Pharmacy Education (2024) 24(1) 403 - 417
https://doi.org/10.46542/pe.2024.241.403417

RESEARCH ARTICLE

Development and a validation study of comprehensive prescription writing rubrics for medical students

Anupong Kantiwong, Sethapong Lertsakulbunlue
Department of Pharmacology, Phramongkutklao College of Medicine, Bangkok, Thailand

Keywords
Drug prescription
Generalisability theory
Medical student
Rubric

Correspondence
Sethapong Lertsakulbunlue
Department of Pharmacology
Phramongkutklao College of Medicine
Bangkok
Thailand
sethapong.ler@pcm.ac.th

Abstract

Background: Prescribing is a complex task for physicians, with many global reports of errors. This study evaluates a comprehensive rubric for medical student prescribing skills regarding validity and reliability. **Methods:** Twenty-one third-year medical students participated in three separate prescribing exams. Two pharmacology professors rated the students' prescriptions using a rubric covering ten criteria. Messick validity framework was utilised to enhance the study's validity. Generalisability theory (G-theory) helped determine the source of variance and the optimal number of raters and test occasions. **Results:** Content validity was ensured by three experts and alignment with the Thai Medical Council criterion. The Inter-rater and test-retest reliability were acceptable. The rubric had a Cronbach's alpha 0.70 with item-test correlation, all above 0.40. G-theory indicated that 54.93% of the total variance was due to performance and 27.57% to the interaction between performance and occasions, with a minimal residual variance of 4.28%. To reach an acceptable Phi-coefficient (≥0.70), three occasions with one rater (Phi-coefficient=0.76) or two occasions with two raters (Phi-coefficient=0.72) are needed. Conversely, the Phi-coefficient was low on a single occasion. **Conclusion:** The study introduces a comprehensive rubric and description of a prescription writing programme to minimise potential prescribing errors in pre-clinical years. Furthermore, more assessment opportunities enhance knowledge retention and assessment reliability.

Introduction

Prescribing medications constitutes a major responsibility for physicians, with newly graduated doctors responsible for a substantial portion of prescriptions in inpatient and outpatient settings. These prescriptions encompass not only specific treatments but also support patient care. Furthermore, in Thailand, the Medical Council has established prescription writing skills as a requisite professional standard that medical students must accurately and comprehensively meet. However, the act of prescribing is a complex task that requires doctors to carefully select the appropriate drug, dosage, frequency, and route of administration while also considering potential drug allergies, interactions, and pre-existing comorbidities (Aronson, 2012).

Previous literature indicates that medical students may not possess adequate prescribing skills (James *et al.*, 2016). A study also revealed that physicians and medical students might have lower prescription writing competency compared to pharmacists and pharmacy students, highlighting the importance of improving prescription writing skills among medical students (Keijsers *et al.*, 2015). This deficiency could stem from limited experience and insufficient assessment during their pre-clerkship years (Heaton *et al.*, 2008). Inadequate training in prescription writing during medical school may result in subpar prescriptions upon graduating as a Doctor of Medicine, which can negatively impact patient care and lead to preventable adverse events (Ryan *et al.*, 2014; Wiernik, 2015). Thailand has reported a higher prevalence of medication errors compared to other Asian countries (Salmasi *et al.*, 2015). In the United States, such errors

can lead to significant health and economic consequences, with estimates suggesting that avoidable medication errors could cost between $17 billion and $29 billion annually (Gautam, 2013).

Given the high stakes involved, medical students must be proficient in prescription writing before encountering patients. Consequently, teaching methods for prescription writing have evolved continuously over the past decade (Linton & Murdoch-Eaton, 2020). Nevertheless, it has been noted that numerous medical schools may not offer thorough instruction in prescription skills, falling short of expectations from students, faculty, and the institutions themselves (Sequeira, 2015; Rothwell *et al.*, 2012). Moreover, a meta-analysis in 2017 that reviewed 47 studies on medical students' prescription writing globally revealed that even final-year students might still not be writing prescriptions at the expected standard (Brinkman *et al.*, 2018).

Several assessment rubrics have been developed to minimise medication errors in clinical contexts and during pre-clerkship years (Fanikos *et al.*, 2014; Baranski *et al.*, 2017; Loskutova *et al.*, 2022). Additionally, efficient tools like using very short answer question formats have proven effective in identifying prescribing errors, outperforming traditional single best answer questions (Sam *et al.*, 2019). However, these tools focus primarily on the medication name, dose, route, and frequency and do not fully encompass the entire spectrum of potential drug prescribing errors, such as handwriting legibility, error corrections, incorrect patient names, and allergies (Salmasi *et al.*, 2015). Hence, there is a need for a comprehensive rubric for medical students that offers clear performance standards and specific feedback to enhance the understanding of prescribing skills and facilitate the identification of potential errors (Hamstra & Yamazaki, 2021; Hill *et al.*, 2022).

According to Messick's validity framework, five key elements can be employed to ensure the validity of the research: 1) content validity, including using previously developed instruments and conducting expert reviews of draft items, 2) response process, emphasising clear test instructions for candidates and rigorous rater training, 3) internal structure, includes reliability calculations such as Cronbach's alpha, inter-rater reliability, test-retest reliability, and generalisability theory, 4) the relations with other variables, and 5) the consequences of testing can be evaluated through associations between survey scores and theoretically related external variables, as well as by analysing the impacts of the tools (Messick, 1995; Hamstra & Yamazaki, 2021; Hill *et al.*, 2022).

The study aimed to determine how valid and reliable the prescription writing rubrics developed for third-year pre-clinical medical students are through Messick's framework. The second objective is to determine the sources of variance and how reliability is affected by the number of occasions and raters by employing Generalisability Theory (G-theory) and Decision Study (D-study). The findings will inform the refinement of prescription writing rubrics and the design of reliable and valid assessments for medical students.

Methods

Study design and subjects

In calculating the sample size, the total sample size is derived from three levels: the number of students, occasions, and raters (students × occasions × raters). A total sample size of 80 was required for an effect size f of 0.45 with 95% power at a significance level of 0.05, with a numerator degree of freedom of 2 and a total of 6 groups (2 raters × 3 occasions) using an ANOVA test in G*Power 3.1.9.7 (Faul *et al.*, 2007; Keijsers *et al.*, 2015). Therefore, dividing 80 by 6 equals approximately 13.33, which rounds up to 14 students. This study analysed the data from prescription writing scores from three separate sessions among 21 students, comprising 15 males and 6 females, who enrolled in an extracurricular clinical readiness course to improve prescription writing at the end of three system blocks at Phramongkutklao College of Medicine (PCM) in Bangkok, Thailand. Due to the analysis of secondary data, the study received an exemption approval from the Medical Department Ethics Review Committee for Research in Human Subjects, Institutional Review Board, Royal Thai Army (IRBRTA) (Approval no. S041h/66_Xmp).

The students were tested with three outpatient case scenarios developed based on common must-know drugs for third-year medical students, as per the Thai Medical Council guidelines. These scenarios aim to assess the student's skills in prescription writing for various drug forms in realistic, common situations. Five pharmacology instructors at PCM reviewed the scenario content iteratively. The instructors assembled and were asked to review the content, including the scenario's realisticity, language, clarity, and importance.

These scenarios included:

(1) A 65-year-old male with hypertension and benign prostatic hyperplasia requiring prescriptions for amlodipine and doxazosin. This aims to assess the

student's ability to prescribe oral drugs and their knowledge that doxazosin should be taken before bed. Furthermore, the drug name poses a challenge, as the trade name with a controlled-release (CR) designation must be mentioned.

(2) A 5-year-old boy with acute exudative tonsilitis experiencing nausea and vomiting, for whom syrup azithromycin and domperidone were prescribed. This aims to assess the student's ability to prescribe and calculate dosages for drugs in syrup form as well as challenge them to calculate the drug quantity for the visit.

(3) A 25-year-old female seeking contraception and managing type 1 diabetes was prescribed subcutaneous insulin and intramuscular medroxyprogesterone. This aims to assess the student's ability to prescribe injectable drugs. Additionally, insulin injections are prescribed in different portions for morning and evening, challenging students to write the drug regimen and quantity in the standard form correctly. Students also need to know the duration of medroxyprogesterone treatment to calculate the drug quantity accurately.

Appendix A shows examples of the case scenario in PowerPoint format, and Appendix B illustrates the specific learning points for each case.

Content

The assessment utilised three-tiered rubrics comprising ten criteria:

1) Personal information: The first name and last name, hospital number and date.

2) Drug allergy: The name of drugs and presentation of the patient's drug allergy.

3) Drug name: Correctly write the generic name of the drug.

4) Drug form: Specify the drug form correctly according to the standard and ensure that the route of administration corresponds to the dosage form.

5) Drug strength: Specify the drug strength correctly according to the standard with the unit, e.g. mg, mcg, g, mg/mL, unit.

6) Drug quantity: Calculate the drug quantity accurately and sufficiently for the entire duration of the treatment or for the period before the next appointment/follow-up.

7) Drug amount for each dose: Calculate the drug amount used for each dose correctly.

8) Dosage regimen: Identified the correct method of drug administration for both the frequency and the time intervals.

9. Prescriber information: Sign their names and provide identification number.

10. Prescription correction: No corrections are made to the prescription or corrections made appropriately with a signature for verification.

In the present rubric, if students do not write their prescriptions clearly and the raters are unable to read them, the unclear prescription is marked as incorrect. Thus, handwriting was not included as an item in the rubric.

The rubrics were developed in accordance with the prescription writing skills standards set by the Thai Medical Council, incorporating components of prescription writing and common prescription errors (Salmasi *et al.*, 2015). Content validity was ensured by three professors from PCM's pharmacology department using the IOC (Item-objective congruence) method. Each item attained a content validity index exceeding 0.67 out of 1.00, indicating good validity (above 0.50) as assessed by the professors. The rubric tiers are assigned as follows: 3 for "*Written Completely*", 2 for "*Partially Written/Incomplete*", and 1 for "*Not Written or Incorrect*". The overall scores were further stratified into three groups: "*Must Improve*", corresponding to scores ranging from 10 to 16, "*Mediocre*" for scores from 17 to 23, and "*Good*" for scores between 24 to 30. The complete assessment form is presented in Appendix C.

Response processes

Five pharmacology professors of PCM reviewed and proofread the case scenario and provided individual comments through IOC methods. Amendments were made before a final meeting, during which approval was granted by all professors. Before the examination, students were acquainted with each item on the rubrics to ensure a shared understanding and agreement on their interpretation. An initial prescribing exercise followed by a practice session to familiarise students with the prescription writing process, examination format, and rating criteria. The tests were conducted separately at the end of each block. Each session was conducted independently, spaced about a month apart. During these sessions, students faced a simulated scenario under exam conditions and were tasked with writing prescriptions on actual pads used at Phramongkutklao Hospital. Subsequently, independent raters evaluated the completed prescription pads. Two pharmacology department teachers from PCM who assisted with verifying the rubrics, each with over ten

years of teaching experience and use of rubrics, served as raters.

Statistical analysis

The data were analysed using IBM SPSS Statistics for Windows, Version 29.0 (Armonk, NY: IBM Corp.). Categorical data were expressed as percentages, while continuous variables were presented as means with standard deviations (SD). The internal reliability of the assessment tool was assessed using Cronbach's alpha. Pearson's correlation was used to calculate inter-rater reliability. Two-way random-effects model Intraclass correlation was used to calculate the test-retest reliability. The cutoff alpha and inter-rater reliability coefficient is ≥0.70 for an acceptable result, ≥0.80 for very good, and ≥0.90 for excellent outcomes (George & Mallery, 2019). For the intraclass correlation coefficient (ICC), values less than 0.50 are indicative of poor reliability, values between 0.50 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability and values greater than 0.90 indicate excellent reliability (Koo & Li, 2016).

To enhance the reliability assessment of the instrument, a generalisability theory analysis using a 3-way ANOVA, or person-by-occasion-by-rater (p×o×r) design, was conducted. This enabled a fully crossed person (P), test occasions (O), and raters (R) design. The analysis identified the variance in measurements due to the facets of the study (Briesch *et al.*, 2014), and variance components were subsequently calculated (Donnon *et al.*, 2013). The present analysis comprised seven components of variance: the principal effects of persons (P), occasions (O), raters (R), two-way interactions between persons and occasions (PO), persons and raters (PR), and occasions and raters (OR). Furthermore, the residual error variance (PRO, e) was incorporated to incorporate the influence of interactions among all facets as well as other unidentified sources of variability. However, the number of items was limited to ten to maintain comprehensiveness, covering all aspects of prescription error; therefore, the number of items was not included as part of the facets.

Furthermore, a two-facet crossed design for a decision study was carried out to explore variations in the G-coefficient under different conditions and to identify the optimal measurement strategy. To assess the reliability of individual facet combinations, the absolute G-coefficient (Phi-coefficient) is selected. The error term incorporates the Phi-coefficient, which corrects for any systematic (primary) effects of the facets that introduce error into the estimate. Because the scores comprising the student's prescription writing score and planned to be accumulated in the grade point average (GPAX) in the future are assessed according to predetermined criteria, as opposed to being relative to one another, the absolute coefficient was employed. Nevertheless, the relative G-coefficient was reported in Appendix D. The established minimum threshold for reliability is set at 0.70 for acceptable reliability (Brennan, 2010).

Results

Characteristics

Twenty-one third-year pre-clinical medical students at PCM underwent three prescription writing evaluations. The average scores assigned by the first rater for the prescription writing tasks were 26.29±2.35, 25.95±3.86, and 24.48±3.33 for the first, second, and third occasions, respectively. The average scores of the second rater for the first to third occasions were 25.38±2.33, 25.19±4.17, and 23.00±3.38, respectively. Stratified by categories, the first rater assigned 48 (76.19%) as "*Good*" and 15 (23.81%) as "*Mediocre*", while the second rater categorised 39 (61.90%) scores as "*Good*", 22 (34.92%) as "*Mediocre*", and 2 (3.17%) as "*Must Improve*".

Inter-rater, internal consistency and test-retest reliability

For the internal structure testing, the inter-rater reliability coefficients were $r = 0.919$ ($p = 0.001$) for the first occasion, $r = 0.972$ ($p = 0.001$) for the second occasion, and $r = 0.948$ ($p = 0.001$) for the third occasion (Table I). The overall Cronbach's alpha for the rubric criteria is 0.71. Additionally, Table II presents the item-test correlation, showing a good correlation range from 0.41 to 0.64. Cronbach's alpha for item deletion revealed that removing any items would reduce the internal consistency reliability. Further analysis of the average inter-item correlation revealed values from 0.18 to 0.21. Table III presents the inter-rater reliability and test-retest reliability for each rubric criterion. Overall, each item showed good inter-rater reliability, except for the "*drug name*" criterion, which had reliability coefficients of $r = 0.389$ ($p = 0.082$), $r = 0.141$ ($p = 0.541$), and $r = 0.375$ ($p = 0.094$) for the respective occasions. Additionally, the overall ICC stratified by occasion was 0.84 ($p = 0.001$). However, in the "*Calculated Drug Amount*" and "*Quantity*" criteria, the ICCs are relatively low at 0.16 ($p = 0.254$) and 0.31 ($p = 0.078$), respectively.

Table I: Scores for the prescription writing by pre-clinical medical students as stratified by rater 1 and rater 2

Occasion	Rater 1		Rater 2		*r*	*p*-value
	min-max	mean±SD	min-max	mean±SD		
1	22-29	26.29±2.35	21-29	25.38±2.33	0.919	0.001
2	17-30	25.95±3.86	16-30	25.19±4.17	0.972	0.001
3	18-29	24.48±3.33	16-29	23.00±3.38	0.948	0.001

SD: standard deviation, r: Pearson correlation coefficient determining the inter-rater reliability

Table II: Internal consistency of the prescription writing scores of pre-clinical medical students

Rubrics criteria	mean±SD	Item-total correlation	Average inter-item correlation	Conbrach's alpha if item deleted
Personal History	2.33±0.55	0.570	0.187	0.674
Drug allergy	2.80±0.59	0.581	0.185	0.672
Drug name	2.75±0.47	0.419	0.209	0.704
Drug form	2.12±0.87	0.615	0.180	0.664
Drug dosage	2.83±0.47	0.502	0.197	0.688
Calculation of drug quantity	2.42±0.83	0.553	0.189	0.678
Calculated the drug amount used for each dose	2.47±0.81	0.413	0.210	0.705
Drug administration	1.74±0.73	0.642	0.176	0.658
Prescriber information	2.89±0.44	0.525	0.193	0.683
Error correction	2.71±0.63	0.416	0.209	0.704

Total Conbrach's alpha: 0.706

Table III: Inter-rater reliability between rater 1 and rater 2 and test-retest reliability of the prescription writing of pre-clinical medical students

Rubrics criteria	Occasion 1			Occasion 2			Occasion 3			ICC	*p*-value
	mean±SD	*r*	*p*-value	mean±SD	r	*p*-value	mean±SD	*r*	*p*-value		
Personal history	2.31±0.56	0.917	0.001	2.29±0.60	0.890	0.001	2.38±0.49	1.000	0.001	0.896	0.001
Drug allergy	2.79±0.61	0.837	0.001	2.90±0.43	1.000	0.001	2.71±0.71	1.000	0.001	0.746	0.001
Drug name	2.67±0.57	0.389	0.082	2.83±0.38	0.141	0.541	2.74±0.45	0.375	0.094	0.412	0.021
Drug form	1.98±0.90	0.916	0.001	2.36±0.91	0.953	0.001	2.02±0.78	0.817	0.001	0.530	0.002
Drug dosage	2.95±0.22	1.000	0.001	2.88±0.45	0.689	0.001	2.67±0.61	0.696	0.001	0.550	0.001
Calculation of drug quantity	2.86±0.35	1.000	0.001	2.21±0.95	0.985	0.001	2.19±0.89	0.935	0.001	0.310	0.078
Calculated the drug amount used for each dose	2.95±0.22	1.000	0.001	2.38±0.85	0.972	0.001	2.07±0.89	0.911	0.001	0.156	0.254
Drug administration	1.71±0.60	0.602	0.004	2.05±0.79	0.983	0.001	1.45±0.67	0.769	0.001	0.571	0.001
Prescriber information	2.95±0.22	1.000	0.001	2.81±0.59	1.000	0.001	2.90±0.43	1.000	0.001	0.536	0.002
Error correction	2.67±0.72	0.775	0.001	2.86±0.35	0.669	0.001	2.60±0.73	0.548	0.010	0.402	0.024
Total	**25.81±2.33**	**0.919**	**0.001**	**25.57±3.99**	**0.972**	**0.001**	**23.74±3.39**	**0.948**	**0.001**	**0.841**	**0.001**

ICC: intraclass correlation coefficient determining the test-retest reliability, r: Pearson correlation coefficient determining the inter-rater reliability

Generalisability study

Table IV shows the results from the two-facet G-study for the p×o×r design, which assessed prescription writing scores across 21 pre-clinical medical students, three occasions, and two raters. The analysis reveals that 54.93% of the total variance is attributable to the students (P), representing the universe score. The variance component due to the occasion (O) accounts for 8.71%, while the raters (R) account for 4.15%. The percentage of variance due to the interaction between students' performance and occasions is relatively higher at 27.57%. Conversely, the variance due to the interaction between students' performance and raters, as well as the interaction between occasions and raters, is minimal. Finally, the residual variance comprises 4.28%, indicating that the current two-facet crossed design accounts for over 95.72% of the variance, representing the universe score.

Table IV: Generalisability study for p×o×r for drug prescription writing, among 21 pre-clinical medical students, 3 occasions and 2 raters

Source of variation	Estimated variance component	% of Total variance
Student (P)	6.945	54.93
Occasion (O)	1.102	8.71
Rater (R)	0.525	4.15
PO	3.486	27.57
PR	0.000	0.00
OR	0.046	0.36
Residual (POR, e)	0.542	4.28
Total	12.645	100.00

Decision study

Appendix D displays the D-Study for the p×o×r design, forecasting the reliability for different assessment occasions and raters combinations. The absolute generalisability (G) coefficient (Phi-coefficient) ranges from 0.55 to 0.87 for combinations of one to five raters across one to five test occasions. To achieve acceptable reliability, three occasions are required for one rater (Phi-coefficient=0.76). In the case of two raters, two occasions suffice to attain a Phi-coefficient of 0.72. For three occasions with two raters, the Phi-coefficient is 0.79. However, the absolute Phi-coefficient is insufficient for a single test occasion, varying only from 0.55 to 0.59 across one to five raters. Figure 1 displays the absolute generalisability coefficient for the p×o×r design, illustrating the changes across varying numbers of occasions and raters.

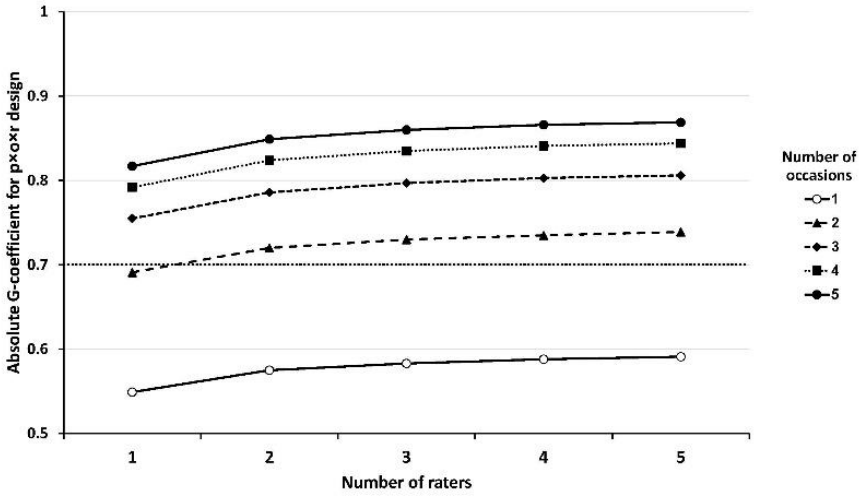


Figure 1: Decision study for the p×o×r design, evaluating drug prescription writing among 21 pre-clinical medical students across three testing occasions and with two raters. The coefficients indicate the projected absolute G-coefficient for various combinations of occasions and raters.

Discussion

This study applied the Messick validity framework to emphasise the validity and reliability of the developed prescription writing rubrics. Furthermore, to the best of the author's knowledge, the present study is the first to reveal the minimum number of occasions and raters to achieve reliable assessment in prescription writing. To achieve acceptable reliability with a phi-coefficient of 0.70 or greater using the current rubric, at least two testing occasions with two raters are necessary. Furthermore, on three occasions, only one rater is needed to facilitate the assessment of prescription writing. The variance in overall scores primarily arises from the student's performance and their performance's interaction across occasions, with a minimal residual variation. These results suggest that the current two-facet crossed design, employing two raters and three occasions, accounts for most of the variance.

Similar to the present study, Thenrajan & Murugan (2016) introduced a prescription writing assessment method that evaluated prescriptions using a 14-point scoring system, comprehensively covering the prescription steps. Moreover, previous studies aimed to assess prescription writing skills or prescription errors among medical students mainly utilised checklists (Gupta *et al.*, 2020; Khaja *et al.*, 2005; Yaman *et al.*, 2012). However, these studies usually did not comprehensively focus on the process of creating their criteria. Furthermore, the checklist approach might lack detailed feedback, hindering students' understanding of their strengths and weaknesses and potentially oversimplifying the complexities of student performance.

Another strength of this study is that it utilised the G-theory, an extension of classical reliability theory, assessing the primary variable of interest and subject performance and comparing it against error variance. It statistically determines the reliability and validity of education assessments by dissecting variance sources like raters, occasions, and students (Bloch & Norman, 2012). G-theory serves as an instrumental measure in evaluating and refining the methodological quality of assessments. Despite its benefits, related studies rarely utilise generalisability theory in similar contexts. A similar study on antibiotic prescription errors examined a reliable generalisability coefficient for the number of occasions and items but did not investigate the sources of variance or conduct a decision study (Martínez-Domínguez *et al.*, 2022). Decision studies help identify specific assessment errors and recommend the best assessment structures. This is especially useful in establishing the required number of occasions and raters to determine reliable and valid assessments (Bloch & Norman, 2012; Andersen *et al.*, 2021).

Several studies in educational context utilised G-theory to evaluate performance-based assessment using scoring rubrics often report a high percentage of unexplained residual variance (Bloch & Norman, 2012; Briesch *et al.*, 2014; Peeters *et al.*, 2021). In contrast, like the current study, written exams typically report a lower percentage of unexplained residual variance (Khodi, 2021). The present study employs a highly detailed and well-structured rubric to assess each step of the students' prescription writing, accounting for much of the potential variance.

The G-study in this research showed that a low percentage of the total variance is attributed to the rater facet, indicating that raters contribute minimally to the variance in scores. Furthermore, the inter-rater reliability in this study is excellent, with a correlation coefficient (r) greater than 0.90 across three occasions. When stratified by each rubric criterion, the inter-rater reliability remains good, except for the "*drug name*" criterion, which assesses the accuracy of the students' writing of generic drug names. This inconsistency arises from the differing approaches of the raters: Rater 1 awards marks to students who have a minor misspell or whose handwriting is difficult to decipher, whereas Rater 2 does not. Hence, conducting regular meetings to foster assessor interactions on a peer-to-peer basis is recommended. Such engagements are intended to cultivate a common interpretive framework and reduce variances (Malau-Aduli *et al.*, 2023).

Even with the increase in raters to five, a single testing occasion proved inadequate for achieving acceptable reliability. This aligns with the G-theory findings, demonstrating that occasions and the interactions between occasions and individuals accounted for over 36% of the variance. This reflects the variation in student performance over time and the varying difficulty of different case scenarios, which may differ according to each student's strengths. Therefore, increasing the number of occasions would enhance the reliability of the evaluation. Similarly, previous research focusing on pharmacology Objective Structured Clinical Examinations (OSCE) stations indicated that seven stations per week are required to achieve good reliability, underscoring the importance of multiple occasions for reaching reliable evaluations (Peeters *et al.*, 2021). Moreover, multiple assessment occasions have improved summative assessment outcomes and facilitated student knowledge retention (Sottiyotin *et al.*, 2023). Therefore, it is recommended that prescription writing assessments be administered on at least two occasions.

Presently, prescribing skills are predominantly evaluated through written examinations with single best answer formats, including OSCEs and Workplace-Based Assessments (WBAs). Nevertheless, these assessments have limitations; they do not fully examine the act of writing a prescription but test the ability to choose the correct prescription from multiple options (Ross & Loke, 2009; Rothwell *et al.*, 2012). Moreover, teaching approaches, such as case-based, patient-based, tutorial-based, and mixed-methods teaching, have been adopted to impart prescribing skills. Despite the heterogeneous nature of these study designs, evidence suggests the superiority of additional educational interventions in prescription writing over no specified supplementary instruction to enhance prescribing skills (Mokrzecki *et al.*, 2023). Nonetheless, a gap remains in the formal introduction and standardised assessment of prescribing skills within medical education.

The current study presents the development of a rubric and description of a programme for assessing prescription writing among pre-clinical medical students. Additionally, it provides significant insights into employing G-theory to determine the requisite number of occasions and raters needed to achieve a reliable assessment. The findings suggest that at least two testing occasions with two raters are recommended for a robust evaluation. However, three testing occasions are necessary if only one rater is available. The rubrics used in this study focus not only on the cognitive dimensions of drug prescribing but also on the professional aspects of the students' prescribing abilities, offering a comprehensive tool to assess the entirety of prescription writing skills in a pre-clinical educational setting. Moreover, the rubrics' interpretation should be assessed as a whole and in each item to give feedback to the students according to their prescription errors.

Limitations

The current study is subject to certain limitations. The sample size consists of volunteers enrolled in the extracurricular course, which may limit internal validity. However, the sample size is adequate, as calculated for generalisability theory analysis and to gather the validity evidence in developing prescription writing. Due to the context-sensitive nature of G-theory, these results might not be generalisable across different educational environments, clinical scenarios and raters (Bloch & Norman, 2012). Hence, external validation should be pursued to assess the broader applicability of these findings in varied educational settings across different academic levels, clinical environments, and diverse cultural contexts. Additionally, other potential facets, such as the number of items, were not included. This decision was made because the rubric content, developed considering potential prescription errors, was validated by experts, making item reduction challenging. Furthermore, this analysis indicated that deleting any item would diminish the internal consistency reliability, as evidenced by the item-test correlation and Cronbach's alpha.

The current study does not include the relations with other variables and consequences elements of the Messick validation framework, as this analysis was limited to student scores and did not extend to evaluating students' prescription performance in their clinical years. Therefore, further research focusing on blueprinting, standard setting, consequences, quality control, prediction of later performance, and the relationship to other measures of prescription writing could prove beneficial. Finally, the rubrics in this study are designed for prescription writing, not electronic prescribing, because medical students may be allocated to community hospitals for their internships, where handwritten prescriptions are still commonly used (Cassidy *et al.*, 2023). In addition, although electronic prescribing can reduce prescription errors, particularly those related to illegible handwriting, issues such as incorrect drug information entry may persist (Odukoya *et al.*, 2014). It remains crucial that prescriptions clearly specify the correct patient, medication, dose, duration, and route; failing to do so can potentially harm patients and, in serious cases, be fatal. Therefore, the current rubrics may still be adapted for learning electronic prescribing in the future (Cassidy *et al.*, 2023).

Conclusion

The study presents a comprehensive rubric and description of a programme that assesses each domain of prescription writing, assists in capturing the potential prescribing errors among medical students, and demonstrates its validity and reliability using the Messick validity framework. Utilising Generalisability Theory, it was determined that the variance primarily originated from the students' performance, followed by differences in occasions. A minimum of two assessment occasions with two raters or three with one rater is recommended to ensure acceptable reliability. Multiple assessment occasions may also enhance knowledge retention and future performance. holding regular meetings to encourage peer-to-peer interactions among assessors may enhance the dependability of evaluations.

Conflict of interest

The authors declare no conflict of interest.

Source of funding

The authors did not receive any funding.

References

Andersen, S. A. W., Nayahangan, L. J., Park, Y. S., & Konge, L. (2021). Use of generalizability theory for exploring reliability of and sources of variance in assessment of technical skills: A systematic review and meta-analysis. *Academic Medicine*, **96**(11), 1609–1619. https://doi.org/10.1097/ACM.0000000000004150

Aronson, J. K. (2012). Balanced prescribing – Principles and challenges. *British Journal of Clinical Pharmacology*, **74**(4), 566–572. https://doi.org/10.1111/j.1365-2125.2012.04413.x

Baranski, B., Bolt, J., Albers, L., Siddiqui, R., Bell, A., & Semchuk, W. (2017). Development of a documentation rubric and assessment of pharmacists' competency for documentation in the patient health record. *The Canadian Journal of Hospital Pharmacy*, **70**(6). https://doi.org/10.4212/cjhp.v70i6.1710

Bloch, R., & Norman, G. (2012). Generalizability theory for the perplexed: A practical introduction and guide: AMEE Guide No. 68. *Medical Teacher*, **34**(11), 960–992. https://doi.org/10.3109/0142159X.2012.703791

Brennan, R. L. (2010). Generalizability theory and classical test theory. *Applied Measurement in Education*, **24**(1), 1–21. https://doi.org/10.1080/08957347.2011.532417

Briesch, A. M., Swaminathan, H., Welsh, M., & Chafouleas, S. M. (2014). Generalizability theory: A practical guide to study design, implementation, and interpretation. *Journal of School Psychology*, **52**(1), 13–35. https://doi.org/10.1016/j.jsp.2013.11.008

Brinkman, D. J., Tichelaar, J., Graaf, S., Otten, R. H. J., Richir, M. C., & van Agtmael, M. A. (2018). Do final-year medical students have sufficient prescribing competencies? A systematic literature review. *British Journal of Clinical Pharmacology*, **84**(4), 615–635. https://doi.org/10.1111/bcp.13491

Cassidy, C. E., Boulos, L., McConnell, E., Barber, B., Delahunty-Pike, A., Bishop, A., Fatima, N., Higgins, A., Churchill, M., Lively, A., MacPhee, S. P., Misener, R. M., Sarty, R., Wells, R., & Curran, J. A. (2023). E-prescribing and medication safety in community settings: A rapid scoping review. *Exploratory Research in Clinical and Social Pharmacy*, **12**, 100365. https://doi.org/10.1016/j.rcsop.2023.100365

Donnon, T., McIlwrick, J., & Woloschuk, W. (2013). Investigating the reliability and validity of self and peer assessment to measure medical students' professional competencies. *Creative Education*, **04**(06), 23–28. https://doi.org/10.4236/ce.2013.46A005

Fanikos, J., Jenkins, K. L., Piazza, G., Connors, J., & Goldhaber, S. Z. (2014). Medication use evaluation: Pharmacist rubric for performance improvement. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, **34**(S1). https://doi.org/10.1002/phar.1506

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**(2), 175–191. https://doi.org/10.3758/BF03193146

Gautam, P. (2013). Minimizing medication errors: Moving attention from individual to system. *Journal of Anaesthesiology Clinical Pharmacology*, **29**(3), 293. https://doi.org/10.4103/0970-9185.117037

George, D., & Mallery, P. (2019). IBM SPSS statistics 26 step by step: A Simple Guide and Reference (16th ed.). Routledge. https://doi.org/10.4324/9780429056765

Gupta, M., Tikoo, D., & Pal, S. (2020). Assessment of Prescription writing skills and impact of an educational intervention on safe prescribing among the first-year postgraduate medical students of tertiary care hospital. *AMEI's Current Trends in Diagnosis & Treatment*, **4**(1), 51–59. https://doi.org/10.5005/jp-journals-10055-0087

Hamstra, S. J., & Yamazaki, K. (2021). A validity framework for effective analysis and interpretation of milestones data. *Journal of Graduate Medical Education*, **13**(2s), 75–80. https://doi.org/10.4300/JGME-D-20-01039.1

Heaton, A., Webb, D. J., & Maxwell, S. R. J. (2008). Undergraduate preparation for prescribing: The views of 2413 UK medical students and recent graduates. *British Journal of Clinical Pharmacology*, **66**(1), 128–134. https://doi.org/10.1111/j.1365-2125.2008.03197.x

Hill, J., Ogle, K., Gottlieb, M., Santen, S. A., & Artino, A. R. (2022). Educator's blueprint: A how-to guide for collecting validity evidence in survey- based research. *AEM Education and Training*, **6**(6). https://doi.org/10.1002/aet2.10835

James, H., Tayem, Y. I. Y., Al Khaja, K. A. J., Veeramuthu, S., & Sequeira, R. P. (2016). Prescription writing in small groups as a clinical pharmacology educational intervention: Perceptions of preclerkship medical students. *The Journal of Clinical Pharmacology*, **56**(8), 1028–1034. https://doi.org/10.1002/jcph.692

Keijsers, C. J. P. W., Leendertse, A. J., Faber, A., Brouwers, J. R. B. J., de Wildt, D. J., & Jansen, P. A. F. (2015). Pharmacists' and general practitioners' pharmacology knowledge and pharmacotherapy skills. *The Journal of Clinical Pharmacology*, **55**(8), 936–943. https://doi.org/10.1002/jcph.500

Khaja, K. A. J. Al, Handu, S. S., James, H., Mathur, V. S., & Sequeira, R. P. (2005). Assessing prescription writing skills of pre-clerkship medical students in a problem-based learning curriculum. *International Journal of Clinical Pharmacology and Therapeutics*, **43**(09), 429–435. https://doi.org/10.5414/CPP43429

Khodi, A. (2021). The affectability of writing assessment scores: a G-theory analysis of rater, task, and scoring method contribution. *Language Testing in Asia*, **11**(1), 30. https://doi.org/10.1186/s40468-021-00134-5

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, **15**(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Linton, K. D., & Murdoch-Eaton, D. (2020). Twelve tips for facilitating medical students prescribing learning on clinical placement. *Medical Teacher*, **42**(10), 1134–1139. https://doi.org/10.1080/0142159X.2020.1726309

Loskutova, N. Y., Lutgen, C., Smail, C., Staton, E. W., Nichols, S. D., & Pinckney, R. G. (2022). Stimulant prescribing error assessment rubric development. *Journal of Patient Safety*, **18**(1), e282–e289. https://doi.org/10.1097/PTS.0000000000000775

Malau-Aduli, B. S., Hays, R. B., D'Souza, K., Saad, S. L., Rienits, H., Celenza, A., & Murphy, R. (2023). Twelve tips for improving the quality of assessor judgements in senior medical student clinical assessments. *Medical Teacher*, **45**(11), 1228–1232. https://doi.org/10.1080/0142159X.2023.2216364

Martínez-Domínguez, J., Sierra-Martínez, O., Galindo-Fraga, A., Trejo-Mejía, J. A., Sánchez-Mendiola, M., Ochoa-Hein, E., Vázquez-Rivera, M., Gutiérrez-Cirlos, C., Naveja, J., & Martínez-González, A. (2022). Antibiotic prescription errors: the relationship with clinical competence in junior medical residents. *BMC Medical Education*, **22**(1), 456. https://doi.org/10.1186/s12909-022-03499-0

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, **50**(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Mokrzecki, S. M., Mallett, A., Sen Gupta, T., Perks, S., & Pain, T. (2023). Do educational interventions improve prescribing skills of medical students compared to no additional learning? A systematic review. *Medical Education Online*, **28**(1). https://doi.org/10.1080/10872981.2023.2259166

Odukoya, O. K., Stone, J. A., & Chui, M. A. (2014). E-prescribing errors in community pharmacies: Exploring consequences and contributing factors. *International Journal of Medical Informatics*, **83**(6), 427–437. https://doi.org/10.1016/j.ijmedinf.2014.02.004

Peeters, M. J., Cor, M. K., Petite, S. E., & Schroeder, M. N. (2021). Validation evidence using generalizability theory for an objective structured clinical examination. *INNOVATIONS in Pharmacy*, **12**(1), 15. https://doi.org/10.24926/iip.v12i1.2110

Ross, S., & Loke, Y. K. (2009). Do educational interventions improve prescribing by medical students and junior doctors? A systematic review. *British Journal of Clinical Pharmacology*, **67**(6), 662–670. https://doi.org/10.1111/j.1365-2125.2009.03395.x

Rothwell, C., Burford, B., Morrison, J., Morrow, G., Allen, M., Davies, C., Baldauf, B., Spencer, J., Johnson, N., Peile, E., & Illing, J. (2012). Junior doctors prescribing: Enhancing their learning in practice. *British Journal of Clinical Pharmacology*, **73**(2), 194–202. https://doi.org/10.1111/j.1365-2125.2011.04061.x

Ryan, C., Ross, S., Davey, P., Duncan, E. M., Francis, J. J., Fielding, S., Johnston, M., Ker, J., Lee, A. J., MacLeod, M. J., Maxwell, S., McKay, G. A., McLay, J. S., Webb, D. J., & Bond, C. (2014). Prevalence and causes of prescribing errors: The Prescribing Outcomes for Trainee Doctors Engaged in Clinical Training (PROTECT) Study. *PLoS ONE*, **9**(1), e79802. https://doi.org/10.1371/journal.pone.0079802

Salmasi, S., Khan, T. M., Hong, Y. H., Ming, L. C., & Wong, T. W. (2015). Medication errors in the Southeast Asian countries: A systematic review. *PLOS ONE*, **10**(9), e0136545. https://doi.org/10.1371/journal.pone.0136545

Sam, A. H., Fung, C. Y., Wilson, R. K., Peleva, E., Kluth, D. C., Lupton, M., Owen, D. R., Melville, C. R., & Meeran, K. (2019). Using prescribing very short answer questions to identify sources of medication errors: A prospective study in two UK medical schools. *BMJ Open*, **9**(7), e028863. https://doi.org/10.1136/bmjopen-2018-028863

Sequeira, R. (2015). Patient safety in medical education: Medication safety perspectives. *Indian Journal of Pharmacology*, **47**(2), 135. https://doi.org/10.4103/0253-7613.153417

Sottiyotin, T., Uitrakul, S., Sakdiset, P., Sukkarn, B., Sangfai, T., Chuaboon, L., & Palee, P. (2023). Effective formative assessment for pharmacy students in Thailand: Lesson learns from a school of pharmacy in Thailand. *BMC Medical Education*, **23**(1), 300. https://doi.org/10.1186/s12909-023-04232-1

Thenrajan, P., & Murugan, Pr. (2016). Impact of patient-based teaching in improving prescription writing skills of II MBBS students. *International Journal of Applied and Basic Medical Research*, **6**(3), 174. https://doi.org/10.4103/2229-516X.186954

Wiernik, P. H. (2015). A dangerous lack of pharmacology education in medical and nursing schools: A policy statement from the American College of Clinical Pharmacology. *The Journal of Clinical Pharmacology*, **55**(9), 953–954. https://doi.org/10.1002/jcph.539

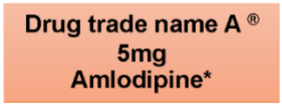
Yaman, M. N., Zamzam, R., Mohamad, N., Besar, M. N. A., & Kamarudin, M. A. (2012). Evaluation of case write-up: Assessment of prescription writing skills of fifth year medical students at UKM medical centre. *Procedia - Social and Behavioral Sciences*, **60**, 249–253. https://doi.org/10.1016/j.sbspro.2012.09.375

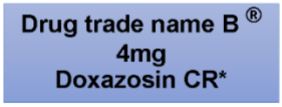
Appendix A1: Example of a case scenario with answer sheets for prescription writing among pre-clinical students. (Question 1)

* The pictures typically feature real drugs used in Thailand

Cardiovascular System	03:00	CASE NO.1

Mr. Jam Butter, a 65-year-old patient with hospital number 95/2558 has underlying conditions of hypertension and benign prostatic hyperplasia. Physical examination revealed a blood pressure of 160/90 mmHg and a pulse rate of 85 beats per minute. The doctor prescribed the medications, as shown in the picture, for a three-month period and scheduled a follow-up appointment.





Take one tablet once a day

Take one tablet once a day

Date:	Name: HN/AN:	Drug Allergy:

R

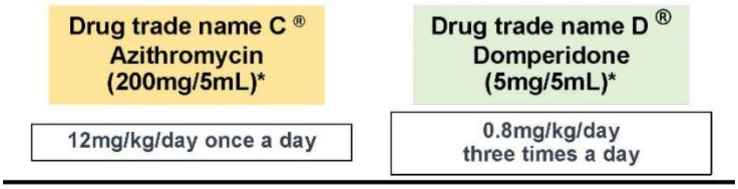
Prescriber information:

Appendix A2: Example of a case scenario with answer sheets for prescription writing among pre-clinical students. (Question 2)

* The pictures typically feature real drugs used in Thailand

Respiratory System	03:00	CASE NO.2

A 5-year-old boy named Black Angus with hospital number 44/2565 has a history of penicillin allergy. Presented with fever, cough and vomiting for three days. Physical examination: Body weight: 12kg, Pale nasal septum, injected tonsil with exudate. The doctor diagnosed exudative tonsilitis and prescribed the medications, as shown in the picture, for a seven-day period.



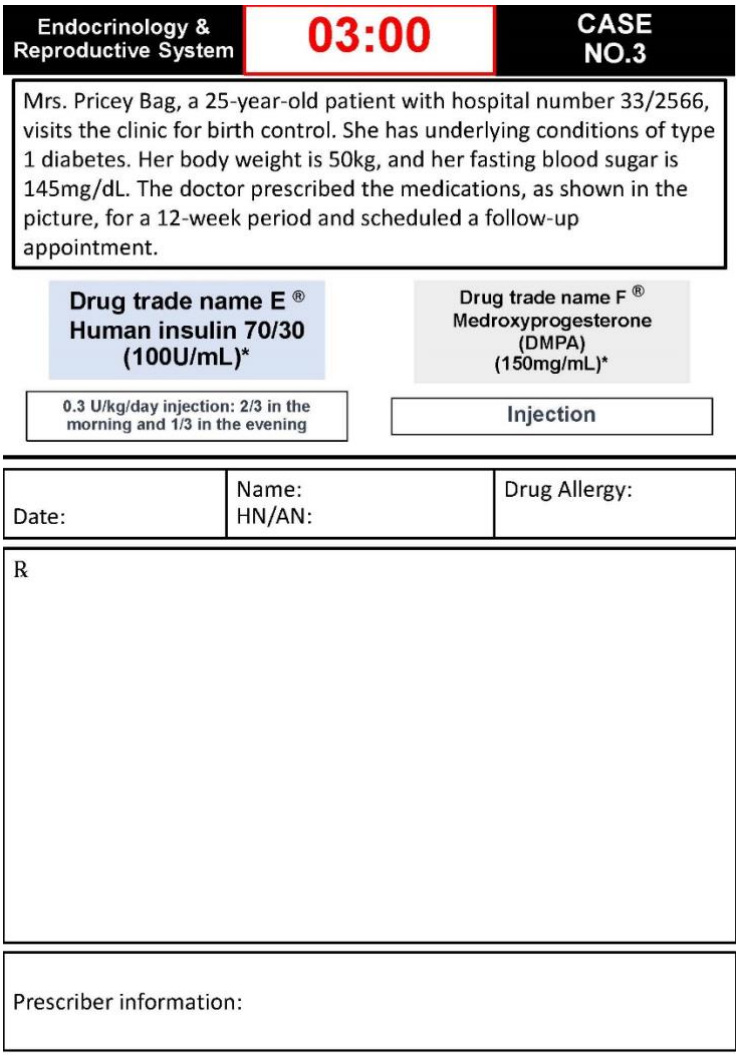
Date:	Name: HN/AN:	Drug Allergy:

℞

Prescriber information:

Appendix A3: Example of a case scenario with answer sheets for prescription writing among pre-clinical students. (Question 3)

* The pictures typically feature real drugs used in Thailand



Endocrinology & Reproductive System	03:00	CASE NO.3

Mrs. Pricey Bag, a 25-year-old patient with hospital number 33/2566, visits the clinic for birth control. She has underlying conditions of type 1 diabetes. Her body weight is 50kg, and her fasting blood sugar is 145mg/dL. The doctor prescribed the medications, as shown in the picture, for a 12-week period and scheduled a follow-up appointment.

Drug trade name E ® Human insulin 70/30 (100U/mL)*	Drug trade name F ® Medroxyprogesterone (DMPA) (150mg/mL)*
0.3 U/kg/day injection: 2/3 in the morning and 1/3 in the evening	Injection

Date:	Name: HN/AN:	Drug Allergy:

℞

Prescriber information:

Appendix B: Essential or standard learning points for each case scenario

Prescription writing skills	Case No.1	Case No.2	Case No.3
Patient information			
1. Personal history	○	○	○
2. Drug allergy history	○	●	○
Prescription writing for the inscription section			
3. Drug name	●	○	○
4. Drug form	●	●	●
5. Drug strength	○	○	●
6. Calculation of drug quantity	○	●	○
Prescription writing for the transcription section			
7. Calculated the drug amount used for each dose	○	●	●
8. Dosage regimen	●	○	●
Prescriber information			
9. Prescriber information	○	○	○
The cleanliness of the prescription form			
10. The correction of the prescription	○	○	○

● = Essential learning point, ○ = Standard learning point

Appendix C: Prescription writing rubric criteria for pre-clinical student (Total score: 30)

Prescription writing skills	3 marks	2 marks	1 mark
Patient information			
1. Personal history	Wrote the patient's full name, outpatient identification number, and the date the patient received the medication completely	Wrote the patient's full name, but other details are missing	The patient's full name was not written, and all other details are missing
2. Drug allergy history	Specified drug allergy history and the name of the drug(s) allergic to	Specified drug allergy history but do not specify the name of the drug(s) allergic to	Did not specified drug allergy history
Prescription writing for the inscription section			
3. Drug name	Wrote the generic drug name correctly	Wrote the trade drug name	Did not specified the drug name
4. Drug form	Specified the drug form correctly according to the standard	Specified the drug form, but not correctly according to the standard	Did not specified the drug form
5. Drug strength	Specified the drug strength correctly according to the standard	Specified the drug strength, but not correctly according to the standard	Did not specified the drug strength
6. Calculation of drug quantity	Calculated the drug quantity accurately and sufficiently for the entire duration of the treatment	Calculated the drug quantity correctly but not in sufficient amounts for the entire duration of treatment	Calculated the drug quantity incorrectly
Prescription writing for the transcription section			
7. Calculated the drug amount used for each dose	Calculated the drug amount used for each dose correctly	Calculated the drug amount for each dose, but the quantity was either insufficient or excessive	Unable to calculate the drug amount for each dose
8. Dosage regimen	Identified the correct method of dosage regimen for both the frequency and the time intervals	Correctly identified only one aspect of the dosage regimen method (either the frequency or the time intervals)	Incorrectly specified the dosage regimen method in both frequency and time intervals
Prescriber information			
9. Prescriber information	Prescriber signature with physician identification number provided	Prescriber signature without physician identification number provided	Prescriber signature and physician identification number are missing
The cleanliness of the prescription form			
10. The correction of the prescription	No corrections are made to the prescription or corrections made appropriately with a signature for verification	Corrections have been made to the prescription appropriately but without a signature for verification	Prescription corrections have been made incorrectly and do not follow the proper format

Appendix D: Decision study of p×o×r design for drug prescription writing

Estimate variance components in decision study																										
Effect	n_r'	1	1	1	1	1	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4	5	5	5	5	5
	n_o'	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
σ_p^2		6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945	6.945
σ_o^2		1.102	0.551	0.367	0.275	0.220	1.102	0.551	0.367	0.275	0.220	1.102	0.551	0.367	0.275	0.220	1.102	0.551	0.367	0.275	0.220	1.102	0.551	0.367	0.275	0.220
σ_r^2		0.525	0.525	0.525	0.525	0.525	0.262	0.262	0.262	0.262	0.262	0.175	0.175	0.175	0.175	0.175	0.131	0.131	0.131	0.131	0.131	0.105	0.105	0.105	0.105	0.105
σ_{po}^2		3.486	1.743	1.162	0.871	0.697	3.486	1.743	1.162	0.871	0.697	3.486	1.743	1.162	0.871	0.697	3.486	1.743	1.162	0.871	0.697	3.486	1.743	1.162	0.871	0.697
σ_{pr}^2		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
σ_{or}^2		0.046	0.023	0.015	0.011	0.009	0.023	0.011	0.008	0.006	0.005	0.015	0.008	0.005	0.004	0.003	0.011	0.006	0.004	0.003	0.002	0.009	0.005	0.003	0.002	0.002
σ_{por}^2		0.542	0.271	0.181	0.135	0.108	0.271	0.135	0.090	0.068	0.054	0.181	0.090	0.060	0.045	0.036	0.135	0.068	0.045	0.034	0.027	0.108	0.054	0.036	0.027	0.022
$\hat{\sigma}^2_\delta$		4.027	2.014	1.342	1.007	0.805	3.757	1.878	1.252	0.939	0.751	3.666	1.833	1.222	0.917	0.733	3.621	1.811	1.207	0.905	0.724	3.594	1.797	1.198	0.899	0.719
$\hat{\sigma}^2_\Delta$		5.700	3.112	2.250	1.819	1.560	5.143	2.703	1.889	1.483	1.239	4.958	2.567	1.769	1.371	1.132	4.865	2.498	1.709	1.315	1.078	4.810	2.457	1.673	1.281	1.046
$E\rho^2$		0.633	0.775	0.838	0.873	0.896	0.649	0.787	0.847	0.881	0.902	0.655	0.791	0.850	0.883	0.905	0.657	0.793	0.852	0.885	0.906	0.659	0.794	0.853	0.885	0.906
Φ		0.549	0.691	0.755	0.792	0.817	0.575	0.720	0.786	0.824	0.849	0.583	0.730	0.797	0.835	0.860	0.588	0.735	0.803	0.841	0.866	0.591	0.739	0.806	0.844	0.869