTX-M}}$ (27.0%). The risk assessment model established based on a decision tree divided the strains into four levels; Level I high-risk strains accounted for 30.6%, and the distribution patterns of strain resistance characteristics across time, region, and host were highly consistent with risk levels. This model integrates multidimensional data such as resistance phenotype, genotype, and strain sampling information, grades strain risk, and can provide a theoretical basis and technical support for zoned precise prevention and control of β-lactam antimicrobial resistance in Escherichia coli from livestock and poultry and for rational veterinary clinical drug use.

Keywords

Escherichia coli, antimicrobial resistance, β-lactam antimicrobials, risk assessment model

Introduction

β-Lactam antimicrobials include multiple subclasses such as penicillins and cephalosporins, and are antimicrobial drugs used extremely widely in human clinical anti-infection and livestock

and poultry breeding production, and their usage on the domestic farming end has long ranked among the top three veterinary antimicrobials[1][2]. With long-term extensive application, the problem of bacterial antimicrobial resistance has become increasingly serious, among which the transmission risk of third-generation cephalosporin-resistant bacteria is particularly severe, and the World Health Organization (WHO) in 2024 included third-generation cephalosporin-resistant bacteria in the critical category of the WHO Bacterial Priority Pathogens List 2024[3]. Escherichia coli is one of the most common intestinal opportunistic pathogens and is also an important reservoir of antimicrobial resistance genes[4], and the β-lactamase resistance genes it carries (such as bla_{TEM}, $bla_{\text{CTX-M}}$, bla_{CMY}, bla_{NDM}, *etc.*) can spread rapidly among different strains and even different species through mobile genetic elements such as plasmids and transposons, exacerbating the spread and prevalence of antimicrobial resistance[5][6]. Existing surveillance data show that the resistance of Escherichia coli from livestock and poultry to β-lactam antimicrobials has gradually expanded from basic penicillin drugs to third-generation cephalosporin drugs [7][8][9][10]. Long-term antimicrobial use pressure has led to the increasingly prominent antimicrobial resistance problem of Escherichia coli from livestock and poultry, becoming an important challenge facing the global public health field.

At present, multiple studies have carried out risk assessment targeting antimicrobial resistance genes and established multi-indicator evaluation processes and corresponding risk assessment systems[11][12][13]. However, current studies are mostly limited to local regions or short-term periods, and do not combine strain phenotype and genotype, making it difficult to comprehensively construct an overall assessment model reflecting the resistance risk of Escherichia coli from livestock and poultry to β-lactam antimicrobials. Therefore, this study, based on 1058 Escherichia coli strains from six provinces and municipalities—Henan Province, Sichuan Province, Liaoning Province, Guangdong Province, Shandong Province, and Shanghai Municipality—spanning 10 years (2011-2021), established a decision tree risk assessment model with resistance phenotype as the core determining condition[14][15]and combined with genotype for comprehensive assessment. It aims to clarify the key control points of β-lactam antimicrobial resistance, transform surveillance data into risk management tools with predictive functions, and provide theoretical basis and technical support for guiding clinical drug use and preventing and controlling the spread of antimicrobial resistance from livestock and poultry.

Results

Resistance phenotype

In this study, phenotypic antimicrobial susceptibility testing was performed on 1058 strains, and the results are shown in Figure 1a. Escherichia coli had the highest resistance rate to AMP, reaching 81.0%; followed by CTX, with a resistance rate of 33.4%; and the lowest resistance rate to FOX, at 11.2%.

Dynamic analysis of antimicrobial resistance from 2011 to 2021 (Figure 1b) found that the AMP resistance problem was the most severe, with the resistance rate remaining consistently in the range of 70.9%–90.0%; the annual fluctuation in CTX resistance rate was pronounced, ranging from 18.9% to 46.2%; the FOX resistance level was relatively low, at 6.8% to 28.7%, and after 2013 it showed a downward trend and stabilized at around 10%.

There were significant statistical differences in strain resistance rates among different regions (χ^2=785.6, P<0.001). As shown in Figure 1c, the AMP resistance rates in all provinces and municipalities were at a relatively high level, averaging 81.1%; CTX was next, averaging 34.7%; and FOX had the lowest resistance rate, averaging 11.4%. Between-group comparisons showed that the CTX resistance rate was highest in Shandong Province (52.7%), and compared with Liaoning Province (37.2%), the difference was extremely significant (P<0.01), and compared with the remaining provinces, the differences all reached extremely significant levels (P<0.001); the FOX resistance rate was highest in Liaoning Province (15.2%), and compared with Guangdong Province, which had the lowest resistance rate (5.0%), the difference was extremely significant (P<0.001). In addition, the AMP resistance rate was at a relatively high level in all provinces and municipalities, among which Henan Province was the highest (86.4%), significantly higher than that of Guangdong Province (75.7%) (P<0.01).

In addition, based on antimicrobial resistance surveillance data from 510 pig-derived and 548 chicken-derived Escherichia coli strains, this study used the Z test for statistical analysis, and the results are shown in Figure 1d. The resistance rate of chicken-derived Escherichia coli to CTX was 46.9%, significantly higher than 18.8% in pig-derived strains (P<0.001); meanwhile, the resistance rates of chicken-derived strains to FOX and AMP were 13.7% and 84.3%, respectively, both higher than those of pig-derived strains (FOX 8.4%, AMP 77.5%).

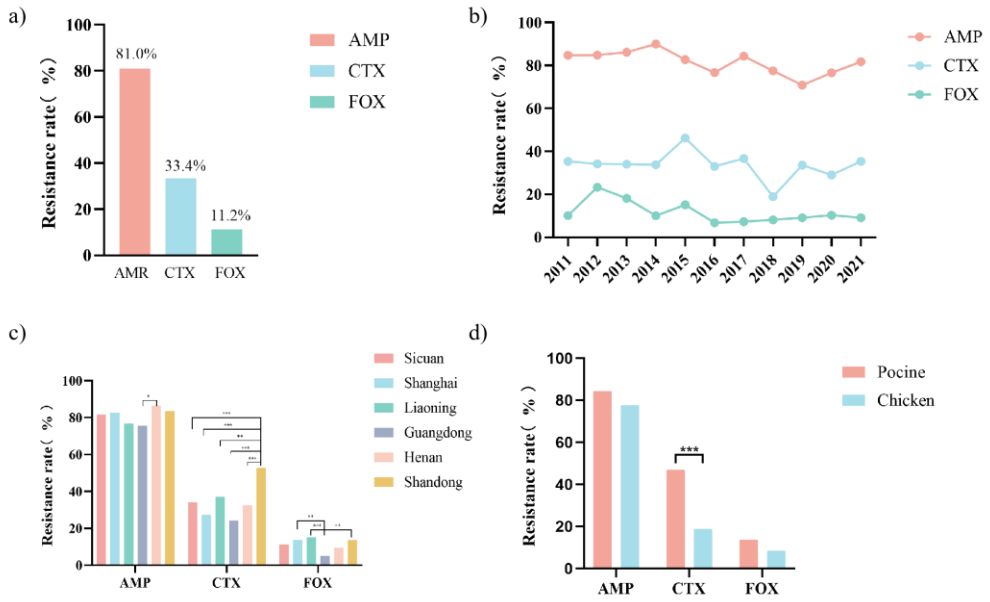

Figure 1 a) Phenotypic antimicrobial susceptibility testing of three antimicrobials b) Annual changes in resistance rates of the three antimicrobials c) Resistance status of the three antimicrobials in different regions d) Resistance rates in different hosts

Resistance genotype

Significant distribution differences were observed among different antimicrobial resistance genes (X2 = 465.2, P < 0.001) (Figure 3a). The *bla*$_{TEM}$ gene dominated with a detection rate of 51.0%, followed by the extended-spectrum β-lactamase gene *bla*$_{CTX-M}$ (27.0%) and *bla*$_{CMY}$ (5.0%). In addition, the detection rates of carbapenemase genes were low, with a detection rate of *bla*$_{NDM}$ of 0.6%, while *bla*$_{KPC}$ and *bla*$_{VIM}$ were not detected. Other rare genes, including *bla*$_{DHA}$, *bla*$_{CARB}$, and *bla*$_{SHV}$, all had detection rates below 0.5%.

Association analysis of phenotypic and genotypic results found that the high resistance rate of CTX (33.4%) may be highly correlated with the high prevalence of the *bla*$_{CTX-M}$ gene (27.0%). Meanwhile, among AMP-resistant strains, the detection rate of the *bla*$_{TEM}$ gene was as high as 51.0%. Kappa agreement test analysis found that there was high agreement between the CTX-resistant phenotype and the *bla*$_{CTX-M}$ gene (Kappa = 0.78, P < 0.001), while the agreement between the AMP-resistant phenotype and the blaTEM genotype was weaker (Kappa = 0.312). In addition, the overall detection rates of AmpC enzyme genes (*bla*$_{CMY}$, *bla*$_{DHA}$) and carbapenemase genes (*bla*$_{NDM}$) were low, but in strains resistant to both CTX and FOX, the detection rates of such genes were significantly elevated (P < 0.001).

Risk assessment model results

Based on the decision tree model constructed in this study, Escherichia coli was divided into

four risk grades, and the results are shown in Figure 2(a). Grade I were high-risk strains, which possessed a resistant phenotype and carried high-risk genes, totaling 325 strains, accounting for 30.6%. Grade II were moderate-risk strains, resistant to key drugs but with no target high-risk genes detected, and such strains totaled 10 strains, accounting for 0.9%. Grade III were low-risk strains, resistant only to basic β-lactam antimicrobials such as AMP, susceptible to third-generation cephalosporins, and not carrying high-risk genes, totaling 529 strains, accounting for 50.1%. Grade IV were susceptible strains, susceptible to all tested drugs and with no high-risk gene risk, totaling 194 strains, accounting for 18.3%.

In terms of temporal distribution, from 2011 to 2021, the distribution of antimicrobial resistance risk grades of Escherichia coli showed obvious annual dynamic changes, as shown in Figure 2(b). The proportion of Grade I high-risk strains was 30.4% in 2011, rose to a peak of 41.9% in 2015, and then fluctuated downward, reaching 31.3% by 2021. Grade III risk strains always occupied a dominant position in each year, exceeding 38% except in 2015, and reached the highest value of 60.4% in 2018. The proportion of Grade IV low-risk strains was relatively stable, mostly fluctuating between 15%–22%, but rose to 29.1% in 2019. The detection rate of Grade II risk strains was overall relatively low, not exceeding 3.2% in any year, and they were not detected in multiple years.

In terms of regional distribution, there were significant differences in the antimicrobial resistance risk distribution of Escherichia coli among different regions, as shown in Figure 2(c). High-risk strains (Grade I + Grade II) were highest in the Shandong region, reaching 51.8%, among which Grade I risk strains accounted for 50%, significantly higher than other regions. The Liaoning region followed, with a proportion of high-risk strains of 36.6%, and the Sichuan and Henan regions were 31.6% and 31.4%, respectively, while the Shanghai and Guangdong regions were relatively low, at 23.7% and 23.2%, respectively. All regions were dominated by Grade III risk strains; by comparison, the proportion of Grade III risk strains in the Shandong region was the lowest (32.7%). The distribution of Grade IV susceptible strains in each region was relatively balanced, with proportions between 12.4% and 24.3%.

In terms of host distribution, the antimicrobial resistance risk of strains from different hosts also differed (Figure 2(d)). The proportion of high-risk strains (Grade I + Grade II) among chicken-derived strains was as high as 44.9%, higher than 17.3% among pig-derived strains. In

terms of risk grade composition, pig-derived strains were dominated by Grade III risk, accounting for 61.0%, higher than 40.0% for chicken-derived strains. The proportion of Grade IV susceptible strains among pig-derived strains (21.8%) was also higher than that among chicken-derived strains(15.1%).

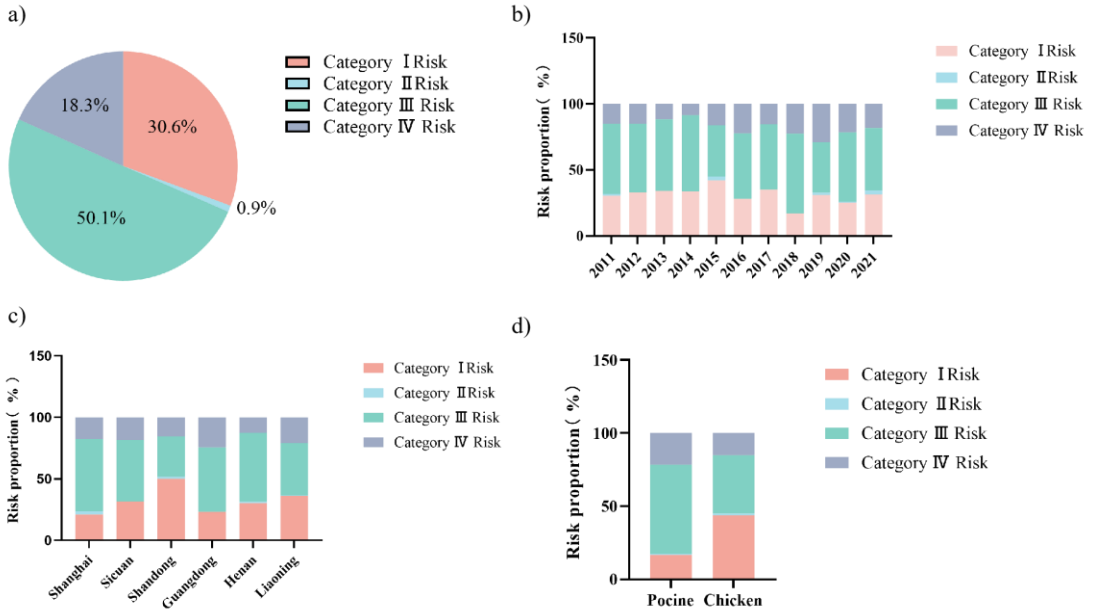

Figure 2 (a) Risk grade distribution chart (b) Annual distribution chart of risk grades (c) Regional distribution of risk grades (d) Risk grade distribution chart of different hosts

Model feasibility verification

Table 1 Overall accuracy

RISK GRADE	ACCURACY (%)	95% CONFIDENCE INTERVAL (%)	PRECISION (%)	RECALL (%)	F1-SCORE(%)
GRADE I RISK	95.94	94.75-97.13	90.77	95.78	93
GRADE II RISK	93.38	91.88-94.88	90	11.54	20
GRADE III RISK	89.32	87.46-91.18	82.04	96.02	88
GRADE IV RISK	94.71	93.36-96.06	92.27	81.36	86

The risk classification model constructed in this study performed well overall, with an overall accuracy of 93.34%, but the performance of each category was unbalanced. The model's

identification of Grade I risk and Grade IV risk was relatively robust, with F1 scores reaching 93% and 86%, respectively. However, the model had two key shortcomings. First, there was severe underreporting for Grade II risk, with a recall of only 11.54%, causing the F1 score to drop sharply to 20%. Second, for Grade III risk there were relatively many false positives, with a precision of only 82.04%.

Discussion

Phenotype and genotype

The AMP resistance results of the strains obtained in this study are consistent with global reports on AMP resistance in *Escherichia coli*[7]**Error! Reference source not found.**. As a classic penicillin, AMP has long been and widely used in both human medicine and the veterinary field, and continuous selection pressure has directly led to the widespread dissemination of its resistance**Error! Reference source not found.**.

Ten years of surveillance data show that the AMP resistance rate has remained at a high level for a long time. This suggests that under continuous environmental selection pressure, the relevant resistance genes may have stably colonized and continuously spread within the bacterial population, forming the current situation of high resistance rates in strains. Compared with the high resistance status of AMP, the overall resistance rate of strains to CTX is not high, but its annual fluctuation is obvious. This phenomenon reflects differences in farming drug use and disease prevention and control management at different stages, suggesting that the use and control of third-generation cephalosporins is not yet standardized. CTX is a key drug for the prevention and treatment of Gram-negative bacterial infections, and the instability of its resistance level will increase the difficulty of clinical treatment and easily cause staged spread of resistant strains. The resistance rate of FOX is overall at a relatively low level; it is stable against most extended-spectrum β-lactamases (ESBLs), and its relatively low resistance rate makes it an important drug for treating infections caused by ESBL-producing Enterobacteriaceae**Error! Reference source not found.**.

In addition, there are significant regional differences in the resistance level of *Escherichia coli* from livestock and poultry to β-lactam antimicrobials in different regions, and this phenomenon may be due to differences in farming models, drug use habits, and the intensity of veterinary antimicrobial control in each region. At the same time, the overall resistance level of

chicken-derived strains is significantly higher than that of pig-derived strains, and this finding is consistent with multiple global studies[Error! Reference source not found.Error! Reference source not found.]. It may be closely related to high-density poultry farming and high-frequency preventive use of antimicrobials.

The resistance prevalence of *Escherichia coli* from livestock and poultry is dominated by broad-spectrum β-lactamase and ESBLs genes, and high-risk carbapenem resistance genes are only distributed in a few cases and have not spread widely. bla_{TEM} is one of the earliest discovered and widely disseminated β-lactamase genes[Error! Reference source not found.], and its widespread presence may be related to the long-term and extensive use of AMP in the medical and agricultural fields. The ESBL encoded by the $bla_{\mathrm{CTX\text{-}M}}$ gene in *Escherichia coli* can produce resistance to drugs such as third-generation cephalosporins[Error! Reference source not found.], and its prevalence means that treatment options will face greater challenges. In this study, the detection rates of carbapenemase genes (bla_{KPC}, bla_{NDM}, bla_{VIM}) were extremely low or not found, suggesting that carbapenem drugs remain one of the drugs with high effectiveness against ESBL-producing strains in the current context**Error! Reference source not found.**.

Association analysis of phenotype and genotype showed that CTX resistance had high agreement with $bla_{\mathrm{CTX\text{-}M}}$ (Kappa = 0.78, $P < 0.001$), indicating that this gene is one of the main molecular mechanisms mediating CTX resistance in Escherichia coli from livestock and poultry. The agreement between the AMP-resistant phenotype and the bla_{TEM} genotype was relatively weak (Kappa = 0.312), suggesting that in addition to bla_{TEM}, there are other more major resistance mechanisms mediating the resistant phenotype. In addition, the detection rates of low-prevalence AmpC enzyme genes and bla_{NDM} were elevated in multi-cephalosporin-resistant strains, indicating that although the overall prevalence of such rare resistance genes is low, they will be enriched in multidrug-resistant strains, exacerbating resistance risk.

Risk assessment model

The core value of this model lies in systematically integrating complex phenotypic and genotypic information and transforming it into easily interpretable visualized risk indicators, thereby establishing a bridge between data and clinical decision-making. This indicator is helpful for guiding the formulation of scientific drug use regimens and also has important theoretical guiding significance for delaying the spread of antimicrobial resistance in *Escherichia coli* from

livestock and poultry. The model results show that Grade I high-risk strains accounted for 30.6%; they are important strains that accelerate the spread of antimicrobial resistance and are also the key surveillance and priority intervention targets of this study. Most of their resistance genes are located on mobile genetic elements and have the potential for horizontal transfer between strains, directly endangering the safety of clinical drug use. Therefore, rapid identification and strict control of Grade I high-risk strains is the key to curbing the continued spread of antimicrobial resistance. Grade II moderate-risk strains accounted for only 0.9%, but they have important research value. Such strains show a third-generation cephalosporin-resistant phenotype but no high-risk genes were detected. This suggests that there may be undiscovered resistance genes and unknown resistance mechanisms beyond the current detection scope, such as other rare ESBL genes, AmpC enzyme genes, drug efflux pumps, or porin mutations. Therefore, continuous tracking and in-depth analysis of such strains will help discover new resistance pathways. Grade III low-risk strains constitute the main body of the bacterial population (50.1%), are resistant only to basic β-lactam antimicrobials such as AMP, are susceptible to third-generation cephalosporins, and do not carry high-risk genes. This reflects the background resistance formed under long-term selective pressure from penicillin drugs. Although the direct clinical hazard is relatively low, they can serve as potential donors for resistance gene transfer and still need to be included in routine surveillance. Grade IV fully susceptible strains accounted for 18.3%, indicating that even in farming environments with relatively high antimicrobial selective pressure, a certain proportion of susceptible strains still exists.

In addition, the spatiotemporal and host distribution characteristics of the resistance phenotypes and genotypes of the strains in this study are highly consistent with the distribution patterns of antimicrobial resistance risk grades. In the temporal dimension, the antimicrobial susceptibility results of strains from 2011 to 2021 show that the resistance rates of AMP and CTX fluctuated dynamically with the year, completely matching the annual distribution trend of risk grades in this study. The peak of the overall resistant phenotype of strains in 2015 corresponded to the highest proportion of Grade I and Grade II high-risk strains in ten years, suggesting that the increase in phenotypic resistance caused by clinical drug use pressure led to the staged enrichment of high-risk resistant strains; the subsequent decrease in resistance rates and the proportion of high-risk strains also confirmed that changes in resistant phenotypes are the core inducement of

the temporal evolution of risk grades. During the ten-year surveillance period, resistance risk grades were always in a state of dynamic fluctuation; even though the proportion of high-risk strains decreased after the peak, it still rebounded to 34.3% in 2021, reflecting the long-term and recurrent nature of antimicrobial resistance prevention and control; and Grade III low-risk strains occupying the main body of the bacterial population also reflects that basic resistant strains have stably colonized within the farming bacterial population, continuously bringing pressure to veterinary clinical anti-infection treatment.

At the regional distribution level, the CTX resistance rates in Shandong, Liaoning, and other places were significantly higher than those in other regions, completely consistent with the model results of a relatively high proportion of regional high-risk strains. At the same time, the overall phenotypic resistance rates in Guangdong, Shanghai, and other regions were low, corresponding to their low proportion of high-risk strains, indicating the background characteristics of the regional phenotypic resistance prevalence level and resistance gene distribution. The differences in risk distribution among regions are speculated to be directly related to local farming models, antimicrobial use intensity, and veterinary drug control intensity; all provinces and municipalities were dominated by Grade III low-risk strains, suggesting that basic-level β-lactam resistance is a nationwide common phenomenon, and also indicating that domestic antimicrobial resistance prevention and control needs to implement differentiated strategies, carrying out key surveillance and precise intervention in high-risk regions.

At the host difference level, chicken-derived strains showed a higher third-generation cephalosporin resistance rate than pig-derived strains, and the carriage of high-risk genes such as $bla_{\text{CTX-M}}$ and bla_{TEM} was more common, explaining from both phenotypic and genotypic levels the core reason why the proportion of high-risk grades among chicken-derived strains was higher than that among pig-derived strains. The drug selection pressure of intensive poultry farming promotes frequent occurrence of resistant phenotypes and widespread dissemination of resistance genes, ultimately resulting in an overall higher antimicrobial resistance risk of chicken-derived strains, while the distribution characteristic of pig-derived strains being dominated by low-risk grades is also consistent with their relatively lower resistance rates and high-risk gene carriage rates. Therefore, strengthening the standardized use of antimicrobials in the farming sector is an important measure to control the spread of antimicrobial resistance from the source.

Conclusion

The risk assessment model established in this study transforms surveillance data into a refined management tool with predictive functions, and reveals the key risk nodes of β-lactam antimicrobial resistance in Escherichia coli from livestock and poultry in China. This system provides a scientific basis and methodological support for the implementation of targeted and differentiated antimicrobial resistance prevention and control strategies, and has important practical significance for safeguarding public health security. Subsequent research can, on this basis, further incorporate clinical data and environmental samples to expand the application scenarios and dynamic early warning capabilities of the model.

Methods

Data collection and preprocessing

This study selected 1058 Escherichia coli strains collected from six provinces (municipalities)—Guangdong, Sichuan, Henan, Shandong, Liaoning, and Shanghai—from 2011 to 2021, and the distribution is shown in Table 2.

Table 2 Distribution of strains and different hosts in six provinces (municipalities) from 2011 to 2021

Region	Pig-derived	Chicken-derived	Total
Guangdong	89	92	181
Sichuan	113	131	244
Henan	84	85	169
Shandong	51	59	110
Liaoning	77	87	164
Shanghai	96	94	190
Total	510	548	1058

Strain resuscitation and identification

Cryopreserved strains were thawed in a water bath and inoculated into BHI broth, then sequentially subjected to stepwise streak culture on *E. coli* chromogenic medium, MacConkey selective medium, and nutrient agar plates to screen and purify target colonies. Subsequently, MALDI-TOF MS was used for strain identification.

Antimicrobial susceptibility testing

This study selected three antimicrobials—ampicillin (AMP), cefotaxime (CTX), and cefoxitin (FOX)—and, according to the standards of the Clinical and Laboratory Standards Institute (CLSI)[Error! Reference source not found.], used the broth microdilution method to determine the minimum inhibitory concentration and judge the resistance status.

Whole-genome sequencing

Strain DNA was extracted, and paired-end sequencing was performed using the MGISEQ-2000RS high-throughput sequencer. The raw sequencing data were preliminarily quality assessed using FastQC (v0.11.9) and MultiQC (v1.23), then quality-controlled using Fastp (v0.23.4) and quality assessed again. Genome assembly was completed using SPAdes (v3.15.5)[Error! Reference source not found.], and genome annotation was completed using Bakta (v1.10.4). For antimicrobial resistance gene analysis, the Resfinder database[Error! Reference source not found.] in the bioinformatics software ABRicate was used, with minimum coverage and minimum identity both set to 90%, and the carriage of antimicrobial resistance genes was counted.

Model establishment

Discrete model establishment

Phenotypic and genotypic data and strain sampling information were stored in Excel format, using a unified data format. All categorical variables were converted into binary numerical format (0 indicates susceptible/negative, 1 indicates resistant/positive), and the conversion rules are as follows: character-type data such as "YES", "Y", "1", "resistant", "TRUE", "T" were uniformly converted to 1, and the rest were converted to 0; logical-type data were directly converted to numerical type; numerical-type data retained their original values. The resistance risk of Escherichia coli to the two classes of drugs was divided into four grades (Grades I–IV). (1) Resistance to at least one of the β-lactam drugs (AMP, FOX); (2) Resistance to third-generation cephalosporin drugs (CTX); (3) Carrying at least one of the following high-risk β-lactamase genes [Error! Reference source not found.][Error! Reference source not found.]: *bla*$_{CMY}$, *bla*$_{CTX-M}$, *bla*$_{TEM}$, *bla*$_{DHA}$, *bla*$_{CARB}$, *bla*$_{SHV}$, *bla*$_{Z}$. In addition, carbapenems possess the broadest antimicrobial spectrum and the strongest antimicrobial activity among β-lactam drugs, and their stability to ESBLs and AmpC enzymes gives them an irreplaceable position in the treatment of fatal resistant infections[Error! Reference source not found.], therefore, in this model, as long as a gene encoding a carbapenemase (*bla*$_{NDM}$, *bla*$_{KPC}$, *bla*$_{VIM}$) [Error! Reference source not found.] is present, it belongs to Grade I risk. Finally, according to

the logic shown in the decision tree flowchart in Figure 3, each strain was assigned to the corresponding risk grade, and the classification details are shown in Table 3.

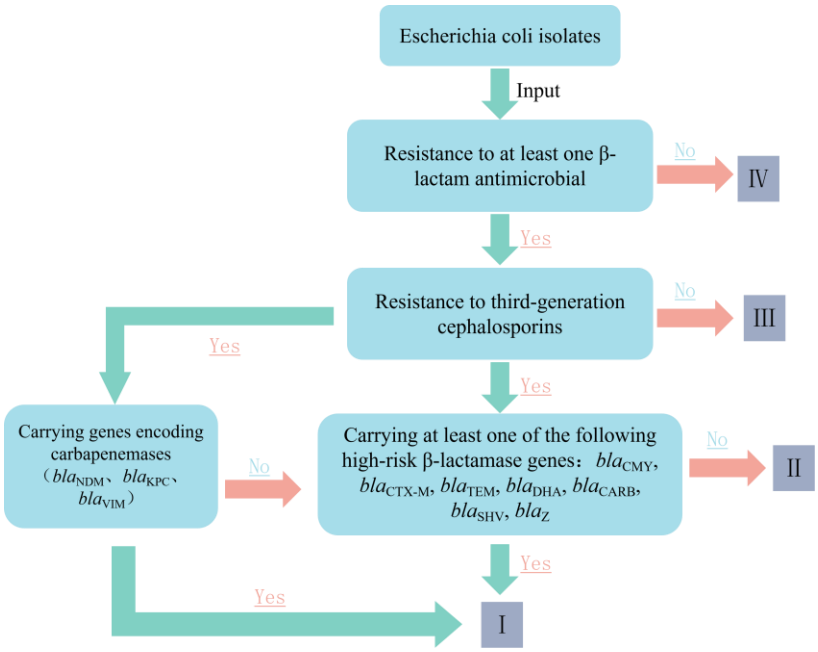

Figure 3 Decision tree for resistance risk grading of Escherichia coli

Table 3 Decision tree for resistance risk grading of Escherichia coli

RISK GRADE	CONDITIONS
GRADE I RISK (HIGH RISK)	1.Resistance to at least one β-lactam antimicrobial; 2.Resistance to third-generation cephalosporins; 3.Carrying at least one high-risk β-lactamase gene (*bla*CMY, *bla*CTX-M, *bla*TEM, *bla*DHA, *bla*CARB, *bla*SHV, *bla*Z). 4.As long as it contains genes encoding carbapenemases (*bla*NDM, *bla*KPC, *bla*VIM).
GRADE II RISK (MODERATE RISK)	1. Resistance to at least one β-lactam antimicrobial; 2. Resistance to third-generation cephalosporins; 3. Not carrying the above-mentioned high-risk β-lactamase genes.
GRADE III RISK (LOW RISK)	1. Resistance to at least one β-lactam antimicrobial; 2. Susceptible to third-generation cephalosporins; 3. Not carrying the above-mentioned high-risk β-lactamase genes.
GRADE IV RISK	1. Susceptible to all β-lactam antimicrobials; 2. Susceptible to third-generation cephalosporins;

(SUSCEPTIBLE STRAINS)	3. Not carrying the above-mentioned high-risk β-lactamase genes.

Model feasibility verification

Overall accuracy

This study used overall accuracy to verify the accuracy of the model's overall prediction, which is the most intuitive indicator for evaluating the performance of a classification model. The calculation method is the proportion of the number of samples correctly predicted by the model to the total number of samples, that is, the sum of the diagonal elements (the sum of true positives and true negatives) divided by the sum of all prediction results. In this study, the calculation of accuracy covered the correctly classified samples in all four categories (Grade I risk, Grade II risk, Grade III risk, Grade IV risk), and the specific calculation method is**Error! Reference source not found.**:

$$\text{Overall accuracy}=\frac{\text{number of correctly classified samples}}{\text{total number of samples}}\times 100\%=\frac{TP+TN}{TP+TN+FP+FN}\times 100\%$$

TP（True Positive）：True positives, the number of samples that are actually positive and are correctly predicted as positive.

TN（True Negative）：True negatives,the number of samples that are actually negative and are correctly predicted as negative.

FP（False Positive）：False positives, the number of samples that are actually negative but are incorrectly predicted as positive.

FN（False Negative）：False negatives, the number of samples that are actually positive b utare incorrectly predicted as negative.

Confidence interval estimation used the Clopper-Pearson exact method to calculate the 95% confidence interval of the overall accuracy **Error! Reference source not found.**, and the formul a is:

$$\left[\text{CI}=\text{Beta}\left(\frac{\alpha}{2};\kappa,n-\kappa+1\right),\ \text{Beta}\left(1-\frac{\alpha}{2};\kappa+1,n-\kappa\right)\right]$$

where κ is the number of correctly classified samples, n is the total number of sampl es,α=0.05.

Multi-dimensional performance evaluation system

To comprehensively evaluate the classification performance of the antimicrobial resistance gene risk assessment model, this study adopted a multi-indicator comprehensive evaluation framework.

Precision is used to measure the proportion of samples that are truly positive among the samples predicted as positive by the model, reflecting the prediction purity of the model, and its calculation formula is as follows **Error! Reference source not found.**

$$Precision=\frac{TP}{TP+FP}$$

Recall is used to evaluate the model's ability to correctly identify positive-class samples and reflects the detection completeness of the model, and the calculation formula is **Error! Reference source not found.**:

$$Recall=\frac{TP}{TP+FN}$$

F1-score is the harmonic mean of precision and recall, providing a single comprehensive indicator to balance precision and recall, and its calculation method is **Error! Reference source not found.**:

$$\mathrm{F1-score}=2\times\frac{Precision\times Recall}{Precision+Recall}$$

This multi-indicator evaluation method can comprehensively reflect the classification performance of the model from different dimensions**Error! Reference source not found.**, and is suitable for scenarios in risk assessment where different types of errors have different importance.

Declarations

Ethics approval and consent to participate（Not applicable）

Consent for publication（Not applicable）

Availability of data and materials

The datasets generated, used and analyzed during the current study are not deposited in a public repository at this stage. All relevant raw data, processed analytical files, supporting laboratory records and related supplementary materials that underpin the findings reported in this manuscript are available from the corresponding author upon reasonable request. Any requests for data access should clearly state the purpose of data use, research background and planned analysis scheme. The corresponding author will evaluate each application, and data sharing will be conducted in accordance with relevant institutional regulations, biosafety requirements and data confidentiality rules. Access to the data may require the signing of a formal data use agreement to restrict data usage solely for non-commercial academic research purposes and prevent unauthorized redistribution.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by the Third Batch of Public Welfare Key Special Project of the Veterinary Drug Industry of China Institute of Veterinary Drug Control (Grant No. GY202408).

Authors' contributions

YM conceived and designed the study, performed experiments, analyzed data and drafted the original manuscript；YC analyzed the data;FL contributed to study design;ZC contributed to study design;MC contributed to study design;TL contributed to study design and revised the manuscript;HW contributed to study design and revised the manuscript.All authors read and approved the final manuscript.

Acknowledgements（Not applicable）

References

[1] Report on the use of veterinary antimicrobials in China in 2020. China Animal Husbandry and Veterinary News. 2021-11-14(003). doi:10.28149/n.cnki.ncxmb.2021.000241.

[2] Bush K, Bradford PA. β-Lactams and β-lactamase inhibitors: an overview. Cold Spring Harb Perspect Med. 2016;6(8):a025247. doi:10.1101/cshperspect.a025247.

[3] Jesudason T. WHO publishes list of medically important antimicrobials. Lancet Infect Dis. 2024;24(5):e284. doi:10.1016/S1473-3099(24)00248-2.

[4] Arbab S, Ullah H, Wang W, Zhang J. Antimicrobial drug resistance against Escherichia coli and its harmful effect on animal health. Vet Med Sci. 2022;8(4):1780-6. doi:10.1002/vms3.825.

[5] Wu B, Qi C, Wang L, et al. Detection of microbial aerosols in hospital wards and molecular identification and dissemination of drug resistance of Escherichia coli. Environ Int. 2020;137:105479. doi:10.1016/j.envint.2020.105479.

[6] Nasrollahian S, Graham JP, Halaji M. A review of the mechanisms that confer antibiotic resistance in pathotypes of E. coli. Front Cell Infect Microbiol. 2024;14:1387497. doi:10.3389/fcimb.2024.1387497.

[7] Roth N, Käsbohrer A, Mayrhofer S, Zitz U, Hofacre C, Domig KJ. The application of antibiotics in broiler production and the resulting antibiotic resistance in Escherichia coli: a global overview. Poult Sci. 2019;98(4):1791-804. doi:10.3382/ps/pey539.

[8] Zheng XR, Sun YH, Chang MX, Jiang HX. Plasmid and chromosomal copies of blaCMY-2

mediate resistance to third-generation cephalosporins in Escherichia coli from food animals in China. Vet Microbiol. 2022;271:109493. doi:10.1016/j.vetmic.2022.109493.

[9] Deza-Cruz I, Vilar MJ, Velasova M, AbuOun M, Anjum MF, Smith RP. Antimicrobial resistance of Escherichia coli in the UK: comparison of single vs. pooled samples from healthy pigs. Lett Appl Microbiol. 2023;76(11):ovad123. doi:10.1093/lambio/ovad123.

[10] National Antimicrobial Resistance Monitoring System. NARMS Now: integrated data. US Food and Drug Administration. https://www.fda.gov/animal-veterinary/national-antimicrobial-resistance-monitoring-system/narms-now-integrated-data. Accessed 26 Oct 2026.

[11] Deng Y, Wang T, Du Y, Zhang L, Wang J, Qi Z, et al. Risk assessment for resistance to fludioxonil in Corynespora cassiicola in Liaoning China. Pestic Biochem Physiol. 2023;197:105622. doi:10.1016/j.pestbp.2023.105622.

[12] Hu Y, Jiang L, Sun X, Wu J, Ma L, Zhou Y, et al. Risk assessment of antibiotic resistance genes in the drinking water system. Sci Total Environ. 2021;800:149650. doi:10.1016/j.scitotenv.2021.149650.

[13] Xu Y, Zhu L, Chen S, Wu H, Li R, Li J, et al. Risk assessment and dissemination mechanism of antibiotic resistance genes in compost. Environ Int. 2023;178:108126. doi:10.1016/j.envint.2023.108126.

[14] Huang M, Zhang X, Luo C, Xu H, Qiu Y, Yang J. Genome and antibiotic resistance characteristics of Shigella clinical isolates in Fujian Province, Southeast China, 2005-2019. Microb Genom. 2024;10(11):001325. doi:10.1099/mgen.0.001325.

[15] Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020;75(12):3491-500. doi:10.1093/jac/dkaa345.

[16] European Food Safety Authority; European Centre for Disease Prevention and Control. The European Union summary report on antimicrobial resistance in zoonotic and indicator bacteria from humans, animals and food in 2016. EFSA J. 2018;16(2):e05182. doi:10.2903/j.efsa.2018.5182.

[17] Levin-Reisman I, Ronin I, Gefen O, Braniss I, Shoresh N, Balaban NQ. Antibiotic tolerance facilitates the evolution of resistance. Science. 2017;355(6327):826-30.

doi:10.1126/science.aaj2191.

[18] Senard O, Lafaurie M, Lesprit P, Nguyen Y, Lescure X, Therby A, et al. Efficacy of cefoxitin versus carbapenem in febrile male urinary tract infections caused by extended spectrum beta-lactamase-producing Escherichia coli: a multicenter retrospective cohort study with propensity score analysis. Eur J Clin Microbiol Infect Dis. 2020;39(1):121-9. doi:10.1007/s10096-019-03701-0.

[19] Skarżyńska M, Leekitcharoenphon P, Hendriksen RS, Aarestrup FM, Wasyl D. A metagenomic glimpse into the gut of wild and domestic animals: quantification of antimicrobial resistance and more. PLoS One. 2020;15(12):e0242987. doi:10.1371/journal.pone.0242987.

[20] Cho H, Kim Y, Hassan A, Park KT. Whole-genome sequence-based comparison of antimicrobial resistant diarrheagenic Escherichia coli in pork and chicken production chains in Korea. Int J Food Microbiol. 2025;431:111085. doi:10.1016/j.ijfoodmicro.2025.111085.

[21] Farooq A, Rafique A, Han E, Park SM. Global dissemination of the beta-lactam resistance gene blaTEM-1 among pathogenic bacteria. Sci Total Environ. 2025;963:178521. doi:10.1016/j.scitotenv.2025.178521.

[22] Ejaz H, Younas S, Abosalif KOA, Junaid K, Alzahrani B, Alsrhani A, et al. Molecular analysis of blaSHV, blaTEM, and blaCTX-M in extended-spectrum β-lactamase producing Enterobacteriaceae recovered from fecal specimens of animals. PLoS One. 2021;16(1):e0245126. doi:10.1371/journal.pone.0245126.

[23] Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, Takebayashi Y, et al. β-Lactamases and β-lactamase inhibitors in the 21st century. J Mol Biol. 2019;431(18):3472-500. doi:10.1016/j.jmb.2019.04.002.

[24] Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing. Wayne (PA): Clinical and Laboratory Standards Institute; 2016.

[25] Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol. 2018;19(1):153. doi:10.1186/s13059-018-1540-z.

[26] Landman F, Jamin C, de Haan A, et al. Genomic surveillance of multidrug-resistant organisms based on long-read sequencing. Genome Med. 2024;16(1):137. doi:10.1186/s13073-024-01412-6.

[27] Zhang AN, Gaston JM, Dai CL, et al. An omics-based framework for assessing the health risk

of antimicrobial resistance genes. Nat Commun. 2021;12(1):4765. doi:10.1038/s41467-021-25096-3.

[28] Yang F, Zhang Z, Li Z, et al. Prevalence of high-risk β-lactam resistance genes in family livestock farms in Danjiangkou Reservoir Basin, Central China. Int J Environ Res Public Health. 2022;19(10):6036. doi:10.3390/ijerph19106036.

[29] Onishi R, Shigemura K, Osawa K, et al. The antimicrobial resistance characteristics of imipenem-non-susceptible, imipenase-6-producing Escherichia coli. Antibiotics (Basel). 2021;11(1):32. doi:10.3390/antibiotics11010032.

[30] Bush K, Bradford PA. Epidemiology of β-lactamase-producing pathogens. Clin Microbiol Rev. 2020;33(2):e00047-19. doi:10.1128/CMR.00047-19.

[31] Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics. 2000;16(5):412-24. doi:10.1093/bioinformatics/16.5.412.

[32] Luo H, Xu G, Li C, He L, Luo L, Wang Z, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. Lancet Oncol. 2019;20(12):1645-54. doi:10.1016/S1470-2045(19)30637-0.

[33] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015;10(3):e0118432. doi:10.1371/journal.pone.0118432.

[34] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-38. doi:10.1097/EDE.0b013e3181c30fb2.

[35] Jia W, Qin Y, Zhao C. Rapid detection of adulterated lamb meat using near infrared and electronic nose: a F1-score-MRE data fusion approach. Food Chem. 2024;439:138123. doi:10.1016/j.foodchem.2023.138123.

[36] Huang Q, Jiang Z, Shi B, Meng J, Shu L, Hu F, et al. Characterisation of cardiovascular disease (CVD) incidence and machine learning risk prediction in middle-aged and elderly populations: data from the China Health and Retirement Longitudinal Study (CHARLS). BMC Public Health. 2025;25(1):518. doi:10.1186/s12889-025-21609-7.