Efficient Monitoring of Fall Risk in the Elderly via Functional Data Analysis

Abstract

**Background:** A disease screening service is a preventive healthcare service. Accurate and efficient disease examination based on continuous monitoring data obtained from diseased subjects is the basis of developing this service. The traditional disease screening method for a specific disease is designed according to the physical meaning of the collected data.

**Methods:** In this paper, a general disease detection statistical model based on monitoring data is proposed. By analyzing the distribution of data obtained from subjects who may have certain diseases, we used functional data analysis to establish a statistical model and obtain an efficient algorithm for parameter estimation.

**Results:** The proposed model is applied to a real example of an elderly fall risk screening service based on plantar pressure data collected from elderly individuals walking over obstacles. Reasonable intervals of the model parameters used to screen the elderly for fall risk are obtained from the training samples, which are used to estimate the fall risk of the elderly with the test samples.

**Conclusions:** The study shows that the foot plantar pressure measured in screening tests can be characterized by functional data analysis, and a linear mixed effect model can be used when time points are fixed. The restricted maximum likelihood technique is used for parameter estimation, and a nonlinear optimization algorithm is employed to iteratively determine the model parameters. This paper is to provide a method of detecting falls in the elderly based on statistical data rather than the physical meaning of collected data.

**Keywords:** Disease screening, Fall risk screening, Functional data analysis, Plantar pressure
Background

Disease screening refers to the implementation of quick physical or laboratory examinations to identify patients with potential disease risk. Specifically, the main objective is to identify patients who are in the subclinical stage or are high-risk subjects [1]. Disease screening is a preventive health care service with a broad spectrum of detection possibilities, including falls in the elderly, cancer genes, childhood autism, osteoporosis, prenatal risk factors and other conditions.

Traditional disease screening evaluation methods include choosing the “gold standard” (the most reliable and authoritative diagnostic method that can reflect the actual disease situation is called the “gold standard”), selecting the research subjects, determining the sample size, designing a synchronous blind test (using the experiment to evaluate subjects from the case group and noncase group determined by the “gold standard”), analyzing data, performing quality control and other tasks [1]. The traditional disease screening procedure mainly includes collecting data with detection devices, analyzing the physical meaning of the data corresponding to the disease itself, and then performing disease screening. However, this traditional research method lacks certain universality.

The emergence of preventive medicine and health care services is a considerable benefit to society. The screening reports generated after participating in a screening service are very helpful for disease detection in subjects. Hence, an accurate and appropriate screening method is not only the key to such services but also an important basis for the evaluation of the sustainable operation of such services in the future. A valuable report will enable more people to use disease screening services. Therefore, it is of great social and practical significance to develop accurate and efficient examination methods to correctly identify potential subjects with diseases.

Fall risk screening for the elderly is a type of disease screening. In the United States, approximately 30% of people over 65 years of age fall at least once a year, and 50% of people over 80 are likely to fall [2]. In China, the likelihood of elderly people
falling in a year is approximately 34%. Approximately 38% of the elderly population in urban communities is injured after a fall, and some individuals may die [3]. Falls have become the leading cause of medical treatment for individuals age 65 and older in China in recent years. In addition to damaging the physical health of the elderly, falls also endanger their mental health by reducing their quality of life and increasing the burden on their families and society [4]. In 2014, the China National Health Planning Commission officially included “prevention of falls” in its “core information on health for the elderly” policy. Since then, the risk of falls has gradually attracted increasing social attention [5].

Qualitative assessments of fall risk for the elderly have mainly been based on two methods: the clinical observation method and the scale evaluation method [6]. The clinical observation method is mainly based on clinical observations by an outpatient doctor, nurse or home care worker for an elderly person who has a risk of falling or has a recent history of falls. Although this method is comprehensive [7], it is highly random and lacks standard criteria; additionally, the evaluation results from different doctors can be very different. Hence, this approach is not an objective and quantitative evaluation method [8]. In contrast, the evaluation results of the scale assessment method are more universal, thus supporting the consistency of this approach. Notably, this type of approach has been adopted by most hospitals, nursing homes and health service centers to assess the risk of falls for elderly people. Scale assessments are based on the Morse Falls Assessment Scale (MFS) [9], the Berg Balance Scale (BBS) [10], the Tinetti Scale (POMA) [11], and the Hendrich II Fall Risk Model [12].

Research on quantitative detection methods of fall risk for the elderly is based on different types of fall detection devices, which are mainly divided into three categories: wearable devices, environmental sensing devices and video image devices. When an elderly person falls from an upright position to a lying position, there is usually a sudden increase in negative acceleration. In this process, waist mounted accelerometer equipment can detect fall behavior [13]. Additionally, falls are detected by triaxial accelerometers mounted at the waist, wrist and head using acceleration
thresholds [14]. Fall detection systems based ground tremors and ground vibration information are also used to assess the fall risk of elderly individuals [15]. A fall detection system is equipped with a camera at a fixed indoor position to analyze the collected image information, assess an individual’s activity status and determine whether an elderly has fallen [16]. An image detection system identifies a fall by calculating the vertical distance between the head of an individual and the ground [17]. The video images of individuals collected by an image detection device are analyzed, and the skeleton structure is studied to determine the occurrence of falls [18].

Functional data analysis involves data modeling using functions or functional parameters. The complexity of functions is not assumed to be known in advance, so approximation methods that are flexible based on the data requirements are used. This approach can be used in the healthcare field [19], such as in the classification of Alzheimer’s patients [20], assessments of the rate of clinical alarms in an intensive care unit [21], the determination of heart failure [22] and the identification of diabetes mellitus [23].

In this paper, a statistical model is proposed to identify potentially diseased patients based on functional data analysis, which is an objective quantitative method. The parameter estimation methods for the model and the corresponding algorithm are given. The proposed model is applied for fall risk screening in an elderly population. A numerical analysis of the model parameters is performed, and a reasonable range of the model parameters is obtained. On this basis, it is feasible to screen the elderly for fall risk. The main contribution of this paper is to provide a method of detecting falls in the elderly based on statistical data rather than the physical meaning of collected data.

**Methods**

**Data Collection and Description**

The data were collected with an elderly fall risk detection device developed by the Shenzhen Institute of Advanced Technology of the Chinese Academy of Sciences.
We recruited elderly volunteers who were willing and able to perform testing in a health service center in Luohu District, Shenzhen. We assigned IDs to these elderly volunteers and obtained their basic information (age, gender, height, weight, etc.). The length of the test section was 6 meters, and the height of the obstacles was 6 centimeters. The elderly volunteers who were tested proceeded to the designated road test area, and the remaining volunteers remained in the rest area. The rest area was approximately 1 meter away from the starting point of the test road. The elderly volunteers wore special insoles during the test. The insoles had plantar pressure sensors that could sense changes in plantar pressure and feed the data back to the corresponding equipment. The test area included flat road followed by a 6-centimeter-high obstacle and then flat road, forming a 6-meter-long test section. Each elderly volunteer walked this section twice, and plantar pressure data were collected and stored.

Each elderly volunteer was required to repeat the test $n$ times. The data from the $j$th measurement were recorded as $y_j(t), j = 1, 2, \cdots, n$. In general, the foot plantar pressure was divided into three phases: the increasing pressure phase, the stable pressure phase and the decreasing pressure phase, which correspond to three phases of walking, i.e., starting, crossing obstacles and stopping, respectively.

**The Proposed Method**

This paper examines the problem from a new perspective. First, we collect detection data for all the screening subjects and investigate the optimal parameter estimation scheme. Second, reasonable ranges of distribution parameters are identified, a unified screening standard is obtained, and all detection objects are classified. This approach does not treat each subject separately. In addition, this approach can be generalized and extended for the screening of various diseases.

We assume that the continuous detection data are collected as a time series. Given a sample of subjects, one key objective is building an appropriate model $y_j(t)$ to fit the data. It is also important to estimate reasonable ranges of model parameters based on training with sample data to distinguish subjects with or without disease. Then, the
the proposed model can be used for detection.

Based on the research questions above and the background of the health screening service, to facilitate the construction of the statistical model $y_j(t)$ and subsequent analysis, we assume that (1) the subjects receiving the health screening services are independent; (2) the subjects are able to complete the entire detection process without human assistance; and (3) the subjects need to complete multiple repeated detection tests from the same disease screening service, with each test being independent of the others.

Here, we define some notations to facilitate model construction. $t$ represents the time points at which the measurements from subjects are taken. Notably, the range of $t$ is $t \in A_j \triangleq [0, T_j]$, where $j = 1, 2, \cdots, n$ and $n$ is the number of repeated tests for the same disease screening service. It is assumed that there are $n_j$ points, that is, $t_{n_j} \triangleq T_j$, $T_j \leq 1$.

$h_i(t)$ is the piecewise mean function of the medical index data when the subjects complete the same detection test $n$ times, where $i = 1, 2, \cdots, m$ and $m$ indicates the number of stages into which a complete service is divided.

$\beta_i$ is the fixed effect, indicating the piecewise mean function parameter for different service stages in the same detection test, and $i = 0, 1, 2, \cdots, m$.

$\delta_{jk}$ is the random effect, which characterizes the variations within tests repeated $n$ times from the same subject, where $k = 0, 1, 2, \cdots, q$. This variable is assumed to be independent of time $t$ and follows a normal distribution as $\delta_{jk} \sim N(0, \varphi_k^2)$, where $j = 1, 2, \cdots, n$. To increase model flexibility, $\delta_{jk}$ is expressed as $\text{Cov}(\delta_{jk}, \delta_{jl}) = \varphi_{kl}$, where $k \neq l$.

$z_k(t)$ is a piecewise function corresponding to random effects, where $k = 1, 2, \cdots, q$.

$\varepsilon_{jt}$ is the error term for each test at different time points; therefore, this error is random if $t$ is a constant. We assume that the error term follows a normal
distribution as \( \varepsilon_{jt} \sim N(0, \sigma_j^2) \) and is independent of \( \delta_{jk} \), where \( j = 1, 2, \ldots, n \). We allow the error term to have different variances for different tests involving the same subject.

\( y_j(t) \) represents the response value of the medical index in the \( j \)th test at time point \( t \in A_j \), where \( j = 1, 2, \ldots, n \).

In this subsection, we employ statistical theory and a model to study disease screening. The model is built to fit the data as follows:

\[
y_j(t) = \beta_0 + \sum_{i=1}^{m} \beta_i h_i(t) + \delta_{j0} + \sum_{k=1}^{q} \delta_{jk} z_k(t) + \varepsilon_{jt}, \quad t \in A_j, j = 1, 2, \ldots, n \quad (1)
\]

\( y_j(t) \) is composed of three parts. The first part \( \beta_0 + \sum_{i=1}^{m} \beta_i h_i(t) \) indicates the fixed effects and reflects the differences among detection tests. The second part \( \delta_{j0} + \sum_{k=1}^{q} \delta_{jk} z_k(t) \) indicates the random effects and reflects the differences within tests. The third part is the error term \( \varepsilon_{jt} \), which represents the error of the measured medical index. Matrix theory is used to simplify and analyze the model

\[
\beta \triangleq (\beta_0, \beta_1, \beta_2, \ldots, \beta_m)^T, H(t) \triangleq (1, h_1(t), h_2(t), \ldots, h_m(t))^T
\]

and

\[
\delta_j \triangleq (\delta_{j0}, \delta_{j1}, \delta_{j2}, \ldots, \delta_{jq})^T, Z(t) \triangleq (1, z_1(t), z_2(t), \ldots, z_q(t))^T. \quad (2)
\]

Then, the model can be written as

\[
y_j(t) = \beta^T H(t) + \delta_j^T Z(t) + \varepsilon_{jt}. \quad (3)
\]

Based on the assumptions that \( \delta_j \) and \( \varepsilon_{jt} \) are independent and follow the normal distributions, we have

\[
y_j(t) \sim N(\beta^T H(t), Z^T(t) \Sigma Z(t) + \sigma_j^2). \quad (4)
\]

where \( \Sigma = (\phi_{kl})_{q \times q} \) is the covariance matrix for the vector \( \delta_j \). As time \( t \) changes, the data can be characterized by a Gaussian process model defined as

\[
\bar{y}_j \triangleq \bar{y}_j(t) = (y_j(t_1), y_j(t_2), \ldots, y_j(T_j))^T. \quad (5)
\]

With the probability distribution function for \( y_j(t) \), the joint probability distribution function of \( \bar{y}_j \) is

\[
\bar{y}_j \sim N(\bar{\mu}_j, \Sigma_j). \quad (6)
\]
with

\[
\mu_j = (\beta^T H(t_1), \beta^T H(t_2), \ldots, \beta^T H(T_j))^T, \\
\Sigma_j = Z \Sigma_\delta Z^T + \sigma_j^2 I,
\]

(7)

(8)

where \( Z \) is an \( n_j \) by \( q+1 \) matrix with the \( i \)th row equal to \( Z(t_i) \) and \( I \) is the identity matrix. As a result, the probability density function of \( \bar{Y}_j \) is as follows:

\[
f(\bar{Y}_j) = (2\pi)^{-\frac{n_j}{2}} |\Sigma_j|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\bar{Y}_j - \mu_j)^T \Sigma_j^{-1} (\bar{Y}_j - \mu_j) \right\}.
\]

(9)

where \( |\Sigma_j| \) is the determinant of \( \Sigma_j \) and \( \Sigma_j^{-1} \) represents the inverse matrix of \( \Sigma_j \).

Now, we have built a statistical model for the data collected from the disease screening tests and derived the probability distribution of the response (6) and the corresponding joint probability density function (9). We note that if the multiple time points \( t_1, t_2, \ldots, T_j \) are fixed, the collected data are longitudinal or panel data. Consequently, the task of Gaussian process model estimation can be converted to longitudinal data modeling using a linear mixed effect model.

**Results**

In this section, we derive the likelihood function of the proposed model (6)–(8) based on the normal probability density function (9). Therefore, the maximum likelihood theory in statistical analysis is used for the parameter estimation.

For convenience, we define two sets of parameters

\[
\sigma \triangleq (\sigma_1, \sigma_2, \ldots, \sigma_n)^T, \varphi \triangleq (\varphi_1, \varphi_2, \ldots, \varphi_q, \varphi_{kl})^T, \ k \neq l.
\]

(10)

The joint probability density function of the responses from all the detection tests, that is, the likelihood function, is

\[
L(\beta, \sigma, \varphi) \triangleq L(\beta, \sigma, \varphi; \bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_n) = \prod_{j=1}^{n} f(\bar{Y}_j)
\]

\[
= (2\pi)^{-\frac{1}{2}} \prod_{j=1}^{n} |\Sigma_j|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} \left[ (\bar{Y}_j - \mu_j)^T \Sigma_j^{-1} (\bar{Y}_j - \mu_j) \right] \right\}.
\]

(11)

Then, the log likelihood function is obtained by taking the natural logarithm of (11)

\[
\ln L(\beta, \sigma, \varphi) = \ln L(\beta, \sigma, \varphi; \bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_n)
\]
\[ -\frac{1}{2} \sum_{j=1}^{n} \left[ n_j \ln 2\pi + \ln |\Sigma_j| + (\bar{y}_j - \bar{\mu}_j)^T \Sigma_j^{-1} (\bar{y}_j - \bar{\mu}_j) \right] \]  

Consequently, maximum likelihood estimation can be used to determine the parameters in the model based on profile estimations of \( \beta, \sigma \) and \( \varphi \). Taking the partial derivative of (12) with respect to \( \beta, \sigma \), and \( \varphi \) yields

\[
\frac{\partial \ln L(\beta, \sigma, \varphi)}{\partial \beta} = 0, 
\] (13)

\[
\frac{\partial \ln L(\beta, \sigma, \varphi)}{\partial \sigma} = 0, 
\] (14)

\[
\frac{\partial \ln L(\beta, \sigma, \varphi)}{\partial \varphi} = 0. 
\] (15)

From (13), we can obtain

\[
\hat{\beta} = (X^T \Sigma_j^{-1} X)^{-1} X^T \Sigma_j^{-1} \bar{y}_j, 
\]

where \( X \) is an \( n_j \) by \( m+1 \) matrix with the \( i \)th row equal to \( H(t_i) \). However, the maximum likelihood estimates of the variance and covariance parameters \( \sigma \) and \( \varphi \) are biased. Hence, we adopt the restricted maximum likelihood estimation (REML) method to obtain \( \sigma \) and \( \varphi \). As a result, equations (14) and (15) are converted to the following forms

\[
\frac{\partial (\ln L(\beta, \sigma, \varphi) - \frac{1}{2} \ln |X^T \Sigma_j^{-1} X|)}{\partial \sigma} = 0, 
\] (16)

\[
\frac{\partial (\ln L(\beta, \sigma, \varphi) - \frac{1}{2} \ln |X^T \Sigma_j^{-1} X|)}{\partial \varphi} = 0. 
\] (17)

We iteratively solve for three parameters \( \beta, \sigma \) and \( \varphi \) until convergence is achieved.

We built a statistical model for the data collected from the disease screening tests and performed detailed parameter estimation. However, solving equations (16) and (17) is computationally extensive since the computational complexity is at least \( O(N^2) \), especially when \( N \) is large for high-dimensional data sets, where \( N = \max_{j=1,2,\ldots,n} n_j \).

Additionally, equations (16) and (17) for obtaining parameters \( \sigma \) and \( \varphi \) do not have analytical solutions because they are nonlinear with respect to these parameters. Therefore, we need to employ a nonlinear optimization algorithm, such as the Newton-Raphson algorithm, the coordinate descent algorithm or the FR conjugate gradient optimization algorithm, to obtain the solutions.
Discussion

In the test, the foot plantar pressure of elderly individuals in each test was recorded when they walked across predesigned obstacles. Each elderly individual was required to repeat the test four times. Therefore, \( n = 4 \) for model (1). The data were collected for individuals living in the communities of Luohu District, Shenzhen city, China, who travel to the health service center to use the fall risk screening service. The service period was from Dec. 1, 2016, to Dec. 31, 2016. During the service period, 71 elderly individuals participated in the tests and completed the entire service. After removing some abnormal data, we obtained 62 samples of 17 males and 45 females. The ages of participants range from 54 to 93 years old, with heights ranging from 145 to 175 cm and weights ranging from 37 to 77 kg. The foot plantar pressure data were recorded as time series. The total number of time points for each test was approximately 200.

The whole data set was randomly split into two sets with 50 training samples and 12 testing samples. We used the R program to perform parameter estimation for the model. After multiple simulations and a comparison of the results, we selected \( m = 3 \) and used \( h_1(t) = t, h_2(t) = t^2 \) and \( h_3(t) = t^3 \). Additionally, we set \( q = 1 \) and \( z_1(t) = t \), which corresponded to a random intercept and slope model. In addition, to reduce the number of parameters and simplify the calculation, we set the maximum number of time points for each test to \( N = 50 \). Hence, the model included the parameters \( \beta_0, \beta_1, \beta_2, \beta_3, \varphi_1, \varphi_2, \varphi_{12}, \sigma_1^2, \sigma_2^2, \sigma_3^2, \text{ and } \sigma_4^2 \). The proposed model was run for each elderly individual, and 50 sets of parameter estimates were obtained corresponding to the 50 training samples. Then, the reference intervals for each parameter were obtained based on the training data.

The reference intervals and average values for each parameter are shown in Table 1. Considering that the randomly divided samples include high-risk falls and low-risk falls in the elderly population, we use the 1/4 quartile and the 3/4 quartile of the parameters estimated based on the training samples as the lower and upper limits of
the reference intervals for each model parameter to reduce the variations resulting from the differences among samples corresponding to a high risk of falls and a low risk of falls. From Table 1, the estimated intervals of $\beta_0, \beta_1, \beta_2$, and $\beta_3$ are $[2.56, 4.25]$, $[23.20, 34.14]$, $[-36.85, -12.88]$ and $[-14.09, 2.90]$, respectively. The estimated intervals of $\varphi_1^2, \varphi_{12},$ and $\varphi_2^2$ are $[2.54, 16.48]$, $[-15.81, -1.74]$ and $[2.42, 19.35]$, respectively. The estimated intervals of $\sigma_j^2$ are $[0.91, 2.05]$, $[0.88, 4.48]$, $[1.01, 2.84]$ and $[0.95, 4.25]$.

Table 1 can be used to screen the fall risk of the elderly; notably, the foot plantar pressure data measured from the elderly can be input into the model constructed in section 2 to obtain the parameter estimates, which can be compared to the corresponding reference intervals in Table 1. If the estimated values of most parameters are within the corresponding intervals, an individual can be considered to have a low risk of falls. In contrast, if most of the parameter estimates are not within the reference intervals, the risk of falls is considered high.

|      | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\varphi_1^2$ | $\varphi_{12}$ | $\varphi_2^2$ | $\sigma_1^2$ | $\sigma_2^2$ | $\sigma_3^2$ | $\sigma_4^2$ |
|------|-----------|-----------|-----------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 1/4  | 2.56      | 23.20     | -36.85    | -14.09    | 2.54         | -15.81       | 2.42         | 0.91         | 0.88         | 1.01         | 0.95         |
| 3/4  | 4.25      | 34.14     | -12.88    | 2.90      | 16.48        | -1.74        | 19.35        | 2.05         | 4.48         | 2.84         | 4.25         |
| average | 3.44 | 29.86     | -28.43    | -4.44     | 10.11        | -9.28        | 14.24        | 2.43         | 3.72         | 2.15         | 2.94         |

The reference intervals of the model parameters are given by 50 training samples. We use these intervals to determine the fall risk of the remaining 12 test samples. The results show that 3 of the 12 test samples are conservatively classified as having a high risk of falls, and the remaining 9 samples are classified as having a relatively low risk of falls. Tables 2 and 3 are obtained by taking 7 representative samples as an example, where Table 2 shows the parameter estimates for elderly individuals with a relatively low risk of falls and Table 3 reports the parameter estimates for those with a high risk of falls. The parameter estimates outside the reference intervals are indicated
by different shades. Table 2 shows that more than half of the parameter estimates are within the reference intervals provided in Table 1, so these individuals can be considered to have a relatively low risk of falls.

Table 2 Parameter estimates for elderly individuals with a relatively low risk of falls based on the testing samples

| Number | \(\beta_0\) | \(\beta_1\) | \(\beta_2\) | \(\beta_3\) | \(\varphi_1^2\) | \(\varphi_1\) | \(\varphi_2^2\) | \(\sigma_1^2\) | \(\sigma_2^2\) | \(\sigma_3^2\) | \(\sigma_4^2\) |
|--------|------------|------------|------------|------------|----------------|-------------|----------------|----------------|----------------|----------------|----------------|
| 27     | 3.73       | 20.12      | -9.40      | -13.69     | 1.72           | -1.84       | 1.74           | 0.97           | 2.51           | 1.16           | 1.08           |
| 39     | 2.80       | 44.74      | -74.45     | 28.81      | 10.08          | -12.38      | 17.33          | 5.94           | 2.71           | 5.61           | 2.23           |
| 41     | 3.62       | 27.31      | -32.07     | 1.41       | 3.53           | -4.04       | 7.46           | 2.84           | 0.80           | 0.56           | 3.39           |
| 62     | 3.64       | 37.28      | -59.19     | 19.39      | 5.51           | -4.60       | 3.86           | 1.49           | 1.10           | 0.86           | 1.36           |

Conversely, Table 3 shows the parameter estimates for three individuals with parameters outside the intervals provided in Table 1. For example, the participants with ID numbers 2 and 57 have 10 parameter estimates outside the reference intervals, and participant number 6 has 9 parameter estimates outside the reference intervals. Thus, these individuals have a relatively high risk of falls.

Table 3 Parameter estimates for the elderly individuals with a relatively high risk of falls based on the testing samples

| Number | \(\beta_0\) | \(\beta_1\) | \(\beta_2\) | \(\beta_3\) | \(\varphi_1^2\) | \(\varphi_1\) | \(\varphi_2^2\) | \(\sigma_1^2\) | \(\sigma_2^2\) | \(\sigma_3^2\) | \(\sigma_4^2\) |
|--------|------------|------------|------------|------------|----------------|-------------|----------------|----------------|----------------|----------------|----------------|
| 4      | 1.30       | -0.94      | 2.58       | -2.69      | 0.94           | -0.62       | 0.49           | 0.03           | 0.17           | 0.00           | 0.24           |
| 6      | 3.15       | 22.77      | -7.37      | -18.70     | 0.13           | 0.32        | 1.66           | 2.88           | 0.78           | 2.49           | 0.86           |
| 57     | 0.19       | -0.34      | 1.19       | -0.73      | 0.05           | -0.04       | 0.07           | 0.01           | 0.02           | 0.03           | 0.00           |

Considering the estimation results for 7 elderly individuals in Tables 2 and 3, Fig. 1 and Fig. 2 display the corresponding foot plantar pressure curves over time. Figure 1 shows the curves of the participants who were screened as having a relatively low risk of falls, as denoted in Table 2. Figure 2 depicts the curves of the individuals who were screened as having a relatively high risk of falls, as denoted in Table 3.
According to the National Health Planning Commission's rankings of the risk of falls among the elderly, the severity of fall risk is divided into high risk, intermediate risk and low risk classes. A comparison of Fig. 1 and Fig. 2 indicates that the foot plantar pressure curves in Fig. 1 are quite different from those in Fig. 2, implying that the elderly individuals can be divided into two fall classes: low risk of falls and high risk of falls. That is, the fall risk of the four individuals with plots in Figure 1 is low, and the fall risk of the three individuals in Figure 2 is intermediate and high. Moreover, we can observe that the foot pressure curves of participants 4 and 57 are much different than those of other participants. Hence, these two individuals can be classified as having a high risk of falls, and participant 6 can be classified as having an intermediate risk of falls.

Fig. 1 The gait plantar pressure curve distribution of the elderly with a relatively low risk of falls

Fig. 2 The foot plantar pressure curve distributions for participants with a relatively high risk of falls
Conclusions

Based on a disease screening service, this paper studies the distribution of data collected from screening tests, and a statistical model for disease detection is built considering both fixed and random effects. The proposed model is applied in the context of a fall risk screening service for elderly individuals. The study shows that the foot plantar pressure measured in screening tests can be characterized by functional data analysis, and a linear mixed effect model can be used when time points are fixed. The restricted maximum likelihood technique is used for parameter estimation, and a nonlinear optimization algorithm is employed to iteratively determine the model parameters. Based on the training samples, the reference intervals for each parameter are constructed, and the elderly can be classified into low or high fall risk classes by determining whether the corresponding parameter estimates are inside or outside the reference intervals.

From this work, we have identified several research directions for the future. (1) In this work, only the foot plantar pressure data from fall detection focused on elderly individuals are used. Notably, we could also consider data from other medical indices and include them in the model. (2) In this work, we use $h_1(t) = t$, $h_2(t) = t^2$ and $h_3(t) = t^3$ as the covariates corresponding to the fixed effects. We could also consider other characteristics of participants, such as height, weight, and age. (3) Due to the limitations of the research conditions, the real data samples are only for elderly individuals who participated in the service within a fixed time and at a certain location. The estimated intervals of parameters obtained through these samples are not representative of the general elderly population. Hence, it may be impossible to accurately identify individuals with a high risk of falls. Thus, the size of the random sample set from multiple cities across China should be increased to obtain the parameter reference intervals.

List of abbreviations

Not applicable.
Declarations

Ethics approval and consent to participate
The study was approved by the Research Ethics Committee of the Southeast University. All participants were informed about the study goals and signed informed consent. For participants who were considered illiterate, written informed consent was also obtained from their legal guardians.

Consent for publication
Not applicable.

Availability of data and materials
The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Funding
This work was supported by the National Natural Science Foundation of China [grant number 71531004]. The grant funders were not involved in the design of the study, the collection, analysis and interpretation of data, and in writing the manuscript.

Authors’ contributions
HW designed the project and wrote the final version of the manuscript. CZ and XK conducted the data analysis. XD conceived the analysis. All authors have read and approved the manuscript.

Acknowledgements
The authors wish to thank Lei Wang of the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, for providing data on falls in the elderly.

References
1. Guo Q. Health management. Beijing: People's Medical Publishing House; 2015.
2. Zhao H H. Related factors and prevention of falls in the elderly. Journal of Clinical Nursing. 2005; 4(3): 43-4.
3. Zhu Y M, Yuan H B, Chen L. A survey on the risk factors of falls in the elderly. Nursing Practice and Research. 2007; 4(10): 5-7.
4. Er Y L, Duan L L, Ye P P, Gao X, Wang Y, Deng X, Ji C R, Jin Y. Analysis on the characteristics of unintentional injuries among the elderly from Chinese National Injury Surveillance System, 2014. Chinese Journal of Health Education. 2016; 32(4): 312-7.
5. Yin Y, Ren J G. The influence factors of older adults’ falls in urban China. Lanzhou Academic Journal. 2017; (11): 157-73.
6. Jia G F. Study on the quantitative evaluation method of human balance ability. Doctoral Dissertation of Zhejiang University, 2015.
7. Wang Y L. Rehabilitation function assessment. Beijing: People’s Medical Publishing House. 2013.
8. Rubenstein L Z, Josephson K R, Osterweil D. Falls and fall prevention in the nursing home. Clinics in Geriatric Medicine. 1996; 12(4): 881-902.
9. Morse J M, Morse R M, Tylko S J. Development of a scale to identify the fall-prone patient. Canadian Journal on Aging. 1989; 8(4): 366-77.
10. Berg K. Measuring balance in the elderly: development of an instrument. Physiotherapy Canada. 2009; 41(41): 304-11.
11. Duncan P W, Wiener D K, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. Journal of Gerontology. 1990; 45(6): M192–7.
12. Hendrich A, Nyhuis A, Kippenbrock T, et al. Hospital falls: development of a predictive model for clinical practice. Applied Nursing Research: ANR. 1995; 8(3): 129-39.
13. Mathie M J, Coster A C F, Lovell N H, et al. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiological Measurement. 2004; 25(2): R1-20.
14. Kangas M, Konttila A, Winblad I, Jamsa T. Determination of simple thresholds for accelerometry-based parameters for fall detection. 29th IEEE Annual International Conference on Engineering in Medicine and Biology Society. 2007; 1367-70.
15. Alwan M, Rajendran P J, Kell S, Mack D, Dalal S, Wolfe M, Felder R. A smart and passive floor-vibration based fall detector for elderly. 2006 2nd International Conference on
16. Vaidehi V, Ganapathy K, Mohan K, Aldrin A, Nirmal K. Video based automatic fall detection in indoor environment. 2011 International Conference on Recent Trends in Information Technology (ICRTIT). Chennai, Tamil Nadu, 2011; 1016-20.

17. Shoaib M, Dragon R, Ostermann J. View-invariant fall detection for elderly in real home environment. 2010 Fourth Pacific-Rim Symposium on Image and Video Technology, Singapore. 2010; 52-7.

18. Chen Y T, Lin Y C, Fang W H. A hybrid human fall detection scheme. IEEE 2010 17th International Conference on Image Processing (ICIP 2010). Hong Kong, 2010; 3485-8.

19. Yang P, He J. Functional Regression with Mode-Sparsity Constraint. IEEE International Conference on Data Mining. IEEE, 2016.

20. Alicia N R, Rafael D, Jose L M, Carmen L. Classification of Alzheimer’s Patients through Ubiquitous Computing. Sensors. 2017; 17(7): 1679-96.

21. Kalavagunta P. The effect of Procedural changes on the rate of clinical alarms in the intensive care unit. Dissertations & Theses Gradworks, 2015.

22. Schluck G, Wu W, Whyte J, Abbott L. Emergency department arrival times in Florida heart failure patients utilizing Fisher-Rao curve registration: A descriptive population-based study. Heart & Lung. 2018; 47(5): 458-64

23. Szczesniak R D, Li D, Duan L L, Altaye M, Miodovnik M, Khoury J C. Longitudinal Patterns of Glycemic Control and Blood Pressure in Pregnant Women with Type 1 Diabetes Mellitus: Phenotypes from Functional Data Analysis. American journal of perinatology. 2016; 33(13): 1282-90.