of least squares

Jia Wu, Zheng Ji, Min Pi and Tiegang Yi

Shenzhen Hospital of Traditional Chinese Medicine, Guangdong Province, China

**Corresponding author**: Tiegang Yi, Shenzhen Hospital of Traditional Chinese Medicine, Hospital leading department, No 1, Fu Hua Road, 518000, Shen Zhen, China.

**Tel.:** +86-13826557880, **E-mail:** kokyo_82@163.com
Abstract

Background

As an important treatment for the treatment of kidney disease, peritoneal dialysis has been widely studied and applied due to its low cost and easy operation. Given that chronic kidney disease is growing globally, peritoneal dialysis is receiving increasing attention. With the development and popularization of mobile network technology, mobile telematics began to become a mainstream trend. The emergence of mobile telemedicine system is an important result of applying the universal computing concept to medical purposes. However, as users are not familiar with the medical field, telemedicine technology depends to a large extent on the patient's acceptance of the use of them.

Methods

By integrating the experience of clinicians, the remote diagnosis and treatment system of peritoneal dialysis developed by Shenzhen Traditional Chinese Medicine Hospital can monitor the whole course of peritoneal dialysis data of patients. We used statistical methods to empirically analyze the peritoneal dialysis data. By exploring data over a standard duration of time, the filtration rate per minute of the peritoneal dialysis patients using a 1.5% low-calcium peritoneal solution was reduced over time and had a power function relationship which can help to remind incorrect data. The linear equation can be obtained by least square regression of the data after the time of peritoneal effusion and the weight of the effluent deformed.

Results

The least squares method was used to regress the patient's peritoneal dialysis data (logarithm of peritoneal dialysis time and filtration rate per minute), and the regression equation R square was equal to 0.95. The regression coefficient passed the T test and the regression equation fits well. According to the result parameters of the regression equation, we calculated the standard range of filtration rate for each peritoneal dialysis. Taking 441 cases of a random patient as an example, 438 cases of diafiltration rate met the standard range. 3 cases were filtered out below the standard.

Conclusions
The system can inform the patients of the results according to the confidence interval of the regression prediction, which greatly strengthens the interaction of the system and increases the patients' compliance.

**Key words:** Remote peritoneal dialysis system; least square method; CAPD; Filtration rate per minute
1. Background

The worldwide prevalence of long-term dialysis continues to rise[1]. In recent years, a concept of renal rehabilitation has become widely known among nephrology specialists, dialysis specialists, kidney transplantation specialists, rehabilitation specialists, nutrition specialists, guideline specialists, nurses, physiotherapists, and representatives of patients[2]. A large proportion of patients use CAPD (continuous ambulatory peritoneal dialysis) to treat renal failure. The role of CAPD in renal replacement programmes has now become more clearly defined whilst patient and technique survival have steadily improved[3]. Peritoneal dialysis has been highly valued by the medical industry, and the government is actively promoting the use of peritoneal dialysis treatment programs. The exciting mushrooming of scientific knowledge on various aspects of peritoneal dialysis has led to a better understanding of the technique[4].

Peritoneal dialysis is one of the most important ways of renal replacement therapy which was usually treated by home treatment, while its efficacy is largely restricted by the patients' normative and follow-up regularity. Poor patient compliance to therapy results in a worsening condition that often increases healthcare costs[5]. However, due to the lack of follow-up data, there are some defects in the current peritoneal dialysis follow-up model, and it is difficult to correct the patients’ irregular operation in time.

With the improvement of computer technology and the improvement of the level of medical information application, the construction of "Medical + Internet” mode provides a feasible solution for remote follow-up diagnosis and treatment of certain diseases[6]. Telemedicine has been used more and more in medical care. It provides patients with relevant information and telemedicine services[7]. This is good news for patients living in remote communities and remote areas. It can support patients living in remote areas and help them maintain good independence while ensuring good results. A telemedicine platform would allow for greater patient independence while instilling a greater degree of confidence that well-trained professionals are closely monitoring the therapy and are readily available for assistance and it may also have Clinical benefit[9].

At present, various remote systems have been developed to assist in the treatment of peritoneal dialysis. Caogen Hong has realized the remote management of peritoneal dialysis by means of information management. They have established a peritoneal dialysis remote management system including medical client, medical APP, and patient APP. It replaced the manual recording of telephone center and telephone follow-up old management mode[10]. Gallar conducted a medical conversation session through a video conferencing device installed in each patient's home[11]. Yajuan Cai, Yan Zhang, etc. can control and real-time guidance through remote monitoring and management, reducing the chance of contact infection in peritoneal dialysis patients. By applying video surveillance technology to the remote management of peritoneal dialysis patients, combined with the knowledge of nursing, they can better take care of peritoneal dialysis patients[12]. Nayak et al. developed an application for tablets that make PDs more accessible and used as an interactive user guide to support patients. They use tablet computers to show patients teaching, complete training, and feedback the patient’s operation back to the patient. The Columbia Remote Automated Peritoneal Dialysis System, used by Alfonso Bunch, provides two-way communication between patients and clinical teams through the introduction of remote patient
monitoring. By comparing the effects of increasing RPM (Remote patient monitoring) and not increasing RPM, it is found that the use of monitoring patients significantly improves the patient's compliance, the patient's treatment effect is also better than the no rpm group [14].

We found that although the past remote peritoneal systems used various methods to improve the interaction between doctors and patients and enhance the efficacy, they were just tools for remote conversations and rarely mentioned how to use the data of remote diagnosis and treatment to improve the quality of diagnosis and treatment. Although the way of online communication and interaction is effective, it is difficult to ensure long-term stability. The model of human interaction monitoring requires medical staff to actively and patiently serve. Long-term observations are prone to fatigue burden, and some minor differences in the process of peritoneal dialysis are sometimes difficult to find. In order to improve the follow-up effect of patients with CAPD at home, reduce the risk of complications caused by irregular follow-up, and improve clinical efficacy and quality of medical services effectively, a peritoneal dialysis telediagnosis and treatment system was developed by the medical staff of Nephrology Department of Shenzhen Hospital of Traditional Chinese Medicine in conjunction with IT Engineer team. The system includes two parts: PC backstage management and mobile APP. By integrating the Internet of Things technology, the data of monitoring indicators such as daily dialysis ultrafiltration volume and body weight change value of home CAPD can be collected. By using the least squares method to return the historical peritoneal data of the patient, it is possible to monitor the abnormality of peritoneal dialysis in one aspect.

2. METHODS

2.1. Introduction to Data Acquisition and Operation of Peritoneal Dialysis Remote System

The peritoneal dialysis remote diagnosis and treatment system in our hospital transmits the peritoneal data to medical personnel in a timely manner through the combination of the Internet of Things and the peritoneal treatment. The medical staff supervises the entire treatment process and can well solve the shortcomings of current CPAD treatment. The weight of peritoneal dialysis fluid is weighed during peritoneal dialysis, and the weight of peritoneal dialysis fluid is obtained by Bluetooth device. Then the data of peritoneal dialysis fluid (weight of peritoneal dialysis fluid, time of peritoneal dialysis, type of peritoneal dialysis fluid, ID of peritoneal dialysis patient) are transmitted to the background system of peritoneal dialysis in real time through mobile network. In addition, medical staff can also manage many aspects of patients' peritoneal dialysis information, including patient file management, information management of peritoneal dialysis treatment, peritoneal dialysis management, follow-up management.

2.2. Patients and Data

From January 2017 to June 2019, 36 patients with peritoneal dialysis in nephrology department of our hospital were enrolled in this study. Among them, 15 were males and 21 were females, aged 20-68 years. All patients were treated with Baxter peritoneal dialysis fluid. The peritoneal dialysis data selected in this study were those of patients using peritoneal dialysis fluid with calcium concentration of 1.5%. In this experiment, we need to obtain data from the peritoneal dialysis system as follows: duration of
peritoneal dialysis $t$, the weight of inflow $\text{inweight}$ and the weight of effluent $\text{effweight}$ (The filtration rate per minute $v_t$ is calculated from the available data). The duration of peritoneal dialysis $t$ is equal to the time point of peritoneal dialysis fluid minus the time point of infusion (unit: minute). Considering the patient's compliance and the regularity of peritoneal dialysis operation, we consider that the operation of peritoneal dialysis time in the range of 180-600 minutes is an effective peritoneal dialysis, and the peritoneal dialysis data that do not meet the time range is considered as an ineffective peritoneum. The standard range of quality control of Baxter peritoneal dialysis fluid is $[2.2\pm0.08]$ kg. If the weight of fluid inflow uploaded by patients is not in this range, we believe that the operation of the patient is wrong and does not conform to the operation specifications, these data will be excluded from this experiment. The filtration rate per minute $v_t$ is equal to the fluid output divided by the peritoneal time. The patients’ original data are taken from the telediagnosis and treatment system of peritoneal dialysis in Shenzhen Chinese Medicine Hospital. After data processing, there are 20,608 valid data that meet the standards in total and shown in Table 1 below:

| Table 1 Descriptive statistics of filtration rate and peritoneal time per minute |
|---------------------------------|-------------|------------|------------|-------------|-------------|-------------|
|                                | Number of cases | range | Minimum value | Maximum value | Average value | Standard deviation | Variance  |
| Filtration rate per minute (kg/min) | 20608 | 1.83 | .96 | 2.78 | 1.8508 | .00216 | .30999 | .096 |
| Peritoneal dialysis time (min)   | 20608 | 420 | 180 | 600 | 367.96 | .748 | 107.374 | 11529.210 |
| Number of valid cases            | 20608 |          |      |      |       |     |       |      |

Through the transformation of the Internet of Things electronic scale equipment, the data of the abdominal communication is collected to the system database by means of the Internet of Things carrier. All patients are aware of the experimental content and have signed an informed consent form, and the patient can freely withdraw at any time. The withdrawal will not have any effect on the conventional treatment. The patients’ identity data are entered into the system by the nursing staff and maintained regularly. The patients themselves are free to choose whether or not to upload the diagnosis and treatment data.

2.3. Least Square Regression for Peritoneal Dialysis Data

The weight of effluent, the filtration rate per minute, and the duration of peritoneal dialysis can be calculated from the data uploaded by the peritoneal system in patients with CAPD. Since the value of the filtration rate per minute is too small, in order to facilitate the observation of the relationship between the data, we increase the filtration rate per minute by a factor of 1,000. Take peritoneal dialysis time (unit:
minute) as the x-axis, and the filtration rate per minute (expanded 1,000 times) to draw a scatter plot for the y-axis. We randomly select one patient’s data and consider using a linear regression function first, let \( v_t \) denote the filtration rate per minute (multiplied by 1000 times), t denotes the peritoneal time, regression of \( v_t \) and t by least square method can be obtained:

\[
v_t = -0.017t + 13.107 \quad R^2 = 0.9014 \quad \text{(Formula 1)}
\]

The relationship between \( v_t \) and t can be displayed well by the above-mentioned formula. It can be seen that \( R^2=0.9 \) has a high overall fitness and the graph is shown in Figure 1 below:

---

Insert Figure 1

---

According to Figure 1, it can be found that the filtration rate per minute is related to time, the Pearson correlation coefficients of the two are calculated by SPSS. The results are shown in Table 2 below:

<table>
<thead>
<tr>
<th>Table 2</th>
<th>The correlation between the filtration rate per minute and the duration of peritoneal dialysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filtration rate per minute</td>
<td>Peritoneal dialysis time</td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>1</td>
</tr>
<tr>
<td>Significant (two-tailed)</td>
<td></td>
</tr>
<tr>
<td>Number of cases</td>
<td>20608</td>
</tr>
<tr>
<td>Peritoneal dialysis time</td>
<td>Pearson correlation</td>
</tr>
<tr>
<td>Significant (two-tailed)</td>
<td></td>
</tr>
<tr>
<td>Number of cases</td>
<td>20608</td>
</tr>
</tbody>
</table>

**. At 0.01 level (double tails), the correlation was significant.

The filtration rate per minute decreases obviously with the increase of time, but the decline rate decreases slowly with the increase of time, and the curve is approximately in the shape of a power function. The Pearson correlation coefficient of the filtration rate per minute and the time calculated in Table 2 is -0.939. The correlation results are significantly effective. There is a strong negative correlation between them.

However, it is obvious that there is a significant difference between the fitted curve and the actual value in the partial interval range, the model cannot show the change trend between \( v_t \) and t. There is room for improvement of linear regression function model. Considering that the actual curve is in the shape of a power function, assuming that:

\[
v_t = at^b
\]

Among them, a and b are unknown parameters, t is the time of peritoneal dialysis, and the filtration rate per minute is \( v_t \) when the peritoneal time is t. After the logarithm transformation on both sides of the formula, we can get:
\[ \ln v_t = lna + b\ln t \]

Let \( v_t^* = \ln v_t \), \( a^* = \ln a \), \( t^* = \ln t \), we can get a linear relationship between \( v_t^* \) and \( t^* \):

\[ v_t^* = a^* + bt^* \quad \text{(Formula 2)} \]

Taking the selected data as an example in the same way, after logarithmizing \( v_t \) and t, the obtained data scatter plot is shown in Figure 2 below:

[Insert Figure 2]

2.4. Data quality control method

Since the regression performed used the least squares method, we can make interval estimation predictions for the results. By setting the confidence interval confidence, we can calculate the confidence interval range of the regression. If the patient's final uploading data is outside the confidence interval, we can remind the patient and the medical staff through the peritoneal system that the data may exceed or fall below the normal data range. The statistical methods used are as follows:

Let the regression equation calculated according to the least square method be:

\[ \hat{Y} = kX + b \]

Where \( \hat{Y} \) is the predicted value of the dependent variable Y, X is the independent variable, and k and b are the estimated values of the regression, respectively. The model's estimate can be given by the linear regression equation of the sample. There is a certain error between this prediction and the actual value. Therefore, giving a confidence interval for the prediction at a certain neutral level can make our predictions better targeted.

\[ \hat{Y} - Y \sim N\left(0, \left[1 + \frac{1}{T} + \frac{(\hat{k} - Y)^2}{\sum(X_i - \bar{X})^2}\right] \sigma^2\right) \quad \text{(Formula 3)} \]

Which \( \hat{Y} \) represents the value of the independent variable to be predicted, which \( \hat{Y} \) corresponds to the predicted value \( \hat{Y} \) calculated by the regression equation, \( \sigma^2 \) represents the total variance, and can be replaced by the sample variance, \( \sigma^2 = \frac{\sum(Y_i - \bar{Y})^2}{(T - 2)} = \frac{\sum(Y - \bar{Y})^2}{(T - 2)} \), T is the number of samples. According to Equation 1, the t statistic can be constructed to obtain a prediction interval with a confidence level of \( 1 - \alpha \):

\[ \hat{Y} - Y \sim N\left(0, \left[1 + \frac{1}{T} + \frac{(\hat{k} - Y)^2}{\sum(X_i - \bar{X})^2}\right] \sigma^2\right) \quad \text{(Formula 4)} \]

Among them, the statistic \( \frac{\hat{Y} - Y}{\frac{1}{T}} \) can be obtained according to the look-up statistical table. All other variables can be calculated. Therefore, we can set a certain deviation range according to the above formula and the expected confidence. In the process of peritoneal dialysis, if the actual peritoneal fluid output data is not within the predicted interval, we can send reminders to the relevant personnel to check...
whether there is a leak in the peritoneal flow. In fact, we only need to know that formula 4 can be used to predict reasonable intervals. Statisticians have shown that the formula is reliable and that all variables can be calculated. What we do is to apply this to the diagnosis of peritoneal dialysis.

3. RESULTS

3.1. Regression results

Logarithmic transformation was performed on the peritoneal time and the filtration rate respectively. Linear regression was performed on the whole data using SPSS 24.0. The regression results were tested as shown in Table 3 and Table 4 below:

<table>
<thead>
<tr>
<th>Table 3</th>
<th>ANOVA³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Sum of Squares</td>
</tr>
<tr>
<td>1</td>
<td>Regression</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
</tr>
<tr>
<td></td>
<td>Total</td>
</tr>
</tbody>
</table>

³. Dependent Variable: Filtration rate per minute(lnVt)

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Regression coefficients and t statistics:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Unstandardized Coefficients</td>
</tr>
<tr>
<td></td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
</tr>
<tr>
<td></td>
<td>Peritoneal dialysis time(lnt)</td>
</tr>
</tbody>
</table>

a. Dependent Variable: Filtration rate per minute(lnVt)

The regression function is:

\[ lnv_t = -1.085ln t + 8.218 \quad R^2=0.953 \quad \text{(Formula 5)} \]

\( R^2 \) is equal to 0.95, and the overall fitness of the equation is extremely high. As there is only one variable in this regression model, according to the F value obtained from the ANOVA table, it can be known the coefficient of the regression equation is not zero. According to the t-statistic of the coefficient table, the constant term and the independent coefficient value are not zero. The influence of the change of independent variable on dependent variable is effective, so the regression function is linearly valid.

According to the test results, the peritoneal dialysis time is an important explanatory variable of the filtration rate per minute, and the logarithm of the filtration rate per minute and the time of peritoneal dialysis is linear. That is, when other conditions remain unchanged, the logarithm of the dialysis time increases by one unit, the logarithm of the filtration rate per minute decreases by 1.085 units. For the deformation of Formula 5 Formula 6 is available:

\[ v_t = 3707.1 \times t^{-1.085} \quad \text{(Formula 6)} \]

Formula 6 is a power function relationship based on the time of peritoneal dialysis as an independent
variable and the filtration rate per minute as a dependent variable. Let $M_t$ be volume of effluent at the time of peritoneal dialysis time $t$, then:

$$M_t = V_t \times t$$  \hspace{1cm} \text{(Formula 7)}

Formula 6 is substituted into Equation 8, which means:

$$M_t = 3707.1 \times t^{-0.085}$$  \hspace{1cm} \text{(Formula 8)}

### 3.2. Quality control results

In the paper, we used the least squares regression method to regress the logarithm of the filtration rate per minute of peritoneal fluid and the logarithm of the peritoneal time. While what we want to emphasize is the clinical practical value of reviewing the results of the equation. In this paper, we explain that we can predict the weight of peritoneal fluid in the future based on historical data, but it is not widely used in the actual treatment. However, by comparing the difference between the predicted value and the actual value, we can judge whether the operation of home-based peritoneal dialysis is standard in real time, which has more realistic significance for remote diagnosis and treatment. From the figure.3 blow we have achieved the time-to-time calibration of abdominal effusion volume errors using statistical methods:

---

Insert Figure 3

---

Fig. 3 is a scatter plot of logarithm ($v$) and logarithm ($t$) of peritoneal dialysis time per minute of a patient in our hospital. The confidence level is 95%. During a peritoneal dialysis, the prediction interval will be calculated according to the regression equation calculated from the previous data. If the final data uploaded by the patient is greater than the maximum value of the prediction interval, we will mark the change point as orange red. If the actual point is less than the minimum value of the prediction interval, it will be marked as green, and the other normal points as yellow. Fig. 3 shows that the patient had three low and none high peritoneal dialysis filtration rates. We can apply this model method to peritoneal dialysis system to realize immediate reminder. People always remember what just happened more clearly, patients are easier to review the previous operation steps, and medical staff can reduce some pressure and concentrate more on other aspects.

### 4. Analysis Results and Discussion

The peritoneal dialysis system can set a certain error range in the background. According to the calculation results of the peritoneal dialysis curve function equation and the actual uploaded data comparison, it can detect whether there are problems in the diagnosis and treatment process of patients by whether the output volume exceeds the error range, and then provide a possible basis for the establishment of the prevention and control points of the corresponding clinical events. At the same time, according to the specific situation, suggestions for improvement of patients’ treatment plan were put forward, and an effective online and offline interactive follow-up mechanism and complication risk prevention and control model were established to improve treatment effect.

For those points beyond the confidence interval, we are very confident that there are some problems
in this peritoneal dialysis. We can find the main cause in time according to the specific situation. However, we can’t guarantee that there is no problem with this peritoneal dialysis for those points that meet the confidence interval. Further exploration is still needed in the future. Besides, the regression equation is only valid for all patients using 1.5% Baxter peritoneal dialysis fluid, and the regression fitting effect for other concentrations such as 2.5% peritoneal dialysis fluid is not ideal. It’s worthy of further research.

5. CONCLUSIONS

In summary, the development of peritoneal dialysis system has promoted the standardized collection of peritoneal data, provided an efficient way for large-scale data census and analysis. However, in the past, remote diagnosis and treatment systems are basically used to record data, provide remote real-time video, audio, image carriers, etc. Few systems use data calculation to achieve the function of automatic decision-making machine. Through the analysis and processing of peritoneal dialysis data, medical staff can reduce the medical risk caused by complications, improve work efficiency, improve treatment effect, and patients also benefit a lot from it. The use of regression equations for interval estimation of peritoneal fluid volume has important clinical significance. Based on the research purpose, the research on the data of the peritoneal dialysis, through continuous optimization and improvement of the model, gradually enhance the accuracy of the prediction of peritoneal data, and broaden the application scenarios of such data. The development and application of peritoneal dialysis system aims to improve the level of peritoneal medical services, realize the real-time transmission of treatment data, promote communication between doctors and patients, improve the mode of follow-up, timely grasp the patient's condition, adjust the treatment plan, and promote the improvement of curative effect. With the development of the peritoneal information system, the integrity and reliability of the peritoneal data are guaranteed, which greatly reduces the risk of data loss and omission, and provides strong support for diagnosis and treatment. In this paper, the regression function between the filtration rate per minute of peritoneal dialysis fluid and the time of peritoneal dialysis can provide certain basis for medical staff to judge the effect of peritoneal dialysis. Our goal is to help realize the automatic diagnosis and treatment relying on data through telemedicine system. Although there are still some shortcomings and limitations in use, we hope that we can promote the intelligent management of data rather than just regard the remote system as a tool without thinking.
Availability of data and materials

The datasets analysed during the current study are available from the corresponding author on reasonable request.

Abbreviations

CAPD: Continuous ambulatory peritoneal dialysis

Acknowledgments

The authors thank the team members of the computer center of Shenzhen Traditional Chinese Medicine Hospital for participating in the system construction, the nephrology department of Shenzhen Traditional Chinese Medicine Hospital for participating in the collaborative peritoneal dialysis team, and the Jirong Zhong team of Shenzhen City Hidohear Technology Ltd. for their technical support and help in the project development.

Funding Information

This work was supported by grant of no. B2019043 from the Guangdong Medical Science and Technology Research Fund Project, China. Compliance with Ethical Standards. The funders played no role in study design, data collection, analysis, and interpretation or manuscript writing.

Competing interests

The authors declare that they have no competing interests.

Ethical declarations

All procedures in the study involving human participants have passed the ethical review of Shenzhen Hospital of traditional Chinese medicine. Written informed consent was obtained from the patient for publication of case report and any accompanying images. A copy of the consent form is available for review and can be provided on request.

Informed consent

Informed consent was obtained from all individual participants included in the study.
REFERENCES


**Figure legends**

**Fig. 1** Scatter plot of filtration rate per minute and peritoneal dialysis time (The data came from a randomly selected patient)

**Fig. 2** Perform a logarithmic transformation on the filtration rate per minute and the peritoneal time to draw a scatter plot (The data came from the selected patient in Fig. 1)
Fig. 3 Evaluation of outlier markers, the dot marked as an orange point as an abnormal point (the data come from a randomly selected patient)

\[
\ln v = -1.0579 \ln t + 1.0941
\]

\[R^2 = 0.9743\]