


RESEARCH

Open Access



# Evaluation of the reliability of the criteria for assessing prescription quality in Chinese hospitals among pharmacists in China

Xiaofang Tan<sup>1</sup>, Jun Zhang<sup>1</sup>, Dongshi Gu<sup>1</sup>, San Ran<sup>1</sup>, Tongtong Gu<sup>1</sup>, Xiaowen Lin<sup>2</sup>, Enxiang Tao<sup>3</sup>, Tetsuya Asakawa<sup>3,4\*</sup>  and Huan Fang<sup>1\*</sup>

## Abstract

**Background:** The reliability of Criteria for Assessing Prescription Quality in Chinese Hospitals (CAPQCH) has never been rigorously verified. This study was designed to verify the reliability of the CAPQCH among pharmacists in China.

**Methods:** Fourteen pharmacists, 5 from hospitals and 9 from the communities were recruited. We randomly selected 200 prescriptions, and made the testing prescriptions including appropriate and inappropriate testing prescriptions. Pharmacists assessed these testing prescriptions according to criteria in CAPQCH. Three test sets (Set 1, Set 2, and Set 3) were evaluated at 6-month intervals. Before administration of Set 3, pharmacists were informed that achievement on Set 3 would be reflected in their performance appraisal. We also evaluated the performance based on prescription comments before and after combining several confusing criteria. Cohen's Kappa statistic, Fleiss' Kappa statistic, and accuracy were employed to evaluate reliability among pharmacists.

**Results:** Median values of Cohen's Kappa were 0.61 in Set 1, 0.66 in Set 2, and 0.80 in Set 3; reliability is thus substantial. Our data indicate no significant differences between Set 1 and Set 2, whereas Set 3 indicates significantly improved performance. Moreover, combinations of confusing criteria contributed little to improvement of performance in prescription comments.

**Conclusion:** Our results verified the reliability of CAPQCH application by working pharmacists. Adding performance based on prescription comments to personal appraisals was effective in improving the quality of prescription comments. These findings may be useful when future modification of the CAPQCH is considered. Moreover, this study contributes to improving the understanding of the prescription assessment situation in China.

**Keywords:** Reliability, Criteria for assessing prescription quality in Chinese hospitals (CAPQCH), Prescription comments, Cohen's kappa, Fleiss' kappa

## Background

Inappropriate prescription is common in aging populations and is a global public health concern because it leads to negative clinical outcomes, such as increasing adverse drug events, morbidity, and costs of hospitalization and care. The global population is aging, and reduction of numbers of inappropriate prescriptions to improve safe drug use is a challenge that must be faced by governments worldwide. Our earlier studies indicated

\*Correspondence: asakawat1971@gmail.com; fanghanyu2010@163.com

<sup>1</sup> Department of Pharmacy, Jinshan Hospital of Fudan University, No. 1508 Longhang Road, Shanghai 201508, PR China

<sup>3</sup> Department of Neurology, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shennanzhong Road 3025, Shenzhen 518033, Guangdong Province, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

that improving the spontaneous reporting system [1] may contribute to reducing medication errors [2], and thereby improving prescription quality. Establishing compelling screening tools based on prescription comments and issued by public health authorities might play a crucial role in reducing inappropriate prescription. The most important tool for identifying inappropriate prescriptions is the Beers criteria, first reported in 1991 [3], and updated in 2003 [4]. Later, several screening tools were developed and used, such as the Screening Tool of Older Peoples Prescriptions (STOPP) and the Screening Tool to Alert doctors to Right Treatment (START) [5], the Pharmaceutical Care Network Europe (PCNE) [6], the Medication Appropriateness Index (MAI) [7], and the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) [8]. The Chinese government also issued the Criteria for Assessing Prescription Quality in Chinese Hospitals (CAPQCH) in 2010 [9]. This guidance is a 28-item criterion that guides development of prescription comments that must be reported by pharmacists (Table S1).

Reliability of prescription comments is a critical issue. The reliability of the widely used tools, such as Beer's criteria, STOPP, START, PCNE, and MAI, has been verified. Inter-rater reliability is typically used to assess reliability of such tools. Higher inter-rater reliability indicates higher consistency among different examiners. This approach can provide compelling evidence for overall reliability of screening tools. Ryan et al. evaluated inter-rater reliability of STOPP and START with 10 pharmacists from hospitals and the communities. Finally, inter-rater reliability of STOPP and START was judged to be "good" for both pharmacists in hospitals and the communities [10]. Hohmann et al. assessed inter-rater reliability of PCNE with nine pharmacists and three pharmacy interns. They found that inter-rater reliability was substantial for main criteria and moderate for sub-criteria [11]. Stuijt et al. evaluated inter-rater reliability of MAI with eight raters and found an overall agreement rate of 83%, the overall chance-adjusted inter-rater agreement was moderate ( $Kappa = 0.47$ ), and the overall intra-group agreement was excellent ( $Kappa > 0.84$ ). The authors concluded that MAI is a reliable tool for evaluation of prescription [7]. Inter-rater reliability is thus a powerful tool for evaluating the overall reliability of a screening tool.

The reliability of CAPQCH has never been rigorously verified. During practical application, some reports in Chinese documented that some criteria, namely "unsuitable usage and dosage," "no or incomplete clinical diagnosis," "improper drug combination," and "incompatibility or bad interactions between drugs," might be confusing [12, 13]. We designed our study based on this knowledge

to 1) verify inter-rater reliability of CAPQCH, 2) assess the relationship of performance of CAPQCH and personal performance appraisals, and 3) examine the impact of a combination of reported confusing criteria on consistency. Our study provides evidence for the reliability of CAPQCH and will be beneficial to the future modification of this tool. The findings of this study will contribute to improving the understanding of the prescription assessment situation in China as well as the methodology for evaluating the reliability using the inter-rater study.

## Methods

### Participants

Fourteen pharmacists were recruited, 5 from the hospital sector (recorded as "HPs" in this study) and 9 from the community sector (recorded as "CPs" in this study). Subjects work in hospitals within the Jinshan district in the south of Shanghai city, China (Table 1). The working contents between HPs and CPs are different in this city. The HPs are mainly in charge of the clinical pharmacy in a local hospital. Prescription assessment is their routine work. Whereas the CPs are working in the community (community medical service). They mainly provide pharmaceutical service to a specific community region and perform prescription assessment for the doctors in community. Because the diseases and patients were different between hospital and community, the prescriptions they reviewed are slightly different. Because the use of the Criteria for Assessing Prescription Quality in Chinese Hospitals is part of a standard practice in the hospital, this study has been granted an exemption from requiring ethics approval by the Ethics Committee of Jinshan Hospital

**Table 1** Characteristics of raters

Number of raters	Position	Gender	Age	Experience in pharmacy (years)
1	Hospital	Female	40	19
2	Hospital	Female	34	14
3	Hospital	Female	34	10
4	Hospital	Male	34	10
5	Hospital	Female	28	5
6	Community	Female	33	11
7	Community	Male	47	29
8	Community	Female	50	31
9	Community	Female	41	20
10	Community	Female	35	16
11	Community	Female	38	16
12	Community	Female	33	11
13	Community	Female	53	30
14	Community	Male	34	13

of Fudan University. The informed consent is also waived by the Ethics Committee of Jinshan Hospital of Fudan University. All the participants were informed the purposes of the study prior to data collection.

### Testing prescriptions

Prescriptions for the study were first selected randomly by two experienced pharmacists from the electronic prescription system. A total of 200 prescriptions written from September to October 2019 were selected, which were then evaluated and modified to form two serials of testing prescriptions, namely “appropriate” prescriptions and “inappropriate” prescriptions using CAPQCH. All inappropriate prescriptions were modified to ensure that there is only “one mistake” in it. Mistakes were chosen to match expected professional knowledge and experience of raters. The “mistake” was defined according to CAPQCH and a previous study concerning CAPQCH [14] as 1) irregular prescription prescriptions, such as no or incomplete clinical diagnosis, no physicians’ signature on the prescription, or inconsistent with sample signature, nonstandard prescription paper, no signature and date for prescription revision, no reason and new signature for overdose use, and outpatient prescriptions for more than 7 days or emergency prescriptions more than 3 days without special conditions, and 2) inappropriate medication, such as unsuitable usage and dosage, prescriptions without indication, improper drug selection, or drug interactions, improper drug combination, incompatibility or dangerous interactions, and improper route of administration. Standard answers for test prescriptions were developed by the two pharmacists and finally checked and reviewed by an expert group of three senior pharmacists. Consensus was reached for all testing prescriptions and standard answers among pharmacists and experts before the testing began. We defined standard answers as “rater A”.

### Setting

All recruited pharmacists participated in three tests, given in September 2019 (Set 1), March 2020 (Set 2), and September 2020 (Set 3). The pharmacists assessed test prescriptions and identified “mistakes” according to guidance of prescription comments. Answers (raters 1–14) were then reviewed by the expert group. Set 1 and Set 2 tests were free from any connection to pharmacists’ appraisal. Before Set 3 (1 month before implementation), all pharmacists were informed that their achievement with Set 3 would be recorded in the assessment of the quality of their prescription comments, which is related to their performance appraisal. A pharmacist who achieved good performance on the test would earn a “good” score on his/her appraisal. Conversely, poor

performance would earn a “poor” score. The performance appraisal is regularly associated with the ranking of their department in the local public health administrative department.

### Statistics

All the answers were entered into a database and analyzed using SPSS software (V26.0.0, IBM, IL, USA). Cohen’s Kappa statistic was used to assess the consistency between raters 1–14 and rater A [15]. The degree of reliability was defined as 1) poor,  $\text{Kappa} \leq 0.20$ ; 2) fair,  $0.21 \leq \text{Kappa} \leq 0.40$ ; 3) moderate,  $0.41 \leq \text{Kappa} \leq 0.60$ ; 4) substantial,  $0.61 \leq \text{Kappa} \leq 0.80$ ; and 5) excellent,  $0.81 \leq \text{Kappa} \leq 1.00$  [10]. The proportion of positive agreement (Ppos) and the proportion of negative agreement (Pneg) were also evaluated [16].

Overall consistency within a single test was assessed using Fleiss’ Kappa statistic. Accuracy was calculated, by pharmacists, as the number of “right raters” (raters in agreement with rater A) divided by the number of ratings.

According to the previous studies [8, 12], some criteria might cause confusion during testing. We thus combined these terms. For example, criteria “inappropriate selection of indications” and “no or incomplete clinical diagnosis” and criteria “improper drug combination” and “incompatibility or bad interactions between drugs” were combined. We assessed changes in reliability with and without these combinations.

Normality of data distributions was assessed using the Kolmogorov–Smirnov test. Normally distributed data with Kolmogorov–Smirnov values over 0.05 were submitted for Levene’s test to verify homogeneity of variances. Paired t-tests were then used to compare differences between raters and rater A for each pharmacist. Independent sample t-tests were used to compare the difference between HPs and CPs. Data not normally distributed were compared using the Wilcoxon test.  $p < 0.05$  was considered statistically significant.

### Results

Seventy-nine percent of raters were female, five pharmacists were from hospitals, and nine were from the community. The average age was  $38.14 \pm 1.94$  y (first quartile–third quartile: 33.75–42.50 y). Average work experience was  $16.79 \pm 2.18$  y (first quartile–third quartile: 10.75–22.25 y) (Table 1).

Inter-rater reliability of pharmacists was evaluated using Cohen’s Kappa statistic. Median values of the statistic for sets 1, 2, and 3 were 0.61 (95% CI, 0.54–0.67), 0.66 (0.55–0.75), and 0.80 (0.74–0.87), respectively. These values exceed 0.6, indicating good reliability for all tests. An increasing tendency in the statistic is seen from sets

1 to 3. No significant difference was found between Set 1 and Set 2. The statistic for Set 3 was significantly greater than those for Set 1 ( $p < 0.001$ ) and Set 2 ( $p < 0.05$ ). Further, the lowest values of Cohen’s Kappa statistic were 0.33 (95% CI, 0.15–0.48) in Set 1 and 0.42 (0.25–0.58) in Set 2, indicating a “moderate” Cohen’s Kappa statistic in the first two tested sets. Accordingly, proportions of values lower than 0.6 were 50% in Set 1, 29% in Set 2, and 0% in Set 3 (Table S2). No significant differences between HPs and CPs were seen, although values of hospital pharmacists were greater than those of pharmacists from the community. The largest Kappa statistic was found for HPs in the Set 3 test—0.86 (0.75–0.88)—achieving “excellent”. The lowest was for CPs in the Set 1 test—only 0.59 (0.48–0.65)—“moderate”. When Kappa statistics among sets were compared, Set 3 was significantly elevated over sets 1 and 2 in CPs ( $p < 0.05$ ), but was significantly higher than only Set 1 ( $p < 0.05$ ) in HPs (Table 2).

Evaluation of overall reliability using Fleiss’ Kappa had the same results as those of reliability indicated by Cohen’s Kappa. Values of Fleiss’ Kappa were lowest in the Set 1 test and increased for Set 2 and increased further for Set 3 for both CPs and HPs.

No significant difference was found between HPs and CPs for accuracy among Set results. Overall, no significant difference was found between sets 1 and 2; however, data from Set 3 were significantly greater than either Set 1 or Set 2 ( $p < 0.05$ ) (Table 3).

Thus, the inclusion of performance on prescription comments in personal performance appraisals may enhance the quality of comments.

**Table 2** Comparison of reliability between pharmacists in hospitals and those in the community on three sets of prescriptions using Cohen’s Kappa

Comparators	Ppos	Pneg	Median Kappa (95% CI)
<b>Set 1</b>			
SA			
HPs	0.64	0.98	0.63 (0.59–0.74)
CPs	0.61	0.98	0.59 (0.48–0.65)
<b>Set 2</b>			
SA			
HPs	0.76	0.99	0.75 (0.62–0.81)
CPs	0.66	0.99	0.64 (0.46–0.71)
<b>Set 3</b>			
SA			
HPs	0.86	0.99	0.86 (0.75–0.88) Δ
CPs	0.76	0.99	0.75 (0.73–0.86) Δ, #

Ppos Proportion of positive agreement, Pneg Proportion of negative agreement, HPs Hospital pharmacists, CPs Community pharmacists, SA Standard answers

Δ means  $p < 0.05$ , Set 3 vs. Set 1; # means  $p < 0.05$ , Set 3 vs. Set 2

**Table 3** The inter-rater agreement among 14 pharmacists for three sets of prescription evaluations

Comparators	TPs	HPs	CPs
<b>Inter agreement Fleiss’ Kappa (95% CI)</b>			
Set 1	0.44 (0.40–0.48)	0.62 (0.50–0.73)	0.38 (0.32–0.44)
Set 2	0.51 (0.47–0.54)	0.62 (0.50–0.73)	0.47 (0.41–0.53)
Set 3	0.68 (0.64–0.72)	0.69 (0.57–0.80)	0.69 (0.63–0.75)
<b>Accuracy Rate (%)</b>			
Set 1	61.22	67.14	57.94
Set 2	65.76	73.10	61.69
Set 3	80.54 Δ, #	82.76 Δ	79.31 Δ, #

TPs Total pharmacists, HPs Hospital pharmacists, CPs Community pharmacists

Δ means  $p < 0.05$ , Set 3 vs. Set 1; # means  $p < 0.05$ , Set 3 vs. Set 2

Reliability evaluated using the Cohen’s Kappa statistic and accuracy after combining confusing criteria based on a previous study [8] indicated no impact in the Set 3 test and small increases in Set 1 and Set 2 tests for TPs. This tendency was unchanged regardless of which test criteria were combined: Set 3 > Set 2 > Set 1 (Table 4). Tendencies in accuracy were in complete agreement with Cohen’s Kappa (Table 5). These data suggested that when performance of consistency is poor (Set 1 and Set 2), combination of the confusing criteria might be able to mildly improve reliability. When the performance of consistency is satisfactory (Set 3), the improvement that results from criterion-adjustment is quite limited.

**Discussion**

We verified inter-rater reliability of CAPQCH with 14 pharmacists from Jinshan district using three sets of prescriptions. Cohen’s Kappa and Fleiss’ Kappa statistics were used to evaluate ratings. Cohen’s Kappa statistic along with accuracy can be used to evaluate the quality of prescription comments and to compare performance between CPs and HPs. The Fleiss’ Kappa statistic can be used to evaluate overall performance on prescription comments. This statistic was used to evaluate overall consistency among pharmacists and to explore dynamic changes in consistency among raters. Median values of Cohen’s Kappa were 0.61 for Set 1, 0.66 for Set 2, and 0.80 for Set 3. Inter-rater reliability was substantial. No significant differences in values of Kappa statistics were observed between Set 1 and Set 2; however, the statistic for Set 3 was significantly higher than statistics for the other two sets. Interestingly, evaluation of overall reliability indicated that the Fleiss’ Kappa statistic in Set 3 was improved for both CPs and HPs. Thus, associating performance using CAPQCH with performance appraisals will significantly increase reliability. In addition, the combination of reported “confusing” criteria contributes little

**Table 4** The inter-rater reliability among 14 pharmacists for three sets of prescriptions compared using Cohen’s Kappa statistics after combining confusing criteria

	TPs	HPs	CPs
<b>Median Kappa 95% CI)</b>			
<b>Set1</b>			
Before combination	0.61 (0.54–0.67)	0.63 (0.59–0.74)	0.59 (0.48–0.65)
After combination	0.63 (0.57–0.70)*	0.63 (0.61–0.76)	0.59 (0.50–0.67)
<b>Set2</b>			
Before combination	0.66 (0.55–0.75)	0.75 (0.62–0.81)	0.64 (0.46–0.71)
After combination	0.68 (0.57–0.75)*	0.75 (0.64–0.84)	0.64 (0.52–0.71)
<b>Set3</b>			
Before combination	0.80 (0.74–0.87) Δ, #	0.86 (0.75–0.88) Δ	0.75 (0.73–0.86) Δ, #
After combination	0.80 (0.74–0.89) Δ, #	0.86 (0.77–0.89) Δ	0.75 (0.73–0.86) Δ, #

TPs Total pharmacists, HPs Hospital pharmacists, CPs Community pharmacists, CI Confidence interval. \* means  $p < 0.05$ , before vs. after combination; Δ means  $p < 0.05$ , Set 3 vs. Set 1; # means  $p < 0.05$ , Set 3 vs. Set 2

**Table 5** Accuracy among 14 pharmacists for three sets of prescriptions after combining confusing items

	TPs	HPs	CPs
<b>Accuracy Rate (%)</b>			
<b>Set1</b>			
Before combination	61.22	67.14	57.94
After combination	62.76*	68.57	59.52
<b>Set2</b>			
Before combination	65.76	73.10	61.69
After combination	67.24*	75.17	62.84
<b>Set3</b>			
Before combination	80.54Δ, #	82.76Δ	79.31Δ, #
After combination	81.03Δ, #	84.14Δ	79.31Δ, #

TPs Total pharmacists, HPs Hospital pharmacists, CPs Community pharmacists, CI Confidence interval. \* means  $p < 0.05$ , before vs. after combination; Δ means  $p < 0.05$ , Set 3 vs. Set 1; # means  $p < 0.05$ , Set 3 vs. Set 2

to improving reliability. To the best of our knowledge, this study is the first to verify the reliability of CAPQCH and provides compelling evidence for its application, as well as any future modification.

A traditional inter-rater test was performed to evaluate consistency among 14 pharmacists by comparing their responses (raters 1–14) to standard answers (rater A) (Tables 2 and S2). Despite possible differences in work experience, the professional abilities of CPs and HPs did not differ for either Kappa statistics or accuracy. After connecting performance with CAPQCH to their personal appraisals (Set 3), the Kappa statistics for CPs and HPs significantly increased. The quality of prescription comments can be improved if performance was associated with appraisals. All pharmacists paid more attention to prescription comments and might take measures to increase their knowledge to achieve better performance.

Perhaps these possibilities explain performance on Set 3. In comparison with previous studies and despite different experimental approaches, Kappa statistics in the present study (Set 1 and Set 2, non-intervention) are consistent with NCC MERP, approximately 0.61–0.63 [8], but lower than values from STOPP and START, typically over 0.85 [10]. This difference might be explained by the objectivity and specificity of STOPP and START, which reduce the subjectivity of the pharmacists during the preparation of prescription comments, thus achieving better consistency [17, 18].

Evaluation of overall consistency using the Fleiss’ Kappa statistic produced analogous results (Table 3). No significant difference was found between CPs and HPs for Fleiss’ Kappa statistic or accuracy. Accuracy for Set 3 was significantly higher than those for sets 1 or 2. The Fleiss’ Kappa statistic exhibited the same tendency, but differences did not reach significance. Again, relating quality of prescription comments and personal performance appraisals will remarkably improve the quality of comments.

With respects to the influence of “confusing items”. Finally, our results imply that these criteria have a little effect on performance for prescription comments. In sets 1 and 2, only TPs displayed significant changes (Tables 4 and 5). We believe that the significant increase in Kappa statistics and accuracy in Set 3 is due to connecting prescription comments with appraisals, rather than combining confusing criteria. This conclusion implies that reinforcing education and training might improve the abilities of pharmacists to implement CAPQCH. These issues require further investigation in future studies. Conversely, our data indicate limited influence of “confusing criteria” on the reliability of the current version of CAPQCH.



In the light of the findings of the present study, we understand that CAPQCH is a reliable tool to evaluate the potentially inappropriate prescription in China. It can be widely used in both HPs and CPs. Adding quality of prescription comments to personal performance appraisals of the pharmacists might improve the performance for the prescription comments, which should be noticed by those who are in charge of the hospital management. In addition, the “confusing items” seldom affected the performance for prescription comments, which should be taken into account in the future modification of the CAPQCH.

There are several limitations in this study. First, regarding the recruitment of pharmacists, although we had recruited all the pharmacists working in our hospital (five) and our community (nine), the sample size was not determined by a strict statistical method. This might lead a sampling bias. In addition, the small size of the participants might lead a biased conclusion. Second, we compared the potentially inappropriate medications between CAPQH and the other tools like NCC MERP, STOPP, and START directly using Kappa values. CAPQH is a list of items to check for formal content. Simply comparison of the Kappa values between CAPQH and the other tools might lead a biased conclusion because the purpose and contents of each tool are very different. We need to consider the results of CAPQH in light of its content. These issues will be fully considered and improved in our future investigations.

## Conclusions

This study verified the reliability of CAPQCH using a traditional inter-rater reliability test, which indicated that CAPQCH is a reliable tool to classify the inappropriate prescription in China. Our results suggest a little influence of reported confusing criteria. These items should be considered in the future modification of the CAPQCH. Instead, adding quality of prescription comments to personal performance appraisals might be more useful for improving comments. These findings contribute to our understanding of reliability of CAPQCH and may be useful for future modifications. In addition, this study contributes to improving the understanding of the prescription assessment situation in China.

## Abbreviations

CAPQCH: Criteria for Assessing Prescription Quality in Chinese Hospitals; CPs: Pharmacists from the community sector; HPs: Pharmacists from the hospital sector; MAI: Medication Appropriateness Index; NCC MERP: National Coordinating Council for Medication Error Reporting and Prevention; PCNE: Pharmaceutical Care Network Europe; START: Screening Tool to Alert doctors to Right Treatment; STOPP: Screening Tool of Older Peoples Prescriptions.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12913-022-07854-1>.

**Additional file 1: Table S1.** CAPQCH items. **Table S2.** The inter-rater reliability of criteria between rater A and raters 1–14 for three sets of prescriptions using Cohen’s Kappa statistics.

## Acknowledgements

The authors would like to thank Enago ([www.enago.cn](http://www.enago.cn)) for the English language review.

## Authors’ contributions

TA and HF designed the study. XT, JZ, DG, SR, TG, XL, ET, TA, and HF performed the research and analyzed the data. XT, TA, and HF wrote the draft. All authors read and approved the final manuscript.

## Funding

No funding was involved in the present study.

## Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

This study has been granted an exemption from requiring ethics approval by the Ethics Committee of Jinshan Hospital of Fudan University. The informed consent is also waived by the Ethics Committee of Jinshan Hospital of Fudan University. All the methods were carried out in accordance with the laws in China.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Department of Pharmacy, Jinshan Hospital of Fudan University, No. 1508 Longhang Road, Shanghai 201508, PR China. <sup>2</sup>Department of Pharmacy, Jinshan Central Hospital, Shanghai 201500, China. <sup>3</sup>Department of Neurology, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shennanzhong Road 3025, Shenzhen 518033, Guangdong Province, China. <sup>4</sup>Research Base of Traditional Chinese Medicine Syndrome, Fujian University of Traditional Chinese Medicine, Fuzhou 350122, China.

Received: 16 November 2021 Accepted: 14 March 2022

Published online: 07 April 2022

## References

- Fang H, Lin X, Zhang J, Hong Z, Sugiyama K, Nozaki T, et al. Multifaceted interventions for improving spontaneous reporting of adverse drug reactions in a general hospital in China. *BMC Pharmacol Toxicol.* 2017;18(1):49. <https://doi.org/10.1186/s40360-017-0159-0>.
- Tan X, Gu D, Lin X, Fang H, Asakawa T. Investigation of the characteristics of medication errors and adverse drug reactions using pharmacovigilance data in China. *Saudi Pharm J.* 2020;28(10):1190–6. <https://doi.org/10.1016/j.jsps.2020.08.008>.
- Beers MH, Ouslander JG, Rollingher I, Reuben DB, Brooks J, Beck JC. Explicit criteria for determining inappropriate medication use in nursing home residents. UCLA division of geriatric medicine. *Arch Intern Med.* 1991;151(9):1825–32.
- Fick DM, Cooper JW, Wade WE, Waller JL, Maclean JR, Beers MH. Updating the Beers criteria for potentially inappropriate medication use in older

- adults: results of a US consensus panel of experts. *Arch Intern Med.* 2003;163(22):2716–24. <https://doi.org/10.1001/archinte.163.22.2716>.
5. Gallagher P, Ryan C, Byrne S, Kennedy J, O'Mahony D. STOPP (screening tool of older Person's prescriptions) and START (screening tool to alert doctors to right treatment). Consensus validation. *Int J Clin Pharmacol Ther.* 2008;46(2):72–83. <https://doi.org/10.5414/cpp46072>.
  6. Eichenberger PM, Lampert ML, Kahmann IV, van Mil JW, Hersberger KE. Classification of drug-related problems with new prescriptions using a modified PCNE classification system. *Pharm World Sci.* 2010;32(3):362–72. <https://doi.org/10.1007/s11096-010-9377-x>.
  7. Stuijt CC, Franssen EJ, Egberts AC, Hudson SA. Reliability of the medication appropriateness index in Dutch residential home. *Pharm World Sci.* 2009;31(3):380–6. <https://doi.org/10.1007/s11096-009-9283-2>.
  8. Forrey RA, Pedersen CA, Schneider PJ. Interrater agreement with a standard scheme for classifying medication errors. *Am J Health Syst Pharm.* 2007;64(2):175–81. <https://doi.org/10.2146/ajhp060109>.
  9. Chen YC, Tao R. The criteria of assessing the prescription quality in Chinese hospital (CAPQCH, the 2nd trial version). *World Clinical Drugs.* 2010;31:259–60.
  10. Ryan C, O'Mahony D, Byrne S. Application of STOPP and START criteria: interrater reliability among pharmacists. *Ann Pharmacother.* 2009;43(7):1239–44. <https://doi.org/10.1345/aph.1M157>.
  11. Hohmann C, Eickhoff C, Klotz JM, Schulz M, Radziwill R. Development of a classification system for drug-related problems in the hospital setting (APS-doc) and assessment of the inter-rater reliability. *J Clin Pharm Ther.* 2012;37(3):276–81. <https://doi.org/10.1111/j.1365-2710.2011.01281.x>.
  12. Zhao HQ. Thoughts and suggestions on unreasonable prescription criteria for prescription comment. *Chinese J Pharmacovigilance.* 2012;9(6):367.
  13. Ping Lin XW, Zhao HQ, Zhen J. Study on the impact of evaluators on results of prescription evaluation. *Chinese Pharmacy.* 2012;23(45):4313–4.
  14. Yanyan W, Chen L. Application of PDCA cycle Management in the Outpatient Prescription Intervention of a hospital. *China Pharmacy.* 2017;28(8):1129–32.
  15. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005;37(5):360–3.
  16. Winters M, Bakker EWP, Moen MH, Barten CC, Teeuwen R, Weir A. Medial tibial stress syndrome can be diagnosed reliably using history and physical examination. *Br J Sports Med.* 2018;52(19):1267–72. <https://doi.org/10.1136/bjsports-2016-097037>.
  17. Lam MP, Cheung BM. The use of STOPP/START criteria as a screening tool for assessing the appropriateness of medications in the elderly population. *Expert Rev Clin Pharmacol.* 2012;5(2):187–97. <https://doi.org/10.1586/ecp.12.6>.
  18. Corsonello A, Onder G, Abbatecola AM, Guffanti EE, Gareri P, Lattanzio F. Explicit criteria for potentially inappropriate medications to reduce the risk of adverse drug reactions in elderly people: from Beers to STOPP/START criteria. *Drug Saf.* 2012;35(Suppl 1):21–8. <https://doi.org/10.1007/BF03319100>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

