J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 3327--3342

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Research on Financial Performance Evaluation of Pharmaceutical Enterprises Based on Big Data Analysis and Their Digital Development Paths

Qingyue Bi[1,✉]

[1] *Liaoning University of International Business and Economics, Dalian, Liaoning, 116052, China*

ABSTRACT

Under the background of the development of digital economy industry, more and more enterprises begin to make attempts of digital change. After constructing the financial performance index system of pharmaceutical enterprises, the study selects 30 pharmaceutical listed companies as the research samples, and evaluates their financial performance by using the principal component analysis method and the collected relevant data. On this basis, the study selects indicators of digitalization degree and puts forward research hypotheses, explores the influence of digitalization degree on the financial performance of pharmaceutical enterprises through correlation analysis, multiple regression analysis and time lag effect analysis, and then puts forward the path of digitalization development of pharmaceutical enterprises in combination with the results of the analysis. The results show that the financial performance of the sample pharmaceutical enterprises is at a medium level, with an average composite score of 0.520, among which pharmaceutical enterprises E10, E6 and E22 have the best performance, with scores above 0.9. The degree of digitization has a negative impact on the financial performance of enterprises at the 1% level, but the coefficient of digital capital investment turns from negative to positive after the lag two period, and there is a time-lag effect of digitization on the financial performance of pharmaceutical enterprises. It is recommended to promote the digitalization of pharmaceutical enterprises by encouraging the cultivation of digital talents, improving the law and cultivating thinking, and building a digital platform.

Keywords: principal component analysis, multiple regression analysis, time lag effect, degree of digitization, financial performance

✉ Corresponding author.
E-mail address: bellebi_21@163.com (Q. Bi).
Received 15 January 2024; Revised 25 March 2024; Accepted 30 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-189

1. Introduction

With the gradual slowdown of the world's economic growth, the digital economy has become a new generation of "leader" to promote economic development, under the double pressure of accelerated population aging and the new crown epidemic, the pharmaceutical industry, as an important industry related to national economy and people's livelihoods, needs to be promoted through the digital transformation of the rapid development of the industry [1-4]. Relevant departments should improve the "Internet + healthcare" service system, improve the "Internet + healthcare" support system, and strengthen industry supervision and security [5-10]. Emphasis on strengthening the application of artificial intelligence, big data and other information technology in the development of intelligent medical equipment and drugs, and actively explore the use of information technology to optimize the process of medical services, creating a comfortable new experience of medical care [11-14]. Innovative development of Internet hospitals, telemedicine, online health consultation, health management and other services, and continue to enhance the popularization of telemedicine facilities and equipment in remote rural areas [15-17]. The importance of digital transformation for the pharmaceutical industry is obvious, and in today's era, the digital development of the pharmaceutical industry will be continuously promoted.

Since the end of the nineteenth century and the beginning of the twentieth century, enterprise performance evaluation has been the focus of attention in the discipline of economics and management, and in the early years, enterprises usually use a series of financial indicators such as return on investment, sales profitability and so on to measure the performance of operation and management, but it is difficult for these indicators to make a correct evaluation of the enterprise's business performance, and it has a certain limitation [18-21]. With the diligent research of countless scholars, the evaluation methods of enterprise performance have been continuously enriched, and scholars have begun to propose a comprehensive system of financial indicators.

The study takes 30 pharmaceutical listed companies as the research object, constructs the index system of financial performance of pharmaceutical enterprises from five dimensions of debt repayment, operation, profitability, growth and cash ability, and collects the index data of relevant samples. After the applicability test, the principal component analysis is used for the determination, interpretation and naming of the common factors, and the comprehensive score of the financial performance of the sample pharmaceutical enterprises is calculated to obtain the financial performance evaluation results of the pharmaceutical enterprises. Then, we introduce the digitization degree index with pharmaceutical enterprises' digital strategy, digital capital investment and digital personnel investment as the three core variables and related control variables, and reflect the financial performance and market performance of pharmaceutical enterprises with the return on net assets and Tobin's Q value to construct the relationship model between the degree of digitization and the performance of pharmaceutical enterprises. Correlation analysis, multiple regression analysis and time lag effect analysis are conducted for each variable to verify the proposed research hypotheses. Finally, the results of the analysis are combined to discuss the digitalization development path of pharmaceutical enterprises.

2. Evaluation and analysis of the financial performance of pharmaceutical companies

In this chapter, 18 financial data indicators are selected to analyze the financial performance of 30 listed biopharmaceutical companies using principal component analysis.

2.1. Principal Component Analysis

Principal Component Analysis (PCA) is an effective method for data dimensionality reduction based on mathematical theory. A large number of indicators with correlation X_1, X_2, Λ, X_p (indicating a total of p indicators) are regrouped and categorized to form a limited number of comprehensive indicators without correlation F_m to replace the original indicators. The extraction of composite indicators is the most critical element, which can ensure that the new indicators remain independent of each other (no overlap of information), but also make it maximize the reflection of the information represented by the original variable X_p, the specific operation method is as follows:

Assuming that F_1 is the first principal component indicator formed by the linear combination of the original variables, i.e., $F_1 = a_{11}X_1 +_- a_{21}X_2 + K + a_{p1}X_p$, according to the relevant basic mathematical theories, it can be found that with the help of the variance of the amount of information extracted from each principal component F_1, the variance $Var(F_1)$ is positively correlated with the amount of information contained therein, and the larger the variance, the more information. Usually, the more information the principal component indicator F_1 has, the more favorable it is for specific research and analysis, so in practice, all linear combinations should be analyzed by ANOVA, and the one with the largest variance among the corresponding linear combinations of X_1, X_2, Λ, X_p should be selected as the principal component indicator F_1, and determined as the first principal component. If the information of the original p indicators is not enough to be represented by the first principal component, it is necessary to further select the second principal component indicator F_2, indicator F_1 and indicator F_2 to ensure the validity of the information and non-repeatable, mathematically speaking, the two indicators are independent of each other and have an irrelevance, that is, the covariance $Cov(F_1, F_2) = 0$, and therefore it is the largest variance among all linear combinations of F_2 and F_1 uncorrelated X_1, X_2, Λ, X_p, determined as the second principal component, and based on this, each principal component of F_1, F_2, Λ, F_m the indicator as the original variable X_1, X_2, Λ, X_p was constructed step by step:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1p}X_p \\ F_2 = a_{21}X_1 + a_{22}X_2 + \ldots + a_{2p}X_p \\ \qquad \ldots\ldots. \\ F_m = a_{m1}X_1 + a_{m2}X_2 + \ldots + a_{mp}X_p \end{cases} \tag{1}$$

The specific steps for the application of principal component analysis are as follows:

1) Calculation of covariance matrix. The covariance matrix is calculated for the sample data:

$$s_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}\left(x_{ki} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_j\right)(i, j = 1, 2, \ldots, p) \tag{2}$$

2) The eigenvalues λ_i of Σ and the corresponding orthogonalized unit eigenvectors a_i are calculated. The first m larger eigenvalues $\lambda 1 \geq \lambda 2 \geq \Lambda > 0$ of Σ correspond to the variances of the first m principal components, and the coefficient λ_i of principal component F_i about the original variable is the corresponding unit eigenvector a_i, which gives the ith principal component F_i of the original variable:

$$F_i = a_i' X \tag{3}$$

The actual amount of information is measured with the help of the variance (or information) contribution of the principal components, a_i for:

$$a_i = \lambda_i \Big/ \sum_{i=1}^{m} \lambda_i \tag{4}$$

3) Selection of principal components. Finally, several principal components need to be selected, i.e., the determination in F_1, F_2, Λ, F_m is $G(m)$ determined by the cumulative contribution of variance (information):

$$G(m)=\sum\lambda/\sum\lambda \tag{5}$$

When the cumulative contribution rate exceeds 75%, it indicates that the information of the original variable can be well reflected, where m indicates the number of principal components extracted.
4) Calculation of principal component loadings. The correlation between principal component F_i and original variable X_j mainly relies on the principal component loadings to measure, and the loadings of original variable $Xj(j=1,2,\Lambda,p)$ on each principal component $F_i(i=1,2,\Lambda,m)$ are:

$$l(Zi,Xj)=\sqrt{\lambda_i a_{ij}}\,(i=1,2,\Lambda,m;\, j=1,2,\Lambda,p) \tag{6}$$

As far as the results of principal component analysis in SPSS software are concerned, the principal component loadings can be directly reflected by the "component matrix".
5) Calculation of principal component scores. The formula for calculating the scores of the samples for different principal components is as follows:

$$F_1=a_{1i}X_1+a_{2i}X_2+\Lambda+a_{pi}X_p, i=1,2,\Lambda,m \tag{7}$$

In most cases, the scale of the indicator will be different in practical application, so the influence of the scale should be eliminated before the calculation of principal components. There are many ways to eliminate the data outline, the common method is to standardize the original data, that is, to do the following data transformation:

$$x_{ij}^*=\frac{x_{ij}-\overline{x_j}}{s_j}i=1,2,\Lambda,m, \text{ of which } \overline{x_j}=\frac{1}{n}\sum_{i=2}^{n}x_{ij}, s_j^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij}-\overline{x_j}\right)^2 \tag{8}$$

2.2. Selection of evaluation indicators

The thesis first combines two aspects of previous research results, from the enterprise profitability, solvency, operating capacity, growth capacity, cash capacity and other five aspects of the pharmaceutical enterprise financial performance index evaluation system to build.

Profitability aspects: return on net assets A11, net operating margin A12, operating profit margin A13, return on assets A14.

For solvency: current ratio A21, quick ratio A22, gearing ratio A23.

For operating capacity: inventory turnover ratio A31, accounts receivable turnover ratio A32, current assets turnover ratio A33, total assets turnover ratio A34.

Growth capacity: net profit growth rate A41, net asset growth rate A42, total asset growth rate A43, operating income growth rate A44.

Cash capacity: cash ratio of operating income A51, cash ratio of operating profit A52, cash current liabilities ratio A53.

2.3. Sample Selection

This article will refer to the relevant information of the relevant enterprises, and select the unlisted formula for the pharmaceutical listed company and other operating conditions, followed by eliminating ST companies and *ST companies, i.e., companies with losses for three consecutive years, as well as companies with incomplete data, and finally selected 30 companies (E1~E30) as the research samples.

2.4. Factor analysis

In this paper, according to the model of factor analysis method, the data of 30 pharmaceutical enterprise indexes will be preprocessed and then factor analyzed, and the process of factor analysis of the paper will be completed in SPSS22.0 system.

1) Suitability test. In the process of using factor analysis, in order to determine the correlation between the selected indicators, it is generally necessary to carry out an adaptability test. The stronger the correlation, the more applicable. In this paper, KMO and Bartlett's test are used to carry out the applicability test.

The results of KMO and Bartlett test are shown in Table 1. Usually, when the KMO value is greater than 0.5 and the significance level of Bartlett value is less than 0.01, it means that there is a high correlation between the variables. From the results, it can be seen that the KMO value is 0.648, which is greater than 0.5, which means that there is a common factor in the original data, and the Bartlett significance is 0.000, which means that the original data is free, so these data can be analyzed by factor analysis.

Table 1. KMO and bartlett test results

KMO sampling availability number		0.648
Bartlett sphericity test	Approximate card	1062.261
	Df	85
	Sig.	0.000

2) Determine the common factor. When the eigenvalues extracted from the factors are greater than 1, they can be called common factors. The total variance explained for each factor is shown in Table 2. There are six factors greater than 1, and the values of the cumulative contribution of the first six factors are 22.564, 39.948, 52.586, 62.228, 69.683, and 75.969, respectively, which indicates that the analysis of the factors is more satisfactory. The cumulative variance contribution rate before rotation was 75.969, and the amount of information covered by the six factors before and after rotation did not undergo a large loss, so it is suitable to extract six factors as public factors.

Table 2. The total variance interpretation of each factor

Constituent	Total	Initial eigenvalue variance%	Cumulation%	Total	Rotation load squared and variance%	Cumulation%
1	4.761	28.451	28.451	3.941	22.564	22.564
2	3.341	16.641	45.092	3.151	17.384	39.948
3	2.167	10.645	55.737	2.374	12.638	52.586
4	1.731	9.761	65.498	1.835	9.642	62.228
5	1.425	6.315	71.813	1.791	7.455	69.683
6	1.052	4.156	75.969	1.673	6.286	75.969

3) Interpretation and naming of common factors. The maximum variance method is used to rotate the component matrix, resulting in a rotated component matrix that can be rationally interpreted for naming the public factors. This paper combines the industry and company will be 60% as the benchmark for the naming of the six public factors. Where the higher the factor loading number, the closer the relationship between the factor and the variable is proved.

Table 3 shows the rotated component matrix. For public factor 1, the return on net assets A11, net operating margin A12, and operating profit margin A13 are all greater than 72%, named as profitability factor. The loadings of factor 2 in current ratio A21, quick ratio A22, and gearing ratio A23 are all greater than 73%, named as solvency factor. The factor loadings of factor 3 in inventory turnover ratio A31, accounts receivable turnover ratio A32, and current asset turnover ratio A33 are 0.708, -

0.603, and 0.618, respectively, and the factor loadings of total asset turnover ratio A34 are 0.890, which are larger, and it is named as the factor of operating ability.

Factor 4 reflects the profitability and profit growth ability of the enterprise in terms of return on assets A14, net profit growth rate A41, and operating income growth rate A44, with factor loadings greater than 65%, and is named the profitability and growth ability factor. Factor 5 is cash ratio of operating income A51, cash ratio of operating profit A52, cash current liabilities ratio A53, the factor loadings are greater than 69%, named as cash ability factor. Factor 6 is Net Assets Growth Rate A42, Total Assets Growth Rate A43, with factor loadings of 0.781 and 0.833, respectively, which reflect the growth of earnings, named as Growth Ability Factor.

Table 3. The component matrix after rotation

	1	2	3	4	5	6
A11	0.728	0.212	0.347	0.263	0.445	0.501
A12	0848	0.324	0.275	0.152	0.557	0.345
A13	0.967	0.584	0.534	0.437	0.243	0.498
A14	0.252	0.445	0.208	0.707	0.376	0.426
A21	0.506	0.733	0.384	0.257	0.412	0.584
A22	0.273	0.862	0.121	0.505	0.508	0.375
A23	0.414	0.775	0.227	0.291	0.192	0.548
A31	0.537	0.116	0.708	0.231	0.519	0.298
A32	0.280	0.350	-0.603	0.478	0.384	0.364
A33	0.204	0.164	0.618	0.561	0.533	0.118
A34	0.399	0.548	0.890	0.405	0.371	0.236
A41	0.277	0.155	0.353	0.825	0.212	0.246
A42	0.487	0.439	0.222	0.491	0.424	0.781
A43	0.487	0.424	0.153	0.212	0.589	0.833
A44	0.264	0.402	0.135	0.651	0.142	0.575
A51	0.578	0.409	0.374	0.145	0.699	0.368
A52	0.362	0.142	0.210	0.399	0.717	0.149
A53	0.452	0.194	0.249	0.406	0.802	0.316

2.5. *Financial performance evaluation results*

The financial performance composite scores and rankings of the companies were calculated by factor analysis of 30 pharmaceutical companies. Table 4 shows the financial performance composite score and ranking of the sample companies. The average financial performance composite score of the sample pharmaceutical companies is 0.520, which is at a medium level. e10, e6 and e22 are ranked in the top three in terms of financial performance level, with composite scores greater than 0.9. 20% of the sample pharmaceutical companies have a composite score of 0.7 or more, and the three companies, e29, e26 and e11, have the worst financial performance, with composite scores lower than 0.2.

Table 4. Financial performance comprehensive score and ranking of the sample company

Enterprise	F	Ranking	Enterprise	F	Ranking
E1	0.486	18	E16	0.665	8
E2	0.515	16	E17	0.608	10
E3	0.542	14	E18	0.758	5
E4	0.386	21	E19	0.709	6
E5	0.501	17	E20	0.838	4

E6	0.916	2	E21	0.367	23
E7	0.451	19	E22	0.909	3
E8	0.389	20	E23	0.339	25
E9	0.316	26	E24	0.665	9
E10	0.927	1	E25	0.538	15
E11	0.129	30	E26	0.138	29
E12	0.379	22	E27	0.571	11
E13	0.543	13	E28	0.362	24
E14	0.264	27	E29	0.181	28
E15	0.545	12	E30	0.676	7

3. Impact of digitization on the financial performance of pharmaceutical companies

As time progresses, the value gained by more and more companies in the market in the development of their digitalization strategies is becoming more and more significant. This section is important to study the impact of the degree of digitization on the financial performance of pharmaceutical companies.

3.1. Methods of multiple regression analysis

Univariate linear regression limits the independent variable affecting a variable to a single variable, but this is not easy to do in most practical problems, and thus the application of correlation statistical analysis to solve practical problems usually requires the use of a mathematical model with a more comprehensive coverage, in which the effects of two or more explanatory variables are estimated separately within the model, which is the main mathematical idea of multivariate statistics.

3.1.1. Multiple linear regression. Let y be a random variable that can be observed and recorded, and it is obviously influenced by $m-1$ a non-random factor $x_1, x_2, \cdots, x_{m-1}$ and a random factor ε. There is a relationship between the non-random, random factors and y based on mathematical and statistical theory, if y and $x_1, x_2, \cdots, x_{m-1}$ have the following linearly related relationship equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{m-1} x_{m-1} + \varepsilon \tag{9}$$

Where $\beta_0, \beta_1, \beta_2, \cdots, \beta_{m-1}$ is the coefficient of the independent variable, ε the mean is zero and the variance is $\sigma^2 > 0$. It is usually assumed that $\varepsilon \sim N\left(0, \sigma^2\right)$. The above model is a multiple regression model where y is the dependent variable and $x_1, x_2, \cdots, x_{m-1}$ is the independent variable.

The first step is to make a reasonable parameter estimation of the unknown coefficients $\beta_0, \beta, \beta_2, \cdots, \beta_{m-1}$, so it is necessary to make $n\left(n \geq p\right)$ observations and record them first, thus obtaining n sets of data:

$$\left(x_{i_1}, x_{i_2}, \cdots, x_{i_{m-1}}; y_i\right)\left(i = 1, 2, \cdots, n\right) \tag{10}$$

They should satisfy $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{m-1} x_{m-1} + \varepsilon$, i.e., have:

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{m-1} x_{m-1} + \varepsilon_1 \\ y = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_{m-1} x_{2m-1} + \varepsilon_2 \\ \vdots \\ y = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_{m-1} x_{nm-1} + \varepsilon_n \end{cases} \tag{11}$$

where $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ are independent of each other and obey a $N\left(0, \sigma^2\right)$-distribution.

Let:

$$Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1},X=\begin{bmatrix} 1,x_{1,1}\cdots x_{1,m-1} \\ 1,x_{2,1}\cdots x_{2,m-1} \\ \vdots \\ 1,x_{n,1}\cdots x_{n,m-1} \end{bmatrix}_{n\times m},\beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m-1} \end{bmatrix}_{m\times 1},\varepsilon=\begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n\times 1} \tag{12}$$

Then the system of multivariate equations can be abbreviated to the following form:

$$\begin{cases} Y=X\beta+\varepsilon \\ \varepsilon\sim N\left(0,\sigma^2 I_n\right) \end{cases} \tag{13}$$

where Y is the observed value and X is the variable observation matrix, which are recorded from the previous data observations, while assuming that X is the variable matrix with full rank of columns, i.e., $r(X)=m$. where β is the column vector of the unknown parameter to be estimated. ε is the column vector of random errors that cannot be effectively observed and recorded. The mathematical formula obtained above is the expression of the matrix of the multiple linear regression model and is denoted as $\left(Y,X\beta,\sigma^2 I_n\right)$.

The central problem to be solved for the above linear model $\left(Y,X\beta,\sigma^2 I_n\right)$ is:

1) Efficient estimation of parameters β and σ^2 to establish the linkage equation between y and $x_1,x_2,\cdots,x_{m-1}$.

2) The assumptions of the proposed linear model and the assumptions of parameter β can be effectively tested.

3) The dependent variable y can be reasonably predicted while the independent variables can be effectively controlled and recorded.

3.1.2. Estimation of regression coefficients. As in the case of the simple regression model, the common estimation method for the linear coefficient $\beta_0,\beta_1,\beta_2,\cdots,\beta_{m-1}$ in the complex model is still the least squares method, denoted as:

$$Q(\beta_0,\beta_1,\cdots,\beta_k)=\sum_{i=1}^{N}\left(y_i-\beta_0-\beta_{1i}-\beta_2 x_{2i}-\cdots-\beta_k x_{ki}\right)^2 \tag{14}$$

Find its minimum value at point $\left(\hat{\beta}_0,\hat{\beta}_1,\cdots,\hat{\beta}_k\right)$, i.e:

$$Q\left(\hat{\beta}_0,\hat{\beta}_1,\cdots,\hat{\beta}_k\right)=\min_{\beta_0,\beta_1,\cdots,\beta_k} Q(\beta_0,\beta_1,\cdots,\beta_k) \tag{15}$$

Then $\hat{\beta}_0,\hat{\beta},\cdots,\hat{\beta}_k$ is the least squares estimate of $\beta_0,\beta,\cdots,\beta_k$.

The linear equation for $\hat{\beta}_0,\hat{\beta}_1,\cdots,\hat{\beta}_k$ is given by the necessary condition for the differential method for the extremum:

$$\frac{\partial}{\partial\beta_j}\sum_{i=1}^{N}\left(y_i-\beta_0-\beta_1 x_{1i}-\beta_2 x_{2i}-\cdots-\beta_k x_{ki}\right)^2=0\left(j=0,1,\cdots,k\right) \tag{16}$$

The solution of the equation. Solve this equation and represent it as a matrix:

$$\hat{\beta}=\left(X^T X\right)^{-1}X^T Y \tag{17}$$

Where: $\hat{\beta}=(\beta_0,\beta_1,\cdots,\beta_k)^T$ and $X^T X$ is required to be invertible.

In this way, the empirical regression equation is obtained as:

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2+\cdots+\hat{\beta}_k x_k \tag{18}$$

As with the simple regression model, Equation σ^2 can be estimated, noting:

$$S_E = \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 = Q\left(\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_k \right) \tag{19}$$

Then the unbiased estimate of σ^2 is:

$$\hat{\sigma}^2 = \frac{S_E}{n-k-1} = \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)}{n-k-1} \tag{20}$$

3.1.3. Tests of significance. The significance of the empirical equation requires a test of significance for the existence of a linear relationship between y and x_i. The original hypothesis can be formulated $H_0 : \beta_0 = 0, \beta_1 = 1, \cdots, \beta_k = 0$. Common statistics used to measure the goodness of fit between the empirical regression equation and the observations are the compound correlation coefficient R and the goodness of fit coefficient, where:

$$R = \sqrt{\frac{U}{U + Q_e}}, R^2 = \frac{U}{U + Q_e} \tag{21}$$

3.2. Research hypotheses and selection of variables

This paper argues that the mechanism of the role of digitalization on the financial performance of pharmaceutical companies exists in the strategic level, the level of infrastructure and human resources and their resources complementary level, and the impact of industry heterogeneity at three levels, which leads to the relevant hypotheses that digitalization mainly affects the financial performance of enterprises from the three aspects of the strategy, the digital capital investment, and the digital personnel investment:

H1: There is a negative relationship between the degree of digitization and financial performance of pharmaceutical companies.

Specifically, the core variables such as the implementation of enterprise digitalization strategy, digital capital investment and digital personnel ratio have a significant negative impact on with the enterprise financial performance (ROE), digitalization investment often in the initial period does not significantly improve the profitability of the enterprise, but may be a burden on the profitability level of the enterprise.

H2: There is a positive relationship between the degree of digitization and market performance of pharmaceutical companies.

Specifically, the core variables of enterprise digital strategy implementation, digital capital investment and digital staff ratio have a significant boost to enterprise market performance (Tobin'Q), which may be mainly due to the fact that the market perceives that the investment in digitization can increase the value of the enterprise after the degree of digitization is increased.

H3: There is a time lag effect of the degree of digitization on the financial performance of pharmaceutical companies.

Enterprises in the first implementation of the data strategy, may be due to the pre-investment is too large, pre-investment and digitalization implementation for the enterprise financial performance has a negative impact, but in the long term there may be a boost to enterprise performance.

Based on the above assumptions, the explanatory variables in this paper are financial performance and market performance, the explanatory variables are digital strategy implementation X1, digital capital investment rate X2 and digital personnel ratio X3, and company growth Z1 (operating revenue growth rate), size Z2 (natural logarithm of total assets), and gearing ratio Z3 (total liabilities at the end of the period/total assets at the end of the period) are the control variables. The pharmaceutical companies in chapter 2.3 are still used as the study sample.

3.3. Correlation analysis

Correlation analysis between the variables is the prerequisite basis for regression analysis, and the correlation statistics of the pharmaceutical enterprise digitization degree indicator variables with the financial and market performance variables are shown in Table 5, with * , **, and *** indicating that they are significant at the 10%, 5%, and 1% significance levels, respectively. Enterprise digitization strategy, digitization capital investment, and personnel investment are all negatively correlated with the financial performance of pharmaceutical enterprises at the 1% or 5% significance level, while they are positively correlated with the market performance of pharmaceutical enterprises at the 1% or 5% significance level, which preliminarily verifies the possibility of the hypotheses H1 and H2 of this paper. In addition, the correlation coefficients between the respective variables are low, indicating that there is no multicollinearity between the variables, which will not interfere with the regression results.

Table 5. Correlation analysis results

	ROE	Tobin'Q	X1	X2	X3	Z1	Z2	Z3
ROE	1.000							
Tobin'Q	-0.236***	1.000						
X1	-0.306***	0.146***	1.000					
X2	-0.107**	0.295**	0.146**	1.000				
X3	-0.197***	0.068**	0.267**	0.357***	1.000			
Z1	0.238***	0.235*	0.291**	-0.059	0.063	1.000		
Z2	0.016**	0.113***	0.082	0.075*	0.049	0.435***	1.000	
Z3	0.142***	0.092**	-0.039	-0.054	0.052*	-0.081*	0.035	1.000

3.4. Regression analysis

Regression analysis of financial performance and market performance of pharmaceutical companies was conducted using the fixed effect model. The results of the regression analysis of the model are shown in Table 6. The financial performance model is significant overall, with a goodness-of-fit R2 of 0.1665, and the core variables are all significant at the 1% level. The effects of digital strategy implementation X1, digital capital investment rate X2 and digital personnel ratio X3 on corporate financial performance were significantly negative at the 1% level, indicating that the degree of digitization is negatively related to the financial performance of pharmaceutical companies, and hypothesis H1 was verified. In addition, the related control variables are also highly significant, and both the debt level Z3 and growth Z1 of pharmaceutical enterprises have positive benefits on the financial performance of enterprises, which indicates that appropriate financial leverage and revenue growth rate have a boosting effect on the financial performance of pharmaceutical enterprises.

The effects of digitalization strategy implementation X1, digitalization capital investment rate X2 and digitalization personnel ratio X3 on pharmaceutical enterprises' market performance under the fixed model are all significantly positive at the 1% or 5% level, with the coefficients of 0.353, 0.242 and 2.337, respectively, which indicates that the degree of digitization of pharmaceutical enterprises is positively related to the enterprises' market performance, and Hypothesis H2 is verified. In terms of economic significance, the market is able to effectively capture the relevant information and perceive that the strategy has an enhancing effect on the value of pharmaceutical companies after the public implementation of their digitization strategy and related inputs.

Table 6. Model regression analysis results

Variable	ROE		Tobin'Q	
	Fixed effect	T statistic	Fixed effect	T statistic

X1	-0.083***	-4.376	0.353***	3.999
X2	-0.005***	-3.938	0.242**	2.199
X3	-0.081***	-3.456	2.337**	2.821
Z1	0.064**	1.904	0.875***	3.098
Z2	0.762	2.092	-0.072***	-2.811
Z3	0.004**	2.803	2.153***	1.003
_cons	-0.092	-0.051	5.502***	4.964
N	653		653	
Prob>F	0.000		0.000	
R2	0.1665		0.1563	
Hausman test value(Prob>chi2)	0.004		0.012	

3.5. Analysis of time lag effects

Since the regression results show that the degree of digitization of pharmaceutical companies has a significant negative impact on financial performance in the current period, in order to explore the possible lagged impact on financial performance, this paper conducts a regression analysis of the model with two lags.

The lagged period regression results of the financial performance model are shown in Table 7. The lagged regression model is still significant, in the model of lag one, digital strategy implementation X1 and digital capital investment rate X2 are still significant, the estimated coefficients are negative, but along with the increase in the number of lag periods and decreasing, which indicates that the negative effect of digitalization is gradually decreasing. In the model of lag two, the estimated coefficient of digital capital investment rate X2 turns from negative to positive and is significant at 1% level, which indicates that there is indeed a certain lag effect on the financial performance of pharmaceutical enterprises, and that digitalization may have an effect of improving the financial performance of the enterprise in one to two years, and the hypothesis H3 is valid.

Table 7. Financial performance model late regression results

Variable	ROEt+1		ROEt+2	
	Fixed effect	T statistic	Fixed effect	T statistic
X1	-0.095**	-4.009	-0.007***	-0.065
X2	-0.051**	-2.377	0.005***	-0.095
X3	-0.036	-1.931	-0.011***	-0.062
Z1	0.064***	6.621	0.091***	4.447
Z2	-1.585**	-2.899	-2.898***	-1.508
Z3	0.182***	3.982	-0.099*	-1.819
_cons	0.358***	2.631	0.485***	2.207
N	537		537	
Prob>F	0.000		0.000	
R2	0.1735		0.1636	
Hausman test value(Prob>chi2)	0.000		0.000	

4. Digital development path for pharmaceutical companies

By analyzing the impact of digital development on pharmaceutical enterprises, although digital transformation in the short term for the enterprise to bring economic benefits is not too significant, or even produce a negative effect, but according to the theory of time lag effect, digital inputs will be in the long term to have a positive impact on enterprise performance, so this paper puts forward the path of digital development of pharmaceutical enterprises.

4.1. Framework construction

The digital development of pharmaceutical enterprises needs to be combined with the global perspective of the entire pharmaceutical industry chain, improve the digital development framework, start from the whole process, and integrate the development of multiple technologies. The digital development framework of pharmaceutical enterprises is shown in Figure 1, starting from the following three aspects:

Digitalization empowers science: empowering scientific research can shorten the cost of drug development and improve the success rate. The current digital technology has been maturely used in the digitization of clinical management, and there is still room for improvement in the use of new drug development.

Accurate digital positioning of pharmaceutical enterprises: Digital transformation of pharmaceutical enterprises has become a hot topic in recent years, but when carrying out digital reconstruction, accurate positioning according to their own characteristics and the environment in which they are located in order to carry out appropriate digital transformation.

Digitalization of medical experience: At present, the user experience of digital technology in the medical field is mainly based on digital diagnostic services and digital medication services. The use of digital technology can not only ease the work pressure of the healthcare team, but also provide patients with a better experience than traditional online consultation.

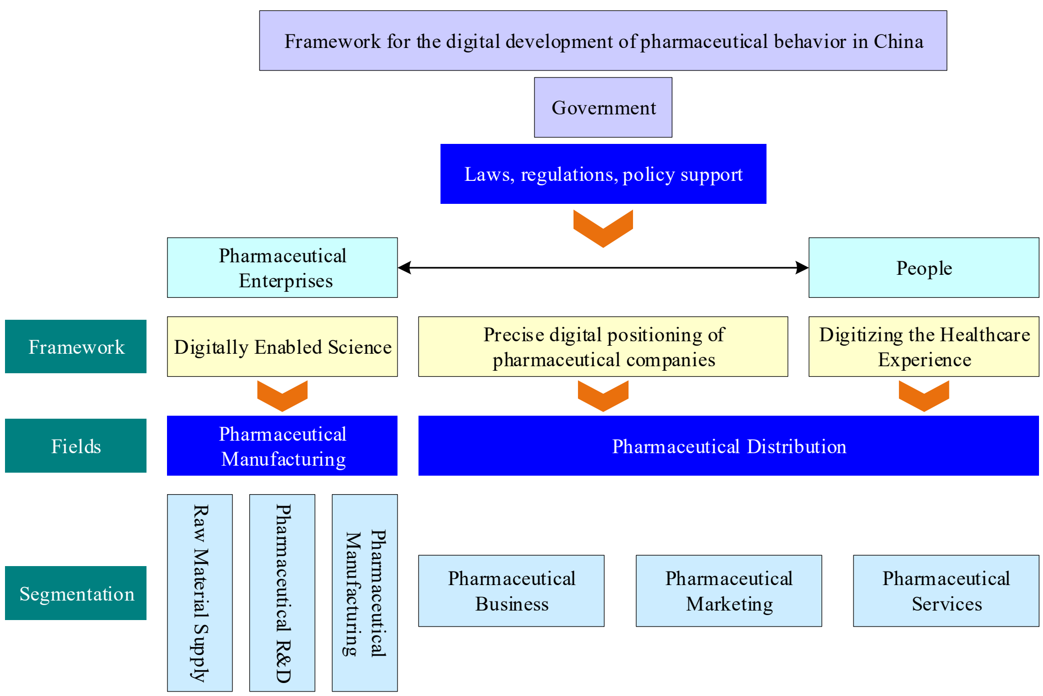

Fig. 1. Framework diagram of the development of digitization in pharmaceutical enterprise

4.2. Specific strategies

The development of the Internet has promoted the change and progress of the pharmaceutical industry, but the development of policies on the use of digitalization is lagging behind, and the digital transformation of pharmaceutical enterprises is faced with the lack of digital technology support for the development of new drugs, the existence of pharmaceutical data security issues, and the lack of service-based platform construction for pharmaceutical enterprises. The specific strategy for the digital development of pharmaceutical enterprises is shown in Figure 2, which promotes the digital development of pharmaceutical enterprises from the three aspects of digital talent training, improving the law and cultivating thinking, and building digital platforms.

4.2.1. Digital talent development. Digital education is conducive to the development of digitalization in the pharmaceutical industry, accelerating the evolution of pharmaceutical enterprises from IT to DT applications. On the one hand, relying on the community, social institutions, universities and other organizations, to carry out "online and offline" integrated education. On the other hand, we are opening up green channels for medical services and integrating digitalization into our lives. In the case of health insurance, it is important to actively create a "one-stop" service for enrollment, payment, and protection, and to simplify the process. At the same time, it is also important to vigorously promote digital technologies such as health insurance e-vouchers and face recognition, and to retain the function of the physical health insurance card, and to carry out multiline services based on the original foundation. In addition, digital education can also be used to cultivate new types of talents and fill the gaps in the field of digital development of pharmaceutical enterprises.

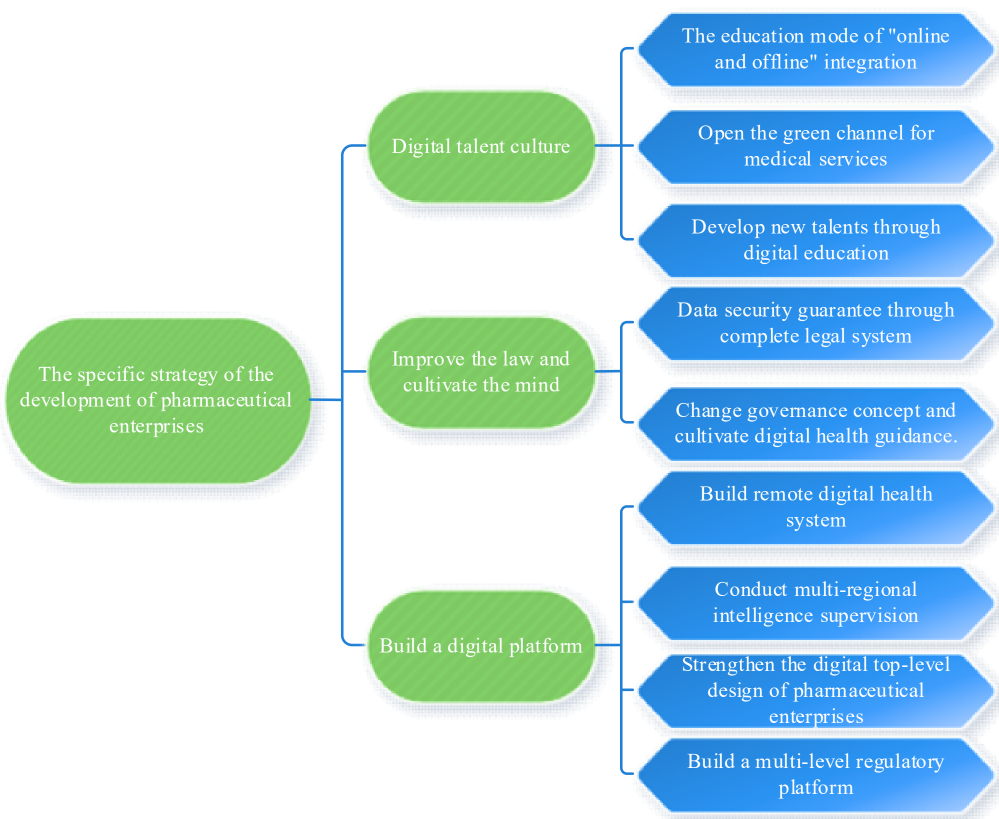

Fig. 2. The specific strategy of the development of pharmaceutical enterprises

4.2.2. Improving Laws and Fostering Thinking. With the digitalization of the pharmaceutical industry, the concept of digital health has emerged. The scope of digital health is broader, involving more complex areas, simply put, there is a need for a complete legal system to build a data "safe house" to ensure that the data does not leakage, but in the data security guarantee, will not be completely blocked to the data, that is, the safe storage of the data at the same time, but also to carry out a reasonable flow of data, to give play to its due value. The value of the data should be realized.

While improving the legal system, it is also necessary to change the concept of governance and develop a digital health guidance mindset. First, based on big data, we will analyse diseases in susceptible populations and regions, carry out early warning of diseases, realize the shift from "single-patient service" to "group service", and speed up service efficiency. Secondly, relying on the data platform and using data rationalization and sharing, we can accurately depict the "health portraits" of individuals and groups, and realize the change from "specific health project management" to "comprehensive health management". Third, we will utilize various digital platforms to carry out data sharing. Thirdly, we use digital platforms to collect, upload and analyze data in real time to achieve real-time dynamic management, realize the change from "passive service" to "active management", and strengthen the deepening of thinking.

4.2.3. Establishment of a digital platform. At the level of the pharmaceutical circulation field, a remote digital medical and health system is constructed, and physicians on the server side can provide convenient medical services for patients in response to the uploaded data. At the level of pharmaceutical supervision, the construction of "drug supervision brain" can be used as a reference to carry out multi-territory intelligent supervision, which can help promote the establishment of drug traceability system. In the future, 5G, blockchain, artificial intelligence and other technologies can be combined to strengthen the construction of the digital top-level design of pharmaceutical enterprises, and at the same time, with the help of the Internet system to create a multi-level regulatory platform of the state, the industry, and the society, and to unify the monitoring, scheduling, and management of all the data, so as to build a more complete regulatory system for better regulatory decision-making.

The construction of the digital platform can be used to solve the difficulties of medical resources in an online diagnosis based on the existing Alibaba health platform and the beauty group health platform.

5. Conclusion

This study evaluates the financial performance of 30 pharmaceutical enterprises using principal component analysis and further proposes a digital development path for pharmaceutical enterprises by studying the impact of the degree of digitization on the financial performance of pharmaceutical enterprises. The related research results are as follows:

1) The composite financial performance score of the sample pharmaceutical enterprises is 0.520, which is at a medium level overall, and 20% of the pharmaceutical enterprises have a composite financial performance score of 0.7 or more, which is at an upper-middle level. Among them, the financial performance composite scores of pharmaceutical enterprises E10, E6 and E22 are all above 0.9, ranking in the top three.

2) Digitalization strategy, digital capital and personnel inputs have a negative effect on corporate financial performance at the 1% level, indicating that the degree of digitization of pharmaceutical enterprises has a significant negative effect on financial performance in the short term. Through the time lag effect analysis, the coefficient of digital capital investment from negative to positive, indicating that the digital input to the financial performance of pharmaceutical enterprises have a certain time lag effect, short-term digital input will have a positive impact in the long term, this effect may be lagged one to two periods of time to show.

3) The development of digitalization promotes the progress of the pharmaceutical industry, which can be better promoted through the construction of top-level design to coordinate the pharmaceutical industry digital platform, improve the law to ensure the safe flow of data, and cultivate digital talents in the pharmaceutical industry with digital education, and other measures to better promote the development of digitalization of pharmaceutical enterprises.

References

[1] Yang, M. , Bento, P. , & Akbar, A. . (2019). Does csr influence firm performance indicators? evidence

from chinese pharmaceutical enterprises. Sustainability, 11.

[2] Kellner, H. , Askari, A. , Kupka, T. , Friccius-Quecke, H. , Furlan, F. , & Hachaichi, S. , et al. (2020). Ab0299real-world assessment of gp2015 (etanercept biosimilar, sdz-etn): an interim analysis of data from the self-injection assessment questionnaire in patients with rheumatoid arthritis in the compact study. Annals of the Rheumatic Diseases, 79, 1449-1449.

[3] Polak, P. , Sinka, I. C. , Reynolds, G. K. , & Roberts, R. J. . (2024). Successful formulation window for the design of pharmaceutical tablets with required mechanical properties. International Journal of Pharmaceutics, 650.

[4] Ahmed, R. R. , Kyriakopoulos, G. L. , Streimikiene, D. , & Streimikis, J. . (2021). Drivers of proactive environmental strategies: evidence from the pharmaceutical industry of asian economies. Sustainability, 13.

[5] Lin, S. . (2021). Financial performance management system and wireless sharing network optimization of listed enterprises under bpnn. Mobile Information Systems.

[6] Yuk, H. , & Garrett, T. C. . (2023). Does customer participation moderate the effects of innovation on cost-based financial performance? an examination of different forms of customer participation1*. Journal of business research.

[7] Aslam, S. , Rehman, R. U. , Naeem, M. A. , & Taghizadeh-Hesary, F. . (2022). Nexus of corporate environmental strategy, environmental performance, and financial performance. The Singapore Economic Review.

[8] Zahid, M. , Rahman, H. U. , Shah, S. A. , & Mata, P. N. . (2021). The interconnection between decent workplace and firm financial performance through the mediation of environmental sustainability: lessons from an emerging economy. Sustainability, 13.

[9] Jin, B. , & Nembhard, I. . (2021). Effects of affiliation network membership on hospital quality and financial performance. Health services research.

[10] Chen, X. U. , & Yuan-Xin, S. . (2017). Drivers of innovation,service innovation and non-financial performance: a test of search strategies and combinative effects in china. R & D Management.

[11] Larsen, Samuel Brüning, Masi, D. , Feibert, D. C. , & Jacobsen, P. . (2018). How the reverse supply chain impacts the firm's financial performance. International Journal of Physical Distribution & Logistics Management, IJPDLM-01-2017-0031.

[12] Ghorbel, H. , & Kolsi, M. . (2022). The effect of governance mechanisms on the financial and stock market performance: the case of canadian companies. European journal of international management.

[13] Adams, D. , Adams, K. , Attah-Boakye, R. , Ullah, S. , Rodgers, W. , & Kimani, D. . (2022). Social and environmental practices and corporate financial performance of multinational corporations in emerging markets: evidence from 20 oil-rich african countries. Resources Policy, 78.

[14] Gartenberg, C. , Prat, A. , & Serafeim, G. . (2019). Corporate purpose and financial performance. Organization Science, 30(1), 1-18.

[15] Kumaran, S. . (2022). Financial performance index of ipo firms using vikor-critic techniques. Finance Research Letters, 47.

[16] Wu, W. , Shi, K. , Wu, C. H. , & Liu, J. . (2022). Research on the impact of information security certification and concealment on financial performance: impact of iso 27001 and concealment on performance. Journal of global information management.

[17] Group, H. B. R. . (2022). The essential link between esg targets and financial performance. Harvard Business Review(5), 100.

[18] Vasileiou, E. , Georgantzis, N. , Attanasi, G. , & Llerena, P. . (2022). Green innovation and financial performance: a study on italian firms. Research Policy, 51.

[19] Dasgupta, R. . (2022). Financial performance shortfall, esg controversies, and esg performance: evidence from firms around the world. Finance Research Letters, 46.

[20] Johl, S. K. , & Toha, M. A. . (2021). The nexus between proactive eco-innovation and firm financial performance: a circular economy perspective. Sustainability, 13.

[21] Zhang, A. , Chen, Y. , Xu, X. , Gao, Y. , & Zhang, L. . (2021). Impacts of resource alertness and change leadership style on financial performance: an empirical study. Journal of Global Information Management, 29(2), 45-60.