An efficacy evaluation method for non-normal outcomes in randomized controlled trials

Yang Li1,2,3, Zhang Zhang2,3, Qian Feng5, Danhui Yi1,2,3 & Fang Lu4

Randomized controlled trials (RCT) are widely used in clinical efficacy evaluation studies. Linear regression is a general method to evaluate treatment efficacy considering the existence of confounding variables. However, when residuals are not normally distributed, parameter estimation based on ordinary least squares (OLS) is inefficient. This study introduces an exponential squared loss (ESL) model to evaluate treatment effect. The proposed method provides robust estimation for non-normal data. Simulation results show that it outperforms ordinary least squares regression with contaminated data. In the mild cognitive impairment (MCI) efficacy evaluation study with traditional Chinese medicine, our method is applied to construct a linear efficacy evaluation model for the difference in Alzheimer’s disease assessment scale-cognitive (ADAS-cog) scores between the final and baseline records (ADASFA), with the existence of confounding factors and non-normal residuals. The results coincide with existing medical literatures. This proposed method overcomes the limitation of confounding variables and non-normal residuals in RCT efficacy studies. It outperforms OLS in estimation efficiency in situations where the percentage of non-normal contamination reaches 30%. These advantages make it a good method for real-world clinical studies.

Mild cognitive impairment (MCI) is a syndrome defined as a cognitive decline, which may affect daily activities. The amnesic subtype of MCI has a high risk of progression to Alzheimer’s disease and could lead to a prodromal stage of this disorder1. Alzheimer’s disease assessment scale-cognitive (ADAS-cog) subscale measures the progression of MCI in 11 relevant fields, namely spoken language ability, comprehension of spoken language, recall of test instructions, word-finding difficulty, following commands, naming, constructions, ideational praxis, orientation, word recall, and word recognition. Detailed information on ADAS-cog subscale can be found in Rosen et al.2.

Institute of Clinical Pharmacology at Xiyuan Hospital conducted a phase III randomized clinical trial to evaluate the efficacy of a traditional Chinese prescription on MCI. The double blinded randomized clinical trial was conducted in eight qualified medical centres across China with 216 patients allocated to the treatment arm and 108 retained as control. Two patients dropped out from each arm, resulting in 320 complete observations in the final dataset. Data on the difference between final ADAS-cog score and baseline scores (ADASFA) were recorded for the efficacy study. Previous literatures have used ADASFA in efficacy evaluation of MCI or Alzheimer’s disease3,4. The most intuitive idea is to test whether the treatment and control means are equal. However, Morgan and Rubin5 argued that the baseline equivalence is not guaranteed although the allocation is randomized. Imbalance in baseline covariates could confound the statistical test when comparing ADASFA between the two arms. Ten variables, specifically age, height, weight, gender, education, ethnicity, occupation, centre, drug (whether the patient took a drug for MCI in the past three months), and ADAS1 (the baseline record of ADAS-cog6), were recorded as potential covariates. Table 1 shows descriptive statistics of these variables. The explorative covariance analysis presented in Table 2 indicates that the variable of centre and ADAS1 may confound the efficacy evaluation of ADASFA. This implies that a linear regression model should be involved rather than using a simple statistical test in this study, that is,

1Center for Applied Statistics, Renmin University of China, Beijing, China. 2School of Statistics, Renmin University of China, Beijing, China. 3Statistical Consulting Center, Renmin University of China, Beijing, China. 4Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing, China. 5School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia. Correspondence and requests for materials should be addressed to F.L. (email: deerfang@126.com)
where \( p \) is the number of covariates. We denote a binary variable \( x_j = 1 \) to represent the treatment arm and \( x_j = 0 \) for the control arm. The efficacy can be evaluated by the corresponding coefficient \( \beta_j \).

Ordinary least squares (OLS) is a general parameter estimation method for simple linear regression which performs as the best linear unbiased estimation when assuming independent identical normally distributed errors:

\[
\varepsilon \sim \mathcal{N}(0, \sigma^2).
\]

However, the QQ plot in Fig. 1 shows that the MCI dataset may not follow a normal distribution and a Shapiro-Wilk test (\( W = 0.9283, p\text{-value} = 2.799e-11 \)) also suggests a similar result. The contaminated non-normal part may come from either measurement error or mixed distribution, which is commonly presented in medical studies. This could lead to inefficient efficacy estimation by using OLS since the contaminated part is not addressed. A more robust estimation method in linear regression is, therefore, required in such studies.

Many robust methods have been discussed in literatures. Bao developed a rank-based estimate in linear regression. Wang et al. proposed an robust estimation via least absolute deviation while Wang et al. introduced an exponential squared loss (ESL) to select variables robustly. Since the breakdown point of ESL is almost 50%, we adopt it in the MCI efficacy evaluation study. Numerical studies show that the proposed method can achieve a more accurate estimation with a large proportion of contamination in the dataset. Additionally, the estimations are consistent with OLS when contamination proportions are relatively low. Therefore, it can be used as a complementary efficacy evaluation method in real-world clinical studies regardless of the presence or lack of contaminations.

Table 1. Descriptive statistics of variables.

| Discrete | Categories | Sample size | Treatment arm | Control arm | Continuous | Treatment arm | Control arm | Median |
|----------|------------|-------------|---------------|-------------|------------|---------------|-------------|--------|
| gender   | male       | 88          | 46            |             | age        | 62.75 ± 7.90  | 63.77 ± 8.28 | 62.00   | 62.50  |
|          | female     | 126         | 60            |             | height (cm)| 163.80 ± 7.27| 164.43 ± 7.45| 163.00  | 163.00 |
| education| primary    | 40          | 26            |             | weight (kg)| 64.34 ± 9.23 | 65.09 ± 9.46 | 65.00   | 65.00  |
|          | middle and above | 174   | 80            | ADAS1       | age        | 14.83 ± 6.40 | 15.11 ± 6.12 | 13.85   | 14.85  |
| ethnicity| Han        | 206         | 105           |             | ADASFA     | 3.99 ± 4.11   | 4.13 ± 3.92 | 3.67    | 4.00   |
|          | Non-Han    | 8           | 1             |             |            |               |             |        |        |
| occupation| Physical | 55          | 31            |             |            |               |             |        |        |
|          | Mental     | 159         | 75            |             |            |               |             |        |        |
| drug     | without    | 155         | 71            |             |            |               |             |        |        |
|          | with       | 59          | 35            |             |            |               |             |        |        |
| centre   | centre1    | 23          | 10            |             |            |               |             |        |        |
|          | centre2    | 8           | 4             |             |            |               |             |        |        |
|          | centre3    | 32          | 16            |             |            |               |             |        |        |
|          | centre4    | 24          | 12            |             |            |               |             |        |        |
|          | centre5    | 24          | 12            |             |            |               |             |        |        |
|          | centre6    | 48          | 24            |             |            |               |             |        |        |
|          | centre7    | 24          | 12            |             |            |               |             |        |        |
|          | centre8    | 31          | 16            |             |            |               |             |        |        |
| Total    |            | 214         | 106           |             |            |               |             |        |        |

Table 2. An example of covariance analysis.

| Degrees of freedom | Sum of squares | Mean square | F-value | P-value |
|--------------------|----------------|-------------|---------|---------|
| group              | 1              | 3.43        | 3.43    | 0.31    | 0.58    |
| centre             | 7              | 375.73      | 53.68   | 4.80    | <0.001  |
| group*centre       | 7              | 104.41      | 14.92   | 1.33    | 0.234   |
| ADAS1              | 1              | 1439.67     | 1439.67 | 128.79  | <0.001  |
| error              | 303            | 3387.00     | 11.18   |         |         |
| Corrected total    | 319            | 5204.11     |         |         |         |
To attain high efficiency, we choose the tuning parameter \( \gamma \) by minimizing the determinant of asymptotic covariance matrix as in Step 2. Since the calculation of \( \det(\hat{V}(\gamma)) \) depends on the estimation of \( \hat{\beta}_n \), we update \( \hat{\beta}_n \) in Step 3 and repeat the algorithm until the convergence condition \( \| \hat{\beta}_n^{\text{old}} - \hat{\beta}_n^{\text{new}} \| < 10^{-2} \) is satisfied.

**Methods**

**Model.** Suppose there are \( n \) subjects, denoted as \( \{(x_{i},y_{i})\}_{i=1}^{n} \), where \( y_i \) is the outcome and \( x_i = (x_{i1}, \ldots , x_{ip})^T \) is a \( p \)-dimensional vector of covariates. A linear regression model is,

\[
 y_i = x_i^T \beta + \varepsilon_i, \quad i = 1, 2, \ldots , n
\]

where \( \beta \) is a \( p \)-dimensional vector of unknown parameters while \( \varepsilon_i \) is independent and identically distributed with some unknown distribution satisfying \( \mathbb{E}(\varepsilon_i) = 0 \) and \( \varepsilon_i \perp x_i \).

The ESL function has been used in AdaBoost for classification problems with success\(^{15}\). Wang et al\(^{14}\) expanded the use of the ESL function for robust variables selection. We now use it to estimate parameters in linear regression without sparsity. The ESL function is defined as

\[
 \Phi_i(t) = 1 - \exp\left(-\frac{t^2}{\gamma}\right),
\]

which is a function of \( t \) and \( \gamma \), where the latter is a tuning parameter. To estimate model parameters \( (\beta) \), the objective function of ESL is to maximize,

\[
 L_n(\beta) = \sum_{i=1}^{n} \exp\left(-\frac{(y_i - x_i^T \beta)^2}{\gamma}\right).
\]

The tuning parameter \( \gamma \) controls the degree of robustness of the estimator. With a relatively large \( \gamma \), the proposed estimator gets close to the OLS estimator while a smaller \( \gamma \) leads to a limited influence of contaminations on the estimator. Since the tuning parameter \( \gamma \) controls the degree of robustness and efficiency of the estimator, a data-driven procedure that yields both high robustness and high efficiency simultaneously is used to select an appropriate \( \gamma \). The entire calculation process in terms of ESL borrows from the idea proposed in Wang et al\(^{14}\):

1. **Find the pseudo outlier set of the sample.** Let \( D_n = \{(x_{i}, y_{i})\}_{i=1}^{n} \). Then, calculate \( r(\hat{\beta}_n) = y_i - x_i^T \hat{\beta}_n \), \( i = 1, 2, \ldots , n \) and \( S_n = 1.4826 \times \text{median}(r(\hat{\beta}_n) - \text{median}(r(\hat{\beta}_n))) \). Take the pseudo outlier set as \( D_{\text{new}} = \{(x_{i}, y_{i})\}_{i=1}^{n} : r(\hat{\beta}_n) \geq 2.5S_n \} \), where \( m \) is the cardinality of \( D_{\text{new}} \) set and \( D_{\text{new}} = D_n - D_{\text{new}} \).

2. **Update the tuning parameter \( \gamma \).** Let \( \gamma \) be the minimiser of \( \det(\hat{V}(\gamma)) \) in the set \( G = \{\gamma : \zeta(\gamma) \in (0, 1)\} \), and \( \zeta(\gamma) = 2m/n + (2/n)\sum_{i=m+1}^{n} \Phi_i(\hat{\beta}_n), \quad \hat{V}(\gamma) = \{I(\hat{\beta}_n)^{-1} \Sigma[I(\hat{\beta}_n)^{-1}]^{-1}\} \), \( \det(\cdot) \) denotes the determinant operator, and

\[
 \hat{I}(\hat{\beta}_n) = 2 \left( \frac{1}{\gamma} \sum_{i=1}^{n} \exp(-r_i^2(\hat{\beta}_n)/\gamma) \right) \left( \frac{2r_i^2(\hat{\beta}_n)}{\gamma} - 1 \right) \times \left( \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T \right)
\]

\[
 \Sigma = \text{cov} \left\{ \exp(-r_i^2(\hat{\beta}_n)/\gamma) \frac{2r_i^2(\hat{\beta}_n)}{\gamma} x_i \times \cdots \times \exp(-r_n^2(\hat{\beta}_n)/\gamma) \frac{2r_n^2(\hat{\beta}_n)}{\gamma} x_n \right\}
\]

3. **Update \( \hat{\beta}_n \).** After selecting \( \gamma \) in step 2, update \( \hat{\beta}_n \) by maximizing (1).

![QQ plot of residuals in the MCI study using OLS.](Image)
In the MCI study, the linear regression model is conducted as follows:

\[
\text{ADASFA} = \beta_{\text{age}} + \beta_{\text{BMI}} + \beta_{\text{ADAS1}} + \beta_{\text{centre2}} + \beta_{\text{centre3}} + \beta_{\text{ADAS2}} + \beta_{\text{ADAS3}} + \beta_{\text{ADAS4}} + \beta_{\text{ADAS5}} + \beta_{\text{ADAS6}} + \beta_{\text{centre7}} + \beta_{\text{group}} + \beta_{\text{gender}} + \beta_{\text{education}} + \beta_{\text{occupation}} + \beta_{\text{drug}} + \varepsilon
\]

Table 3. Average results over 100 replications of ESL and OLS for 10%, 20%, and 30% contamination proportions, respectively.
We include these variables in the model based on our clinical experience and existing literatures. In addition, we transform weight and height into a new variable BMI, since there are discussions on whether BMI has an effect on MCI. We exclude ethnicity and marital status from our model mainly because these two variables are extremely unbalanced between the treatment and control arms, and also due to the fact that almost no literature suggests that these two variables have an effect on MCI. Since 11 out of 16 variables are categorical variables, we do not consider interaction effects. The by-centre descriptive analysis is presented in Table 4 and 5.

Table 6 shows the parameter estimations using ESL and OLS. The empirical 95% confidence interval is calculated by the bootstrap approach. When the bootstrap confidence interval does not include 0, it indicates that the corresponding covariate has a significant effect on the primary outcome. Note that there are some differences between the ESL and OLS estimations. For example, the effects of centre5 and centre8 on ADASFA are opposite. Given the non-normal residuals, the ESL estimators are more accurate. From the results, we can conclude that

(1) ADAS1 and centre 6 have significant influences on ADASFA since their bootstrap confidence intervals do not contain 0. From the medical view, higher ADAS1 means patients are in worse health situation, which can have a positive effect on ADASFA.

(2) ESL and OLS both show that ADAS1 has a positive effect on decreasing ADAS-cog. For age, ESL shows that age has no effect on decreasing ADAS-cog because its bootstrap confidence interval contains 0 while OLS shows that age has a negative effect on decreasing ADAS-cog. From a medical viewpoint, it is verified that 'age' has a significant effect on MCI. Prior work has demonstrated that rates of dementia increase exponentially with age. However, the significant effect of age on MCI does not mean that it also influences the treatment effect.

(3) The ESL group coefficient is $-0.141$ and its bootstrap confidence interval contains 0. This result makes sense because this project is a non-inferiority trial and the treatment group was not worse than the control group.

(4) The ESL shows that centres 3, 6, and 7 have significant effects on the outcome. However, OLS shows that centres 3 and 7 have no significant impact but centre 6 has a significant effect on the outcome. According to Table 4, the average ADAS1 of centre 6 is much lower than that of centres 3 and 7, which implies...
### Table 4. Continuous variable descriptive statistics.

| Centre | Group       | Treatment | Control |
|--------|-------------|-----------|---------|
|        |             | Mean     | Sd      | Mean     | Sd        | Median | Median |
| 1      | age         | 56.39    | 7.48    | 58.90    | 8.37      | 53.00   | 58.00  |
|        | bmi         | 23.57    | 1.67    | 23.44    | 1.72      | 23.92   | 23.13  |
|        | ADAS1       | 12.36    | 4.99    | 12.03    | 6.91      | 13.00   | 12.70  |
|        | ADASCHA     | 2.60     | 2.00    | 2.18     | 2.11      | 2.40    | 2.00   |
| 2      | age         | 66.00    | 7.60    | 66.75    | 8.30      | 67.50   | 67.50  |
|        | bmi         | 23.82    | 1.33    | 20.71    | 2.47      | 23.84   | 22.70  |
|        | ADAS1       | 8.08     | 3.78    | 8.25     | 3.52      | 7.30    | 8.00   |
|        | ADASCHA     | 2.63     | 3.19    | 4.68     | 3.30      | 5.55    | 5.00   |
| 3      | age         | 64.25    | 6.41    | 62.63    | 7.39      | 64.00   | 61.00  |
|        | bmi         | 24.00    | 2.60    | 24.29    | 2.47      | 24.24   | 23.94  |
|        | ADAS1       | 17.70    | 5.57    | 18.53    | 5.63      | 17.00   | 19.15  |
|        | ADASCHA     | 5.68     | 4.45    | 4.86     | 4.16      | 4.87    | 4.55   |
| 4      | age         | 64.88    | 7.92    | 64.50    | 9.02      | 66.50   | 63.50  |
|        | bmi         | 24.68    | 2.53    | 24.85    | 2.66      | 24.97   | 24.37  |
|        | ADAS1       | 20.71    | 6.49    | 21.11    | 5.76      | 19.65   | 20.85  |
|        | ADASCHA     | 4.22     | 3.31    | 6.55     | 6.12      | 4.14    | 5.55   |
| 5      | age         | 65.71    | 9.07    | 60.75    | 6.77      | 66.00   | 59.50  |
|        | bmi         | 24.16    | 2.50    | 25.64    | 2.62      | 24.49   | 26.02  |
|        | ADAS1       | 18.39    | 5.52    | 17.91    | 5.13      | 19.33   | 19.35  |
|        | ADASCHA     | 4.53     | 3.63    | 4.91     | 1.43      | 4.37    | 5.05   |
| 6      | age         | 64.63    | 7.55    | 67.83    | 6.55      | 60.00   | 69.00  |
|        | bmi         | 24.55    | 3.27    | 23.84    | 3.09      | 24.57   | 23.80  |
|        | ADAS1       | 10.55    | 4.75    | 10.87    | 3.37      | 9.40    | 10.30  |
|        | ADASCHA     | 3.94     | 4.46    | 4.59     | 3.33      | 3.64    | 4.00   |
| 7      | age         | 58.17    | 5.95    | 57.42    | 4.87      | 57.50   | 56.00  |
|        | bmi         | 22.69    | 1.99    | 22.96    | 2.80      | 22.99   | 22.23  |
|        | ADAS1       | 15.69    | 5.84    | 14.66    | 3.60      | 14.85   | 14.95  |
|        | ADASCHA     | 3.21     | 6.69    | 2.78     | 5.06      | 4.70    | 3.80   |
| 8      | age         | 61.81    | 6.96    | 67.63    | 9.54      | 60.00   | 70.50  |
|        | bmi         | 23.45    | 3.48    | 23.77    | 3.59      | 22.66   | 23.14  |
|        | ADAS1       | 14.09    | 3.07    | 15.41    | 5.77      | 15.30   | 14.55  |
|        | ADASCHA     | 3.67     | 3.49    | 2.41     | 2.70      | 2.60    | 1.45   |

### Table 5. Discrete variable descriptive statistics.

| Centre | Group       | Gender | Education | Ethnicity | Occupation | Drug |
|--------|-------------|--------|-----------|-----------|------------|------|
|        |             | Male   | Female    | Primary   | Middle and above | Han | Non-han | Physical | Mental | Without | With |
| 1      | Treatment   | 8      | 15        | 4         | 19         | 23   | 0       | 7        | 16     | 22      | 1    |
|        | Control     | 3      | 7         | 3         | 7          | 10   | 0       | 3        | 7      | 9       | 1    |
| 2      | Treatment   | 2      | 6         | 1         | 7          | 8    | 0       | 1        | 7      | 4       | 4    |
|        | Control     | 4      | 9         | 1         | 3          | 4    | 0       | 0        | 4      | 0       | 4    |
| 3      | Treatment   | 11     | 21        | 2         | 30         | 32   | 0       | 3        | 29     | 22      | 10   |
|        | Control     | 7      | 9         | 3         | 13         | 16   | 0       | 2        | 14     | 12      | 4    |
| 4      | Treatment   | 15     | 9         | 6         | 18         | 23   | 1       | 9        | 15     | 18      | 6    |
|        | Control     | 6      | 6         | 4         | 8          | 12   | 0       | 5        | 7      | 9       | 3    |
| 5      | Treatment   | 10     | 14        | 3         | 21         | 23   | 1       | 7        | 17     | 14      | 10   |
|        | Control     | 3      | 9         | 2         | 10         | 12   | 0       | 3        | 9      | 9       | 3    |
| 6      | Treatment   | 22     | 26        | 9         | 39         | 43   | 5       | 16       | 32     | 30      | 18   |
|        | Control     | 13     | 11        | 10        | 14         | 23   | 1       | 10       | 14     | 13      | 11   |
| 7      | Treatment   | 8      | 16        | 5         | 19         | 23   | 1       | 3        | 21     | 22      | 2    |
|        | Control     | 4      | 8         | 0         | 12         | 12   | 0       | 3        | 9      | 10      | 2    |
| 8      | Treatment   | 12     | 19        | 10        | 21         | 31   | 0       | 9        | 22     | 23      | 8    |
|        | Control     | 6      | 10        | 3         | 13         | 16   | 0       | 5        | 11     | 9       | 7    |
that patients in centres 3 and 7 are in worse conditions. Moreover, patients in different centres may have different non-compliance levels, which may also contribute to the result that some centres have significant effects on the outcome while others do not.

(5) Since we have shown that the data is not normally distributed, we can have greater confidence in the ESL results.

**Conclusion**

In this paper, we discuss a method to evaluate efficacy in a randomized control MCI study. As many covariates may influence the outcome, a linear regression model is considered rather than comparing group means using t test or ANOVA. An exponential squared loss function, which is superior to OLS when dealing with non-normal residuals, is introduced in this study. Simulation results show that the ESL model yields more efficient estimation than OLS in non-normal data. The proposed method is also robust in the case of data with outliers. These advantages of the ESL model become more noticeable when the contamination percentage increases. The proposed method does not require the normal distribution assumption, offering new insight in the efficacy evaluation for practical researchers.

**References**

1. Gauthier, S. et al. Mild cognitive impairment. *The Lancet* 367, 1262–1270 (2006).
2. Rosen, W. G., Mohs, R. C. & Davis, K. L. A new rating scale for Alzheimer’s disease. *The Am. J. Psychiatry* (1984).
3. Doraiswamy, P. M., Kaiser, L., Bieber, F. & Garman, R. L. The Alzheimer’s Disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer’s disease. *Alzheimer Dis. & Assoc. Disord.* 15, 174–183 (2001).

**Table 6.** Estimation results in MCI study using ESL and OLS.

| Variables                  | ESL Estimate (95% CI) | OLS Estimate (95% CI) |
|----------------------------|-----------------------|-----------------------|
| Intercept                  | 0.241 (−4.412, 4.203) | −0.127 (−4.961, 4.554) |
| Age                        | −0.05 (−0.098, 0.01)  | −0.055 (−0.108, −0.003) |
| BMI                        | 0.072 (−0.088, 0.22)  | 0.023 (−0.112, 0.158)  |
| ADAS1                      | 0.31 (0.221, 0.398)   | 0.434 (0.361, 0.508)   |
| Center2                    | 1.82 (−0.331, 5.253)  | 3.011 (0.683, 5.336)   |
| Center3                    | 1.428 (0.127, 2.683)  | 0.725 (−0.848, 2.299)  |
| Center4                    | −0.524 (−2.192, 1.106)| −0.988 (−2.711, 0.733) |
| Center5                    | 0.553 (−0.604, 1.571) | −0.151 (−1.828, 1.526) |
| Center6                    | 2.396 (1.457, 3.308)  | 2.733 (1.263, 4.201)   |
| Center7                    | 1.474 (0.451, 2.514)  | −0.71 (−2.303, 0.889)  |
| Center8                    | −0.001 (−1.284, 1.407)| 0.158 (−1.372, 1.688)  |
| Group (treatment = 1)      | −0.141 (−0.8, 0.567)  | −0.096 (−1.553, −0.022) |
| Gender (female = 1)        | −0.565 (−1.19, 0.275) | −0.786 (−1.118, 1.947) |
| Education (middle and above = 1) | 0.383 (−0.441, 1.285) | 0.921 (−1.275, 0.589)  |
| Occupation (mental = 1)    | −0.245 (−0.991, 0.456)| −0.341 (−0.812, 0.892) |
| Drug (with = 1)            | −0.08 (−1.197, 0.912) | 0.035 (−1.197, 0.912)  |

**Figure 4.** Error bars of ESL and OLS with 30% contamination.
4. Petersen, R. C. et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. *New Engl. J. Medicine* **352**, 2379–2388 (2005).
5. Morgan, K. L. & Rubin, D. B. Rerandomization to improve covariate balance in experiments. *The Annals Stat.* **40**, 1263–1282 (2012).
6. Hong Yan, N. Z. & Yongyong X Medical statistics. *People's Med. Publ* (2010).
7. Jianan Fang, Y. L. Advanced medical statistics. *People's Med. Publ. House* (2002).
8. Qin, Y. & Priebe, C. E. Maximum Lq-likelihood estimation via the expectation-maximization algorithm: A robust estimation of mixture models. *J. Am. Stat. Assoc.* **108**, 914–928 (2013).
9. Di Zhang et al. Clinical study of insulin resistance for patients after selective operation in department of general surgery. *Chin. J. Bases Clin. Gen. Surg.* **9**, 951–954 (2010).
10. Yang Dai, X. Y. J. X. T. Z. Y. S. C. L. L. W. & Lijun, W. To study the effect and safety of lhxcxiaofeng drug detoxification. *Chin. J. Clin. Res.* **8**, 1087–1089 (2015).
11. Hekimoglu, S., Erenoglu, R. C. & Kalina, J. J. Outlier detection by means of robust regression estimators for use in engineering science. *J. Zhejiang Univ. Sci.* **10**, 909–921 (2009).
12. Bao, Y. Research on robust regression analysis and diagnosis method based on rank order. *Shanxi Med. Univ.* **10**, 10 (2003).
13. Wang, H., Li, G. & Jiang, G. Robust regression shrinkage and consistent variable selection through the LAD-lasso. *J. Bus. Econ. Stat.* **25**, 347–355 (2007).
14. Wang, X., Jiang, Y., Huang, M. & Zhang, H. Robust variable selection with exponential squared loss. *J. Am. Stat. Assoc.* **108**, 632–643 (2013).
15. Friedman, J., Hastie, T. & Tibshirani, R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The Annals Stat.* **28**, 337–407 (2000).
16. Yohai, V. J. High breakdown-point and high efficiency robust estimates for regression. *The Annals Stat.* **15**, 642–656 (1987).
17. Emre, M., Ford, P. J., Bilgic, B. & Uc, E. Y. Cognitive impairment and dementia in Parkinson's disease: practical issues and management. *Mov. disorders* **29**, 663–672 (2014).
18. Hall, C. B. et al. Dementia incidence may increase more slowly after age 90 results from the Bronx Aging Study. *Neurol.* **65**, 882–886 (2005).
19. Alzheimer's Association et al. 2009 Alzheimer's disease facts and figures. *Alzheimer’s & Dementia* **5**, 234–270 (2009).

**Acknowledgements**

We thank Wenjing Kong for her suggestions on the writing of this manuscript. This study is partially supported by the Ministry of Science and Technology of the PRC: Significant New Drug Development Construction of Technology Platform used for Original New Drug Research and Development (2012ZX09303-010-002), the Ministry of Education Project of Key Research Institute of Humanities and Social Sciences at Universities (16JJD910002), and the National Natural Science Foundation of China (81774146).

**Author Contributions**

Yang Li, Fang Lu, and Danhui Yi were responsible for the concept and study design; Yang Li, Zhang Zhang, and Qian Feng wrote the main manuscript text; and Zhang Zhang and Qian Feng prepared all the figures and tables. All authors reviewed the manuscript.

**Additional Information**

**Competing Interests:** The authors declare no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019