

Quality of reporting of multivariable logistic regression models in Chinese clinical medical journals

Ying-ying Zhang, MD, MPH^a, Xiao-bin Zhou, MD^{a,*}, Qiu-zhen Wang, PhD^b, Xiao-yan Zhu, MD, MPH^a

Abstract

Multivariable logistic regression (MLR) has been increasingly used in Chinese clinical medical research during the past few years. However, few evaluations of the quality of the reporting strategies in these studies are available.

To evaluate the reporting quality and model accuracy of MLR used in published work, and related advice for authors, readers, reviewers, and editors.

A total of 316 articles published in 5 leading Chinese clinical medical journals with high impact factor from January 2010 to July 2015 were selected for evaluation. Articles were evaluated according 12 established criteria for proper use and reporting of MLR models.

Among the articles, the highest quality score was 9, the lowest 1, and the median 5 (4–5). A total of 85.1% of the articles scored below 6. No significant differences were found among these journals with respect to quality score ($\chi^2=6.706$, $P=.15$). More than 50% of the articles met the following 5 criteria: complete identification of the statistical software application that was used (97.2%), calculation of the odds ratio and its confidence interval (86.4%), description of sufficient events (>10) per variable, selection of variables, and fitting procedure (78.2%, 69.3%, and 58.5%, respectively). Less than 35% of the articles reported the coding of variables (18.7%). The remaining 5 criteria were not satisfied by a sufficient number of articles: goodness-of-fit (10.1%), interactions (3.8%), checking for outliers (3.2%), collinearity (1.9%), and participation of statisticians and epidemiologists (0.3%). The criterion of conformity with linear gradients was applicable to 186 articles; however, only 7 (3.8%) mentioned or tested it.

The reporting quality and model accuracy of MLR in selected articles were not satisfactory. In fact, severe deficiencies were noted. Only 1 article scored 9. We recommend authors, readers, reviewers, and editors to consider MLR models more carefully and cooperate more closely with statisticians and epidemiologists. Journals should develop statistical reporting guidelines concerning MLR.

Abbreviations: CI = confidence interval, MLR = multivariable logistic regression.

Keywords: Chinese clinical medical journal, multivariable logistic regression, quality evaluation

1. Introduction

The most widely used approaches to multivariable models in clinical studies are multivariable linear regression, multivariable logistic regression (MLR), and proportional hazards regression^[1]. The independent variables in MLR model can be continuous, categorical, or ordinal. Moreover, their coefficients

can be easily converted into odds ratios (ORs) with straightforward explanation. Normal distribution is not required in logistic regression. Due to these advantages, MLR is widely used in medical research mostly for adjustment of confounding factors, screening of relevant variables, predicting, and discrimination.^[2] However, failure to perform multivariable regression appropriately, which is manifested, for instance, violating or neglecting assumptions and preconditions, or ambiguous coding of variables, can potentially lead to inaccurate, misleading, or even erroneous conclusions; or render the conclusions difficult to interpret.^[3] Appropriate assumptions, correct interpretation, and complete reporting of MLR have been studied since 1993.^[4] The study of Kumar and Chhabra^[5] indicated that understanding of the inherent assumptions and important limitations of the model before application can significantly enhance the quality and reliability of MLR analysis. However, several deficiencies in reporting quality can be found in many fields, such as obstetrics and gynecology,^[2] pulmonary and critical care,^[6] and clinical epidemiology.^[7,8]

As early as 1993, Concato et al^[3] pointed out the reporting quality problems of multivariate statistical analysis in medical research, found that 6 important assumptions in logistic regression analysis are ignored or unreported, and recommended improving the reporting and application guidelines for multivariate analysis in medical research. In the following 20 years, a large number of such studies emerged.^[7–12] However, the reporting quality of MLR in the literature remained uneven. Recently, there has been significant

Editor: Vijayaprasad Gopichandran.

Supplemental digital content: All collected data and obtained scores are available upon request, <http://links.lww.com/MD/B707>

The authors have no funding and conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

^a Department of Epidemiology and Health Statistics, ^b Department of Public Health, Medical College of Qingdao University, Qingdao, China.

* Correspondence: Xiao-bin Zhou, Department of Epidemiology and Health Statistics, Medical College of Qingdao University, No. 38, Dengzhou Road, Qingdao 266021, China (e-mail: xiaobin_zhou@126.com)

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution-NoDerivatives License 4.0, which allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to the author.

Medicine (2017) 96:21(e6972)

Received: 15 November 2016 / Received in final form: 27 April 2017 / Accepted: 27 April 2017

<http://dx.doi.org/10.1097/MD.0000000000006972>

improvement; however, there are insufficiencies in, for instance, the collinearity test, checking for outliers, reporting guidelines or standards, and participation of statisticians.^[5,9,12] Scholars in China have only focused on theoretical research. Sun^[13] discussed the significant influence of sample size and screened methods on model parameters and paired data variables, respectively; Liu et al.^[14] developed a generalized ad-logistic regression theory; Liu^[15] found several common problems in logistic regression analysis and proposed methods for performing collinearity tests and identifying outliers in the sample.

So far, research on reporting quality evaluation of MLR is rare worldwide, and has not been available in Chinese literature. Thus, we evaluated studies using MLR published in 5 leading Chinese clinical medical journals from 2010 to 2015 according to 12 established criteria, hoping to provide reference for reasonable application of MLR and correct reporting of results for authors, reviewers, and journal editors.

2. Methods

2.1. Data collection

The following journals were selected due to their high Chinese Science Citation Database (CSCD) impact factor (IF) and their large circulations: Chinese Journal of Cardiology (IF=0.5065 of year 2012), Chinese Journal of Oncology (IF=0.4474), Chinese Journal of Neuromedicine (IF=0.6729), Chinese Journal of Pediatrics (IF=0.7669), and Chinese Journal of Digestive Surgery (IF=0.7318). We performed a manual search of all articles published between 2010 and 2015. The 6-year interval was used to indicate any trends over time. A total of 316 articles that contained the word “logistic” in the title, abstract, or keywords list published in Chinese Journal of Cardiology (119 articles), Chinese Journal of Oncology (57 articles), Chinese Journal of Neuromedicine (77 articles), Chinese Journal of Pediatrics (31 articles), and Chinese Journal of Digestive Surgery (32 articles) were included in the research.

2.2. Assessment criteria of reporting quality

In line with related research,^[5,9,10,12] we applied 12 criteria to evaluate the quality of reporting of MLR, namely: (1) selection of independent variables; (2) fitting procedure; (3) coding of variables; (4) interactions; (5) collinearity; (6) statistical significance (OR, 95% confidence interval [CI], *P* value); (7) goodness-of-fit; (8) checking for outliers; (9) complete identification of the statistical software application that was used; (10) sufficient events (>10) per variable; (11) participation of statisticians and epidemiologists; and (12) conformity with linear gradient. The articles were evaluated according to the first 11 criteria. Each criterion contributed one point to the score if it was fulfilled. Criterion (12) was not included in the calculation of the total score, since continuous or rank variable were not used in every article. Ethical approval was not involved, as the subject of our study is literature, not human being or animal.

2.2.1. Selection of independent variables. Variable selection can seriously affect the model estimation. Therefore, it should be justified. Usually, variables are selected according to professional knowledge and previous studies, or statistically significant association in a univariate analysis.

2.2.2. Fitting procedure. The variables may be determined by model selection methods, such as automatic procedures (eg,

forward inclusion, backward elimination, stepwise selection, or best subset selection), nonautomated backward selection, or a priori specification (either collectively or in “hierarchically” grouped subsets).^[8] Preferably a description of the model test methods (eg, conditional parameter estimation, maximum partial likelihood estimation, and the Wald chi-square test) should be included. An article was considered to fulfill this criterion if it reported one of the methods above.

2.2.3. Coding of variables. Reporting the coding of variables properly is important, since the interpretation of regression coefficients depends on the coding of variables and the measurement of units. For example, the coefficient for the impact of age on mortality will be very different if age is coded in 1-year increments, in 10-year increments, or dichotomously as <65 versus ≥65 years.^[3] Hence, reporting the coding of variables plays an important role in correctly understanding the model parameters (OR). We evaluated an article as satisfying this criterion if a detailed classification of the independent variables or the reference group was presented. An explicit list is, of course, preferable.

2.2.4. Interactions. When the effect of an independent variable on the outcome variables can be affected by other variables, we have interactions among independent variables. This can conceal the true correlation between independent and dependent variables.^[16] Generally, their statistical significance and effect on the model must be tested and reported, according to professional knowledge or previous studies. Articles including explicit tests for interaction, mentioning the concept of interaction anywhere in the text, or justifying the exclusion of interaction from the final model were regarded as fulfilling the criterion.

2.2.5. Collinearity. Collinearity is high correlation between 2 or more covariates. Multicollinearity would occur if some covariates are partially or totally explained by other covariates. Collinearity is necessary to be checked before establishing logistic models, otherwise unreliable estimates of coefficients and wide CIs may appear. Methods for tackle the multicollinearity can be found in the textbook by Allison.^[17] However, collinearity is often ignored. The criterion was considered to be fulfilled as long as the concept of collinearity was discussed anywhere in article.

2.2.6. Statistical significance. Articles that reported the OR and CI correctly, preferably with *P* values, met this criterion.^[18] The use of *P* values only should be discouraged.

2.2.7. Goodness-of-fit. Goodness-of-fit can provide information about how well the entire model matches the observed data. Despite controversy, several methods are available for goodness-of-fit testing including Pearson test, deviance, likelihood ratio, and Hosmer–Lemeshow statistic (equivalents of R^2 in linear regression).^[19] Moreover, there are strategies for assessing the model’s predictive performance based on Pseudo R^2 , the fraction of correct predictions, or receiver-operating characteristic (ROC) curves. These 2 procedures (methods for goodness of fit and predictive performance) are different concepts, and both should be reported in a model. Nevertheless, article was classified as meeting this criterion if any of the measures above was calculated.

2.2.8. Checking for outliers. Outliers are variables whose residuals (observed-predicted) are significantly greater than those of other variables. If undetected, they may distort the model’s robustness due to their effect on the coefficients of the variables.

Obviously, outlier-checking procedures are essential to ensure the accuracy of MLR analysis. Useful techniques are available for detecting outliers, such as residual examination (eg, Pearson residuals and deviance residuals), and for measuring the impact of outliers on the regression model, such as the Cook distance and DFBETA (s).^[5,20] Specific SAS procedures can be found in the literature. We classified the articles based on whether any method for outliers-checking was mentioned.

2.2.9. Complete identification of the statistical software application that was used. Identifying the software application can help other researchers to reproduce and test the MLR model. It is necessary to report the version of the program due to the continuous revisions of the same application.^[5]

2.2.10. Sufficient events (>10) per variable (ratio of outcome events to independent variables). Generally, including more relevant variables can result in a model that fits the data better.^[8] However, large standard error, unreliable coefficients estimates, and wider CIs may appear when an excessively large number of variables are used for small number of outcome events. This is called over-fitting.^[5] Despite controversy about the number of outcome events required for MLR models, the ratio of at least 10 outcome events per independent variable is widely approved.^[8,21] A small ratio implies greater potential of biased estimation and invalidity of MLR.^[21] We categorized the articles with a ratio greater than 10:1 (or greater than 20 for conditional MLR) as fulfilling the criterion.

2.2.11. Participation of statisticians and epidemiologists. Proper use of statistical methods combined with professional knowledge can avoid bias and reduce defects during statistical analysis and reporting procedures. Similar studies have indicated that inadequate reporting was less frequent if an author was affiliated with a Department of Statistics, Epidemiology, or Public Health.^[22] Therefore, the participation of statisticians and epidemiologists is important for properly using and appropriately reporting the MLR model.

2.2.12. Conformity with linear gradient for continuous or rank variables. In the case of continuous or rank independent variables, linear relationship of log-odds may be imposed upon the model. Any specific unit change in the respective covariate should have the same impact on the outcome,^[3] otherwise, there is a risk of misspecification and false inferences,^[9] especially when the relationship is U-shaped, J-shaped, or parabolic.^[5] To avoid this problem, the most common strategy is to convert a continuous variable into a categorical variable. This is likely to result in loss of useful information. Therefore, more flexible modeling procedures are required for handling continuous variables, such as spline regression, multivariable fractional polynomials, and generalized additive models. If there was any description of the linear assumption, preferably, some type of justification (eg, professional modeling approaches, exploratory plot of the data, or prior clinical knowledge) should also be provided, the article was considered to meet the criterion.

2.3. Statistical analysis

The criteria were evaluated by 2 authors (YZ and XZ) and possible disagreement was discussed. The frequencies of fulfilling the criteria were separately calculated for each journal, and the Pearson χ^2 test and Fisher exact test were used for testing the differences among the journals. The quality scores were

Table 1

Distribution of quality scores of the 316 articles using multivariable logistic regression (MLR).

Scores for each article	Number	Percentage, %
≤5	269	85.1
6	39	12.4
7	7	2.2
8	0	0
9	1	0.3
Total	316	100.0

represented by median (interquartile range), and their differences among the journals were tested by the multiple independent samples Kruskal–Walis H test. All calculations were performed in IBM SPSS Statistics 19.0, and P value <.05 was considered significant.

3. Results

3.1. Basic characteristics of the quality of the selected articles

The highest score was 9, the lowest was 1, the median (interquartile range) was 4 (4–5). A total of 85.1% of the articles scored less than 6, and 14.9% scored between 6 and 9 points, as seen in Table 1. The distribution of quality scores among journals is summarized in Table 2. The number of articles using MLR and the quality score had no apparent change over time (Fig. 1). Through the Kruskal–Walis H test, the difference of quality scores among 5 journals was no statistically significant ($\chi^2=6.706$, $P=.15$), whereas there is a significant difference over the 6-year period ($\chi^2=26.388$, $P<.05$).

3.2. The quality of reporting of articles using MLR

The results for the fractions of fulfillment of the 11 criteria are summarized in Table 3. The criterion of conformity with linear gradients for continuous or rank variables was applicable to 186 articles; however, only 7 (3.8%) mentioned or tested this conformity.

3.2.1. Selection of independent variables. A total of 219 articles (69.3%) described the selection of independent variables. Specifically, 56 were based on previous studies and experience, and 128 were based on statistical significance of the single factor test. Twenty-one articles used both methods, and 14 full-model.

Table 2

Distribution of articles in 5 journals and their descriptive statistics for quality score.

Journal	Number of MLR	Quality score		χ^2	P
		Median (interquartile range)	Range (min–max)		
CJC	119	4 (4–5)	7 (2–9)	6.706	.15
CJO	57	4 (3–5)	5 (2–7)		
CJN	77	4 (3–5)	6 (1–7)		
CJP	31	4 (3–5)	5 (1–6)		
CJDS	32	5 (4–5)	4 (2–6)		
Total	316	4 (4–5)	8 (1–9)		

CJC=Chinese Journal of Cardiology, CJDS=Chinese Journal of Digestive Surgery, CJN=Chinese Journal of Neuromedicine, CJO=Chinese Journal of Oncology, CJP=Chinese Journal of Pediatrics, MLR=multivariable logistic regression.

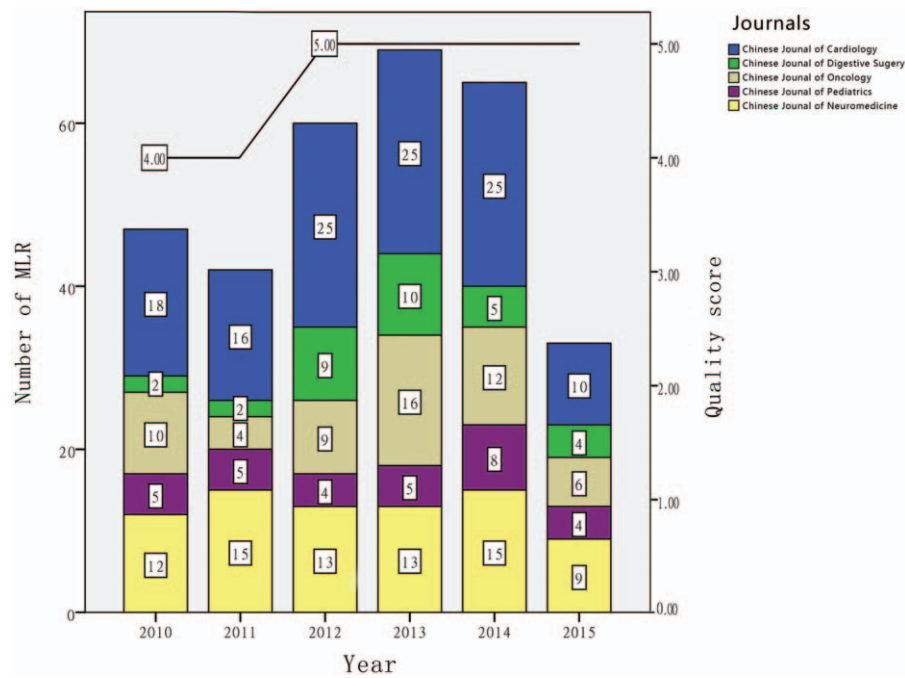


Figure 1. Number of articles using multivariable logistic regression and their quality score. The bar chart represents the number of articles using multivariable logistic regression (MLR). The line chart represents the quality score of articles using MLR. The difference among the 5 journals had no statistical significance ($\chi^2=6.706, P=.15$). The difference was significant over the 6-year period ($\chi^2=26.388, P<.05$).

3.2.2. Fitting procedure. In 185 articles (58.5%), the fitting procedure was reported. Only 22 articles reported both the selection and test methods.

3.2.3. Coding of variables. Classification or coding of independent variables was described completely in 59 articles (18.7%).

3.2.4. Interactions. Twelve articles (3.8%) met the criterion. The product term was included as an independent variable in all models in order to isolate the interaction and examine its impact.

3.2.5. Collinearity. The criterion of collinearity among independent variables was met in 6 articles (1.9%). BKW^[23] was used to assess the degree of collinearity.

3.2.6. Statistical significance (OR, 95% CI, P value). A total of 273 articles (86.4%) met this criterion. Specifically, 261 reported all indices for each variable’s coefficients in the final model, and 12 reported the OR and 95% CI. Among the remaining articles, 32 reported either the OR or the 95% CI, and 11 reported only the P value.

3.2.7. Goodness-of-fit. Thirty-two articles (10.1%) evaluated the model. Specifically, 5 conducted goodness-of-fit (Hosmer–Lemeshow was used in 3 articles and likelihood ratio in 2); 23 evaluated the model’s predictive performance based on the fraction of correct predictions (4 articles), or receiver operating characteristic curves (19 articles); 4 used the C++ programming

Table 3

Quality of reporting of multivariable logistic regression criteria (n, %).

Criteria	Number of articles that met the criteria					Total (n=316)	95% CI	χ^2	P
	CJC (n=119)	CJO (n=57)	CJN (n=77)	CJP (n=31)	CJDS (n=32)				
Selection of independent variables	73 (61.3)	43 (75.4)	53 (68.8)	20 (64.5)	30 (93.8)	219 (69.3)	69.30–74.42	13.884	<.01
Fitting procedure	69 (58.0)	32 (56.1)	44 (57.1)	18 (58.1)	22 (68.8)	185 (58.5)	53.08–64.01	1.509	.81
Classification or coding of independent variables	26 (21.8)	1 (1.8)	17 (22.1)	8 (25.8)	7 (21.9)	59 (18.7)	14.35–22.99	13.378	.01
Test for interactions between independent variables	5 (4.2)	4 (7.0)	2 (2.6)	1 (3.2)	0 (0.0)	12 (3.8)	1.68–5.92		.56*
Collinearity among independent variables	4 (3.4)	1 (1.8)	0 (0.0)	0 (0.0)	1 (3.1)	6 (1.9)	0.39–3.41		.42*
Statistical significance	112 (94.1)	48 (84.2)	64 (83.1)	21 (67.7)	28 (87.5)	273 (86.4)	82.59–90.19		<.01*
Goodness-of-fit	18 (15.1)	2 (3.5)	9 (11.7)	2 (6.5)	1 (3.1)	32 (10.1)	6.78–13.47		.08*
Checking of outliers	5 (4.2)	2 (3.5)	3 (3.9)	0 (0.0)	0 (0.0)	10 (3.2)	1.22–5.11		.84*
Name of statistical software with its version	115 (96.6)	56 (98.2)	76 (98.7)	28 (90.3)	32 (100.0)	307 (97.2)	95.31–99.00		.19*
Sufficient events (>10) per variable	101 (84.9)	47 (82.5)	54 (70.1)	26 (83.9)	19 (59.4)	247 (78.2)	73.58–82.74	13.877	<.01
The participation of statisticians and epidemiologists	1 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.3)	–0.31–0.94		1.00*

CI = confidence interval, CJC = Chinese Journal of Cardiology, CJDS = Chinese Journal of Digestive Surgery, CJN = Chinese Journal of Neuromedicine, CJO = Chinese Journal of Oncology, CJP = Chinese Journal of Pediatrics, MLR = multivariable logistic regression.

* Refers to Fisher exact probability.

language to predict model's performance. It seems that 68% of the articles may be over-fitted.

3.2.8. Checking for outliers. Ten articles (3.2%) met the criterion of checking for outliers.

3.2.9. Identification of the statistical software application. The majority, that is, 307 articles (97.2%) specified the application and its version. The most frequently used program is SPSS (88.3%), followed by SAS (7.9%).

3.2.10. Sufficient events (>10) per variable. In 247 articles (78.2%), the ratio of outcome events to independent variables was 10:1 or higher. The entire range was from 1.50 to 2459.75. There were 6 articles using conditional logistic regression with an optimal ratio over 20, namely, between 25.00 and 183.43.

3.2.11. Participation of statisticians and epidemiologists. Only 1 article (0.3%) reported participation of statisticians and epidemiologists.

4. Discussion

It was found that 5 criteria were fulfilled in more than 50% of the articles. This result is not particularly encouraging. For example, classification and coding of variables were described in only 18.7% of the articles. By contrast, Mikolajczyk et al^[9] found that 86 of the 104 articles (83%) that were published in obstetrics and gynecology journals in 2005 and 2006 used coding of potential covariates. The percentage was 41.28% in the study conducted by Kumar et al,^[12] where 109 articles using MLR were selected from 8 Indian journals from 1994 to 2008. Interactions between independent variables, conformity with linear gradient, and collinearity were reported in 3.8%, 3.8%, and 1.9%, respectively, in this study. Even though these are similar to the corresponding percentages of the study of Kumar (3.67%, 1.83%, and 0, respectively),^[12] they are far below those of the study conducted by Ottenbacher et al^[7] (39%, 17%, and 19%) in 2 epidemiology journals from 2000 to 2001, and those in the study of Kalil et al^[10] (19%, 4.7%, and 25%) from 6 major journals on organ transplant from January, 2005 to January, 2006. The small number of reportings for these assumptions may be due to the lack of statistical knowledge, unskillful software application, or absence of guidelines for appropriate reporting.

In this study, the encouraging finding was that the statistical software application and statistical significance, which are the most basic requirements of statistical analysis, were provided in a large number of articles. However, this was not the case with relatively complex tests. We found that 88.3% of the articles used SPSS for statistical analysis. However, there was no output information on multicollinearity in SPSS. It is likely that researchers ignore the multicollinearity problem when SPSS is used to construct MLR models. One solution is to use the same dependent and independent variables to conduct multiple-linear regression and the corresponding collinearity diagnosis. Yang^[24] and Zhao et al^[25] have given the implementation process for the diagnosis of multicollinearity of logistic regression by Stata and SAS, respectively. This can partly resolve the problem. However, in this study, 98.1% of the articles did not consider the potential of multicollinearity, which may result unstable regression coefficients or large CIs, and even affect the selection of variables. Currently, the most frequently used methods, such as principal component analysis, partial least squares estimation, and ridge regression, can be employed more efficiently to overcome the multicollinearity issue; however, there may be defects. More

adaptable methods are required for analyzing collinearity completely.^[26]

Only 10 articles (3.2%) detected outliers; however, no article provides a detailed analysis. Outliers, that is, observations beyond what is expected, may be identified by statistical variations, face validity, or consensus based on clinical reasons.^[27] MLR models are very sensitive to outliers, as outliers may cause or cover multicollinearity between independent variables and affect the model's robustness and parametric estimation.^[5] Therefore, outliers should be treated with caution when they appears: first, remove outliers caused by data collection or recording error. Subsequently, important covariates, interactions, sufficient sample size, and other issues should be considered when outliers are corrected.^[20] Finally, conduct 2 complete independent studies. In 1 study, retain the outliers. In the other, remove them. The conclusions from both studies are then compared.^[10]

The linear gradient for continuous or rank variables is not generally considered.^[12] In our study, only 7 articles fulfilled the corresponding criterion. This may be due to nonavailability of an automatic option for this test in current statistical software. Moreover, it may be related to the level of technical expertise. The linear gradient of continuous covariates can be easily resolved through a flexible modeling approach, such as spline regression (or segmented regression), multivariable fractional polynomials,^[28] and generalized additive models.^[29] However, these modeling procedures are more complex and cannot be achieved by common statistical software. Moreover, they are difficult to interpret and understand, even for experts. These problems limit the use of these complex statistical models in the medical field.^[7,10-12]

The selected journals in this study have high impact factor and large circulation. Since the evaluation criteria are currently acceptable, the same problems are likely to be found in other Chinese clinical medical journals. Undoubtedly, there are limitations in our study. We evaluated only articles from 5 leading journals. This is not likely to reflect conformance with the criteria comprehensively. Moreover, we cannot ensure whether certain data were missing due to actual failure to perform the corresponding test, or space limitations in the article.^[2] The authors might feel that it was unnecessary to perform these test or had good reasons to make exemptions.^[9] Finally, a comparison study between Chinese and non-Chinese journals may be necessary to decide whether this is a global issue.

In conclusion, despite the fact that reporting quality of MLR used in Chinese clinical medical journals has increased, severe deficiencies were noted. The main reasons behind these deficiencies may be: lack of statistical expertise and software application ability; nonavailability of automatic option for complex analysis in current statistical software; inadequate training in statistical methods among medical researchers or medical professionals; and absence of guidelines for appropriate reporting. Moreover, researchers may occasionally be unwilling to perform complex statistical tests and collaborate with biostatisticians.

The reporting quality and reliability of MLR models can be improved by editorial amendments, peer review, and a statistical review system.^[30,31] When MLR models are inaccurately constructed and improperly reported, it is difficult for researchers and readers to understand the results, and reproduce the models for future research. Hence, we recommend authors, readers, reviewers, and editors to become more acquainted with the use and reporting of MLR models. Journal editors should be more specific and proactive about the requirements for the publication of MLR and relax the word limit in the statistical analysis section,

where interaction, multicollinearity, and linear gradient should be reported. Moreover, journals develop statistical reporting guidelines concerning MLR and encourage researchers to collaborate with statisticians and epidemiologists to improve accuracy and the quality of reporting.

Acknowledgments

The authors thank Guoyi Yu, director of the editorial board of Journal of Qingdao University Medical College and Jianqiang Li, vice director of the editorial board of Journal of Qingdao University Medical College for providing helpful comments on early draft of this article. The authors also thank the reviewers for their insightful comments and suggestions.

References

- [1] Karan J, Goyal JP, Bhardwaj P, et al. Statistical reporting in Indian pediatrics. *Indian Pediatr* 2009;46:811–2.
- [2] Khan KS, Chien PF, Dwarakanath LS. Logistic regression models in obstetrics and gynecology literature. *Obstet Gynecol* 1999;93:1014–20.
- [3] Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993;118:201–10.
- [4] Campillo C. Standardizing criteria for logistic regression models. *Ann Intern Med* 1993;119:540–1.
- [5] Kumar R, Chhabra P. Cautions required during planning, analysis and reporting of multivariable logistic regression. *Curr Med Res Pract* 2014;4:31–9.
- [6] Moss M, Wellman DA, Cotsonis GA. An appraisal of multivariable logistic models in the pulmonary and critical care literature. *Chest* 2003;123:923–8.
- [7] Ottenbacher KJ, Ottenbacher HR, Tooth L, et al. A review of two journals found that articles using multivariable logistic regression frequently did not report commonly recommended assumptions. *J Clin Epidemiol* 2004;57:1147–52.
- [8] Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *J Clin Epidemiol* 2001;54:979–85.
- [9] Mikolajczyk RT, DiSilvestro A, Zhang J. Evaluation of logistic regression reporting in current obstetrics and gynecology literature. *Obstet Gynecol* 2008;111(2 Pt 1):413–9.
- [10] Kalil AC, Mattei J, Florescu DF, et al. Recommendations for the assessment and reporting of multivariable logistic regression in transplantation literature. *Am J Transplant* 2010;10:1686–94.
- [11] Tanboga IH, Kurt M, Isik T, et al. Assessment of multivariate logistic regression analysis in articles published in Turkish cardiology journals. *Turk Kardiyol Dern Ars* 2012;40:129–34.
- [12] Kumar R, Indrayan A, Chhabra P. Reporting quality of multivariable logistic regression in selected Indian medical journals. *J Postgrad Med* 2012;58:123–6.
- [13] Sun SG. The current situation and comment of logistic regression. *Chin J Health Stat* 1986;03:62–5.
- [14] Liu YY, Liu J, Hu LS, et al. The Theory and applications of generalized logistic regression for risk analysis of fuzzy states (5) – generalized age-dependent logistic regression model theory. *Chin J Health Stat* 2001;17.
- [15] Liu QJ. Logistic regression model and its research progress. *J Prev Med Inform* 2002;05:417–9.
- [16] Katz MH. Multivariable analysis: a primer for readers of medical research. *Ann Intern Med* 2003;138:644–50.
- [17] Allison PD. *Logistic Regression Using the SAS System: Theory and Application*. 1999;SAS Publishing.
- [18] Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J (Clin Res Ed)* 1986;292:746–50.
- [19] Hosmer DW, Taber S, Lemeshow S. The importance of assessing the fit of logistic regression models: a case study. *Am J Public Health* 1991;81:1630–5.
- [20] Feng GS, Chen JW, Zhou CL. Several easily overlooked issues in regression applications. *Chin J Epidemiol* 2004;06:92–3.
- [21] Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
- [22] Mullner M, Matthews H, Altman DG. Reporting on statistical methods to adjust for confounding: a cross-sectional survey. *Ann Intern Med* 2002;136:122–6.
- [23] Belsley DA, Kuh E, Welsch RE. Regression diagnostics: identifying in data and sources of collinearity. *J Operat Res Soc* 1981;2:157–8.
- [24] Yang JY, Yang HT. Diagnosis of collinearity of logistic regression using stata software. *Chin J Health Stat* 2005;03:174–6.
- [25] Zhao YD, Liu R, Liu YL, et al. Collinearity analysis of multivariate logistic regression. *Chin J Health Stat* 2000;05:3–5.
- [26] Yang M, Xiao J, Cai H. Multi-collinearity and corresponding processing method in multivariate analysis. *Chin J Health Stat* 2012;04:620–4.
- [27] Thiess MS, Arnold ZC, Walker SD. The misuse and abuse of statistics in biomedical research. *Biochem Med (Zagreb)* 2015;25:5–11.
- [28] Sauerbrei W, Meier-Hirmer C, Benner A, et al. Multivariable regression model building by using fractional polynomials: description of SAS, STATA and R programs. *Comput Stat Data Anal* 2006;50:3464–85.
- [29] Rao SJ, Stefanski LA. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. *J Am Stat Assoc* 2005;51:257–8.
- [30] Goodman SN, Altman DG, George SL. Statistical reviewing policies of medical journals: caveat lector? *J Gen Intern Med* 1998;13:753–6.
- [31] Goodman SN, Berlin J, Fletcher SW, et al. Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med* 1994;121:11–21.