



# Towards a “prescribing license” for medical students: development and quality evaluation of an assessment for safe prescribing

David R. M. Jansen<sup>1</sup> · Carolina J. P. W. Keijsers<sup>2</sup> · Michiel O. Kornelissen<sup>3</sup> · Marcel G. M. Olde Rikkert<sup>1</sup> · Cornelis Kramers<sup>4</sup> · (on behalf of the education working group of the Dutch Society for Clinical Pharmacology and Biopharmacy)

Received: 1 January 2019 / Accepted: 22 April 2019 / Published online: 18 May 2019

© The Author(s) 2019

## Abstract

**Purpose** This report describes the development and validation process of an assessment with national consensus in appropriate and safe pharmacotherapy.

**Methods** A question-database on safe prescription based on literature of pharmacotherapy-related harm was developed by an expert group from Dutch medical faculties. Final-year medical students concluded a 2-year education program on appropriate and safe prescription by one of nine assessment variants of 40 multiple-choice questions each. An expert panel of professionals ( $n = 10$ ) answered all database questions and rated questions on relevance. Questions were selected for revision based on lack of relevance or poor test and item characteristics.

**Results** A total of 576 final-year medical students of the Radboud University was assessed. There was no significant difference in performance between students and content expert group ( $p = 0.7$ ), probably due to learning behavior. Out of 165 questions, 59 were selected for revision.

**Conclusion** Joint national effort from a team of experts in prescription and pharmacotherapy is an appropriate way to achieve a valid and reliable last-year student drug prescription assessment.

**Keywords** Prescribing · Medication safety · Education

---

Carolina J. P. W. Keijsers and Michiel O. Kornelissen contributed equally to this work.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00228-019-02686-1>) contains supplementary material, which is available to authorized users.

---

✉ David R. M. Jansen  
david.jansen@radboudumc.nl

<sup>1</sup> Department of Geriatric Medicine, Radboud University Medical Center, P.O. box 9101, 6500 HB Nijmegen, The Netherlands

<sup>2</sup> Department of Geriatric Medicine, Jeroen Bosch Hospital, Den Bosch, The Netherlands

<sup>3</sup> Radboudumc Health Academy, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>4</sup> Department of Pharmacology and Toxicology and Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands

## Introduction

Teaching junior doctors to be prepared to write appropriate and safe prescriptions requires integration of knowledge on drug selection, dosage, side effects, interactions with other drugs, co-morbidities, and individual (genetic) variability in pharmacokinetics. Many studies report a high rate of inappropriate prescribing as a potential cause of preventable morbidity and mortality [1, 2]. In this light, it is relevant to note that for the hospital setting, most drugs are prescribed by junior doctors with limited clinical experience [3]. For example, in the United Kingdom, half of the prescriptions were made by foundation Y1 doctors, so the doctors with the least experience [3]. Although the cause of a prescribing error is probably multifactorial, and related to for example time pressure or interruptions, it is also shown that junior doctors tend to copy the drug treatment choices of their supervisors instead of basing their choices on their own independent analysis of the [4]. One way or the other, it is prerequisite that junior doctors are well prepared for the high-risk task of prescribing, before

entering the ward. Therefore, sufficient education on appropriate and safe prescribing is conditional and an assessment on whether this education had led to sufficient knowledge could be recommended. On top of that, it is shown that assessments drive the learning of the students [5]. Unfortunately, appropriate learning outcomes for such an education program are lacking in many (European) countries [6, 7].

It is known that the majority of adverse events are caused by a relatively small group of drugs, i.e., pain medication (non-steroidal anti-inflammatory drugs (NSAIDs) and opioids), (combination of) antithrombotics, antibiotics, cardiovascular agents, and drugs which are renally excreted [8–12]. Experts from seven out of eight Dutch University Medical Centers, brought together via the Dutch Society for Clinical Pharmacology and Biopharmacy (further referred to as ‘expert working group’), reached consensus on the learning goals as shown in Table 1, based on the drugs causing the most drug-related harm in clinical practice.

For each of these drugs, the expert working group decided specifically what learning goals should be mastered to ensure appropriate and safe prescribing. These included mechanism of action, the side effects that may harm the patient, risk factors for these side effects (both patient factors and co-medication), measures to prevent these side effects, and how to act if these side effects occur. Since drug-related harm was the starting point, the learning goals concern mainly safety. However, the most important indications of the drug causing drug-related harm had to be mastered. For antibacterial drugs, the learning goals mainly cover efficacy and not safety. The expert working group has taken the initiative to develop an educational program in pharmacotherapeutics which is followed by a formal assessment of students’ comprehension and application of knowledge on safe prescribing. On the Blooms taxonomy on cognitive domains, learning domains ranging from recall of information to applying and synthesis of information should be incorporated in this educational program and assessment on safe pharmacotherapy [13].

The educational program consists of a reader with literature review, practice assessments, and interactive lectures, which can be seen at <https://www.youtube.com/user/fteindtoets/> (in Dutch). For better understanding of the scope of the reader, the learning goals for the domain of anticoagulants is added as an e-Appendix.

After the educational program, students could obtain a “medical prescribing license” by passing the formal assessment [14], which is part of the final-year medical examination program. By passing this assessment, the students demonstrate a certain basic level of mandatory knowledge for appropriate and safe prescribing analogous to the driving license theory test.

It is the ultimate vision of the expert working group to develop a compulsory assessment on all medical faculties in the Netherlands, so that medical students cannot graduate if they do not control both the knowledge and the skills needed for appropriate and safe drug prescription. Therefore, we describe in this paper the development and validation of a formal assessment in appropriate and safe pharmacotherapy by joint national effort and report its external validation process by using an independent group of experts for assessing content validity (further referred to as “content expert panel”).

## Methods

This study was designed to develop and study the quality of a newly developed 40 questions assessment on pharmacotherapy knowledge for last-year medical students.

### Development of the assessment: content

First, national consensus was achieved on the knowledge base required for the medical students by an expert working group ( $n = 7$ ), consisting of representatives from seven out of eight Dutch medical faculties [14]. All of the members of the expert

**Table 1** Consensus-based domains for pharmacology training in last-year medical students across The Netherlands

	Domains	Subjects per drug-related group
Drug-related topics	Analgesics	1. Medications and mechanisms of action
	Anticoagulants	2. Main indications
	Cardiovascular agents	3. Relevant kinetic data
	Drugs for diabetes mellitus	4. Main problems or side effects: including mechanism of action and main clinical presentation per agent
	Antidepressants	5. Patients who are most at risk per problem or side effect
	Benzodiazepines	6. Key interactions and mechanisms contributing to problem or side effect
	Antibiotics	7. Measures to prevent problems or side effects
		8. Measures to take if a problem arises or side effect occurs
General knowledge	Basic pharmacokinetics	
	Drug allergies	
	Laws and regulations	
	Appropriate drug use	

working group were physician or pharmacist, all registered clinical pharmacologists.

In order to design the curriculum to prepare the students for applying the underlying knowledge in this assessment, a stepped approach was chosen: (1) identify the drug groups which cause most harm, (2) identify the learning goals per group, and (3) identify additional important knowledge to be assessed.

For the first step, the expert working group identified the drug groups known to cause the majority of preventable serious adverse reaction based on the Dutch HARM-Wrestling Task Force targeting outpatient drug safety [15]. The following drug groups were identified and clustered by indication into seven drug-related domains, namely, (1) analgesics, (2) anticoagulants, (3) cardiovascular agents, (4) drugs for diabetes mellitus, (5) antidepressants, and (6) benzodiazepines. By consensus, the expert working group added (7) antibiotics as domain. Antibiotics are widely used and, in the opinion of the expert working group, knowledge about good antibiotic practice is relevant because injudicious use may not only cause harm in individual patients but may also induce antimicrobial resistance at the societal level.

Second, the expert working group systematically assessed the following categories to identify learning goals connected to these drug groups: (1) medications and mechanisms of action, (2) main indications, (3) relevant kinetic data, (4) main problems or side effects: including mechanism of action and main clinical presentation per agent, (5) patients who are most at risk per problem or side effect, (6) key interactions and mechanisms contributing to problem or side effect, (7) measures to prevent problems or side effects, and (8) measures to take if a problem arises or side effect occurs.

Additionally, general knowledge domains concerning related general pharmacotherapy were defined by the expert working group, based on expert opinion and added to the previously defined seven drug groups, i.e., (1) basic pharmacokinetics including simple dosage calculations; (2) drug allergies; (3) laws and regulations including medicine and

driving, and drug prescribing; and (4) appropriate drug use consisting of topics like the WHO-six-step method [16], pharmacotherapy in case of pregnancy, and lactation and transfer of information to pharmacist and primary care physician.

For the final step, Table 1 shows the domains and subjects per drug group as a result of expert consensus. Table 2 shows the test matrix of the assessment based on these domains. After three steps of content development, a total of 11 assessment domains were identified by expert consensus: core prescription knowledge on seven drug groups and on four related general pharmacotherapy domains. For details about drugs to be studied by the students per domain, see e-Appendix.

### Development of the assessment: from questions to assessment

The expert working group created a database of multiple-choice questions covering basic and applied knowledge of the selected drugs and general topics. All experts contributed questions which were subsequently reviewed by experts from two other medical faculties. Second, all questions were appraised by an assessment expert which ultimately resulted in a database with 165 peer-reviewed and approved questions.

Next, from this database, assessments were drawn consisting of 40 questions each, covering the previously mentioned 11 domains with 2 to 6 questions per domain. In Table 1, the number of questions per category in each assessment is shown. These assessments were checked for duplicate questions and contamination between the questions, e.g., if one question could be helpful to answer another question. If needed, small adjustments were made to the assessment and questions were replaced by other questions. From the database, nine different assessments were drawn, which were alternately used for the test taking.

### Participants and data collection procedure

In the period of March 2015 to December 2016, the assessment took place monthly in groups of on average 30 final-year medical students from the Radboud University Faculty of Medical Sciences. In total, 576 students were assessed, during 1-h sessions. The assessments consisted of 40 multiple-choice questions per assessment with on average four (range: two to five) alternatives resulting in at least 150 answer options per assessment, which is considered as lower limit for reliable psychometric analysis.

The assessment was summative as part of their regular curriculum; students had to pass it to become a physician. Those who failed the assessment had to repeat the assessment until they acquired a sufficient score. The pass rate was set at 80% correct answers after correction for guessing.

To study concurrent validity, a content expert panel, consisting of physicians with regular prescribing activities

**Table 2** Test matrix

	Domains	Questions ( <i>n</i> )
Drug-related topics	Analgesics	6
	Anticoagulants	6
	Cardiovascular agents	6
	Drugs for diabetes mellitus	4
	Antidepressants	4
	Benzodiazepines	2
	Antibiotics	4
General knowledge	Basic pharmacokinetics	2
	Drug allergies	2
	Laws and regulations	2
	Appropriate drug use	2
		40

and pharmacists, all registered clinical pharmacologists, was composed. The experts were approached by e-mail in October 2016 with the request to participate. The goal of this content expert panel was twofold: first to verify whether these experts considered the questions a precondition to be able to perform safe pharmacotherapy in clinical practice and second for external validation by comparison of the students' results with the experts' scores. The content experts examined the full set of 165 questions. All questions were answered and rated on relevance by choosing "essential," "useful, but not essential," or "not necessary" according to Lawshe [17]. Additionally, there was an option for the expert to note comments per question. The mean values per assessment for the experts were reduced from the total set of 165 questions.

The answers of the students, the answers of the experts, and experts' relevance scores and comments were anonymously collected and analyzed.

## Data analyses

### Construct validity

In order to study construct validity of the assessment as a whole, the following analyses were performed:

Concurrent validity was studied by comparing the means of the medical students with the experts from the content expert group by unpaired *t* tests for the overall comparison.

Content validity was tested according to Lawshe [17]; all experts rated all questions as "relevant" (value 1) or either one of the other two options: useful, but not essential, or not necessary, and in that case scored the item as "not relevant": value -1. If the mean score was positive, an item was scored as "precondition to be able to perform safe pharmacotherapy". Content validity was tested with a focus on the most frequent occurring drug-related problems, by which the larger part of safe prescribing was tested.

Criteria for content validity, discriminant analysis and test characteristics were combined to evaluate the appropriateness of all items separately (see below).

### Reliability

Different aspects of reliability were analyzed, namely, the internal consistence of the assessment as a whole and several parameters per question. To start with, the internal consistency was measured by Cronbach's alpha [18] per assessment, which should be 0.50 or higher, as proposed for multidimensional assessments and relatively short test lengths [19, 20]. Internal consistency was only analyzed in the student group because the internal consistency differs per test population, and the students were the population aimed for. Next, for all questions, discriminant analysis was performed per item calculating the "Difficulty Index" with correction for guessing on

multiple choice; probability ( $p'$ ), reflecting the percentage of students who answered the item correctly, including correction for guessing ( $p' = p - (1 - p) / (\text{number of options} - 1)$ ). The scores were interpreted as follows: > 0.9: easy question, 0.9 to 0.5: medium difficulty, < 0.50: difficult question for three-option multiple-choice questions, and for four-option multiple-choice questions: > 0.9, between 0.9 and 0.44, and < 0.44, respectively to correct for chance. Next, item-rest correlation ( $R_{ir}$ ) was calculated, which is the correlation between the question score and the overall assessment score (excluding that question), referring to how well a question differentiates between participants who master the material and those that do not know) [21].  $R_{ir}$  scores were categorized as: < 0.19: poor, 0.20–0.29: adequate, 0.30–0.39: good, and > 0.40: very good) [22, 23]. Standard deviation was calculated for all assessments for students and content expert panel as a group.

For adequate interpretation of the quality of the questions, each question should preferably at least be used 100 times [23]. To accomplish this, we reduced the amount of assessment variants to the first four (of a total of nine). The other test variants were used less often, and so, for the questions in these variants, this criterion was not met.

### Selection procedure of potentially inappropriate items

Finally, potentially inappropriate questions were selected for revision. A question was considered potentially inappropriate if at least one of the following criteria was met; (1) difficulty index ( $p'$  value) < 0.8 for experts, (2) rated "irrelevant" by a majority of the content expert panel, or (3)  $p$  value < 0.65 for students and item-rest-correlation  $\leq 0$ .

All assessments were entered in the online assessment tool TestVision (Teelen Kennismanagement, The Netherlands). Psychometric analysis was performed with Microsoft Excel 2007 (Microsoft) and IBM SPSS 22.0 (IBM). Statistical significance was set at  $p < 0.05$ .

## Results

From March 2015 to December 2016, 576 unique students (66% female) were assessed, as well as 10 independent experts from the content expert panel (40% female). The students completed a total of 673 assessments.

Table 3 shows the main results. The vast majority (87.3%) of the students passed the assessment at the first attempt. A total of 73 students (12.7%) had to retake the assessment (8.9% once, 3.5% twice). The mean score of the students was 90.5% and of the experts 90.7%.

**Table 3** Validity and reliability of the nine assessments of 40 questions each for student and clinical pharmacology experts

Assessment (Q <sup>b</sup> = 40)	Groups							
	Medical students <i>n</i> = 576					Content expert panel <sup>a</sup> <i>n</i> = 10		
Tests completed	Mean score in % (SD <sup>c</sup> )	Mean <i>p</i> <sup>d</sup>	Item-rest correlation <sup>e</sup>	Cronbach's alpha <sup>f</sup>	Mean score In % (SD <sup>c</sup> )	Mean <i>p</i> <sup>d</sup>	<i>p</i> value <i>t</i> test	
1	131	90.80 (6.85)	0.87	0.21	0.65	91.50 (6.03)	0.88	<i>p</i> = 0.66
2	101	87.13 (7.10)	0.81	0.18	0.60	88.25 (3.13)	0.83	<i>p</i> = 0.62
3	110	92.65 (5.70)	0.90	0.25	0.54	91.75 (6.68)	0.89	<i>p</i> = 0.64
4	100	91.70 (7.45)	0.89	0.23	0.71	94.25 (4.73)	0.92	<i>p</i> = 0.29
5	65	87.80 (10.01)	0.83	0.28	0.77	90.75 (3.13)	0.87	<i>p</i> = 0.37
6	73	90.85 (7.48)	0.87	0.26	0.72	88.25 (4.58)	0.83	<i>p</i> = 0.29
7	32	90.15 (7.03)	0.87	0.29	0.68	91.75 (5.78)	0.89	<i>p</i> = 0.52
8	29	94.48 (5.60)	0.92	0.20	0.61	92.50 (2.63)	0.90	<i>p</i> = 0.29
9	32	90.00 (6.25)	0.86	0.20	0.62	87.50 (5.00)	0.82	<i>p</i> = 0.26
Mean assessment result		90.45 (7.05)			0.66	90.73 (4.63)		<i>p</i> = 0.73

<sup>a</sup> A panel of physicians with regular prescribing activities and pharmacists, all registered clinical pharmacologists

<sup>b</sup> Number of questions per assessment

<sup>c</sup> Standard deviation

<sup>d</sup> *p*' Difficulty Index with correction for guessing on multiple-choice items

<sup>e</sup> Item-rest correlation: index of how well a question differentiates between participants who master the material and those that do not know. (range 0.20–0.29: adequate)

<sup>f</sup> Cronbach alpha: desired range was 0.50 or higher

## Validity

Three domains of validity were studied. First, the experts were compared with the students for concurrent validity. There was no significant difference between the students and the experts (mean score 36.18 out of 40, (SD 2.97) versus mean 36.29 (SD 2.03) ( $t(761) = -0.352$ ,  $p = 0.73$ ). Overall, the mean rating of the questions was positive meaning “essential” by a majority of the experts. Still, 24.4% ( $n = 40$ ) of the questions were considered “useful, but not essential,” or “not necessary.” All these 40 questions were marked as potentially inappropriate.

## Reliability

First, the analysis of complete assessments was considered. Internal consistence as measured by the Cronbach alpha (a measure of internal consistency) how closely related a set of items are as a group. It is considered to be a measure of scale reliability ranged from 0.54 to 0.77 and was in the range 0.5 or above per test, which is adequate to good.

Next, discriminant analyses on the level of the individual questions were performed. The majority of the questions (87.9% ( $n = 145$ )) turned out to be easy  $p' > 0.90$  based on students results. Ten questions (6.1%) were

medium and five items (3.0%) turned out to be a difficult question. The item-rest correlation was very good for 13.9% ( $n = 23$ ) of the questions, good for 16.4% ( $n = 27$ ) of the questions, adequate for 23.6% ( $n = 39$ ) of the questions, and poor for 43.0% ( $n = 71$ ) of the questions.

## Qualitative analysis of potentially inappropriate questions

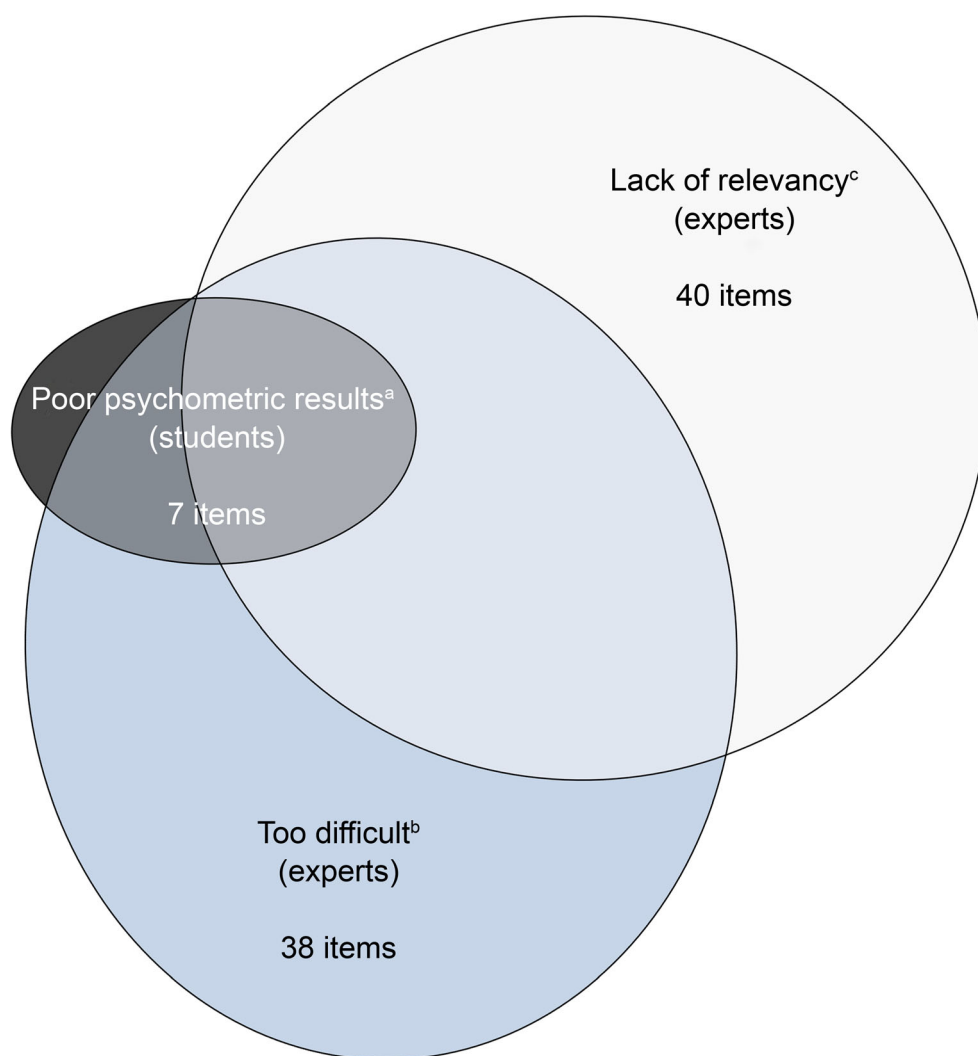
Figure 1 shows which questions were included for what reason in further adjustments based on potential inappropriateness. A total of 59 out of 165 questions (36%) were marked for revision based on at least one of the three defined criteria; 38 questions based on difficulty for the experts ( $p' < 0.8$ ), 40 questions based on lack of relevancy as rated by the majority of the experts, and seven questions based on student criterion. The vast majority of the questions was marked for revision based on the first two criteria. The third criterion only added one unique question to the set.

As a result of this discriminant analysis, all marked items were revised for future use by the expert working group.

## Discussion

This report describes the development of an assessment on pharmacotherapy knowledge for medical students. Based on

**Fig. 1** Diagram of items for revision ordered by selection criterion. Additional 52 items selected for revision based on expert criteria in addition to classic test analysis criteria' from students' results. <sup>a</sup>Poor psychometric results indicating that students who overall did poorly on the test did better on that item than did students who overall did well ( $p$  value  $< 0.65$  for students and item-rest-correlation  $\leq 0$ ). <sup>b</sup>Too high difficulty for the experts (difficulty index:  $p' < 0.8$ ), <sup>c</sup>Lack of relevancy as rated by the majority of the content expert panel



literature on frequently occurring drug-related problems, learning goals were set by a nationwide expert working group, resulting in 11 domains. Using a database of 165 questions, assessments consisting of 40 questions were tested in over 500 students and 10 content experts. It appeared that the overall performance of last-year medical students on this assessment on appropriate and safe prescription was comparable with the performance of experts in the field of appropriate and safe pharmacotherapy. The assessment had an acceptable internal consistence shown by a mean Cronbach alpha of 0.66. In the analyses of the individual questions by the experts of the content expert panel, still, many questions should be considered for revision namely 35.8%, with rather large overlap between the questions as considered non-relevant by the experts and those with low classic psychometric parameters based on expert data. This is an example of a stepwise approach of the development of a prescribing assessment.

A relevant finding is the fact that students score near the level of experts. Generally, summative assessments are designed to discriminate between those who do have enough knowledge

and those who do not. The caesura, or cut of point to pass, is adjusted depending on the mean score of the subjects tested to assure a weighted distribution of scores independent of the difficulty of the questions. In other words, a student has to be in the upper half of scores to pass. For this assessment, however, the national expert working group set a fixed relatively high caesura of 80%, to be sure the students demonstrated a basic standard of competence in safe pharmacotherapy, in analogy with the driving license theory test. Consequently, a fair amount of 87.9% of the questions was considered “easy” as measured by the easiness index and moderate item-rest correlation index. Also, the high caesura might have stimulated the external motivation to study hard on this topic, with high assessment scores as a result [24]. However, whether and how these knowledge levels may again start to differ over time during a medical career remains to be studied [25].

More initiatives are taken worldwide to improve education on safe pharmacotherapy and to assess pharmacotherapy knowledge or skills [26]. For example, recently, there was a report about a nationwide initiative for an online prescribing

assessment for final-year medical students in the UK, published by Maxwell et al. [27]. However, a widely accepted consensus is lacking on (1) the optimal strategy in education of future prescribers to reduce prescription errors in primary care and hospitals and (2) appropriate learning goals for this education program and concluding assessment. This study shows how a national collaboration can lead to a nationally accepted assessment, which will gradually be introduced in all medical faculties in the Netherlands.

Despite consensus in the expert working group who initially developed the assessment, it was surprising that there was a lack of consensus between the experts that designed the assessment and the content expert panel as used in this study for external reference. This content expert panel rated 24.2% of the individual questions as non-relevant though these questions were developed by expert colleagues. Since prescribing is a complex skill, this often results in more than one adequate prescription and in more than one correct answer for a clinical problem, depending on the circumstances. This lack of concordance between the expert working group and the content expert panel is a clear indication that context affects assessment.

The assessment in this study shows to be a valid and reliable assessment on pharmacotherapy knowledge for medical students. However, before nationwide application, some considerations should be taken into account. One could discuss whether the Cronbach's alpha is high enough. Although it is certainly acceptable, for an assessment with such a high caesura of 80%, a higher alpha may be required. The moderate alpha level may be explained by the fact that these assessments had (1) many knowledge domains, (2) relative limited number of items per assessment, and (3) a rather uniform performance of the group or ceiling effect due to learning behavior [20]. All three aspects however were chosen as a part of the design and therefore higher Cronbach's alpha will probably not be obtained as long as students intensively study for their assessment, which is aimed for. An alternative would be to extend the assessment by adding more questions [19]. Since the end of data inclusion for this study, the assessment is extended to 60 questions per assessment. Regarding the item-analyses, it would have been preferable to have had more assessments in the analysis. The student numbers however were set, and within the time frame of the study, more assessments were not possible.

It should be noted that the development of this assessment is the result of a nationwide Dutch collaboration. However, because of differences in laws, regulations, and guidelines, learning goals are not one-on-one transferable to other countries. Nevertheless, we believe that the procedure described here, with joint national effort from a team of experts in the field, may be feasible for implementation in other countries.

## Conclusions

This study describes a valid and reliable assessment, developed by national effort to assess final-year medical students' comprehension and application of knowledge on appropriate and safe prescribing. With a high cut-off point to pass, students seem to be very motivated to pass. After an education program on appropriate and safe prescribing, students achieved a score comparable to experts in the field. In analogy of the "driving licence theory test" an additional practical test "on the road" should be obligate: this should assess whether students can apply clinical pharmacology knowledge in clinical practice. Whether such an assessment ultimately improves safe prescription and overall patients' outcomes has yet to be investigated.

**Acknowledgements** The assessment was developed by the education working group of the Dutch Society for Clinical for Clinical Pharmacology and Biopharmacy: J. Tichelaar (VU University Medical Center, Amsterdam), B. Janssen (Maastricht University Medical Center), A. Maassen van den Brink and L. Peeters (Erasmus Medical Center, Rotterdam), G. Dumont (Amsterdam Medical Center, Amsterdam), M. Hessel and R. Rissmann (Leiden University Medical Center, Leiden), and W. Knol (Utrecht Medical Center, Utrecht).

**Author contributions** DJ, CKe, MK, MO, and CKr wrote the manuscript; DJ and CKr designed the research; DJ and MK analyzed the data.

## Compliance with ethical standards

The Dutch Ethics Review Board of Medical Education (NVMO-ERB) approved conduct of this study.

**Conflict of interest** The authors declare that there are no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Leendertse AJ et al (2008) Frequency of and risk factors for preventable medication-related hospital admissions in the Netherlands. *Arch Intern Med* 168(17):1890–1896
2. Miguel A, Azevedo LF, Araújo M, Pereira AC (2012) Frequency of adverse drug reactions in hospitalized patients: a systematic review and meta-analysis. *Pharmacoepidemiol Drug Saf* 21(11):1139–1154
3. Ryan C, Ross S, Davey P, Duncan EM, Francis JJ, Fielding S, Johnston M, Ker J, Lee AJ, MacLeod MJ, Maxwell S, McKay GA, McLay JS, Webb DJ, 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

4. Tichelaar J, Richir MC, Avis HJ, Scholten HJ, Antonini NF, de Vries TPGM (2010) Do medical students copy the drug treatment choices of their teachers or do they think for themselves? *Eur J Clin Pharmacol* 66(4):407–412
5. Wood T (2009) Assessment not only drives learning, it may also help learning. *Med Educ* 43(1):5–6
6. Brinkman DJ, Tichelaar J, Graaf S, Otten RHJ, Richir MC, van Agtmael MA (2018) Do final-year medical students have sufficient prescribing competencies? A systematic literature review. *Br J Clin Pharmacol* 84:615–635
7. Brinkman DJ, Tichelaar J, Schutte T, Benemei S, Böttiger Y, Chamontin B, Christiaens T, Likic R, Ma'ulaitis R, Marandi T, Monteiro EC, Papaioannidou P, Pers YM, Pontes C, Raskovic A, Regenthal R, Sanz EJ, Tamba BI, Wilson K, Vries TP, Richir MC, Agtmael MA, on behalf of the Working Group Research on CPT Education of the European Association for Clinical Pharmacology and Therapeutics (EACPT) (2017) Essential competencies in prescribing: a first European cross-sectional study among 895 final-year medical students. *Clin Pharmacol Ther* 101(2):281–289
8. Davies EC, Green CF, Taylor S, Williamson PR, Mottram DR, Pirmohamed M (2009) Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS One* 4(2):e4439
9. Hoonhout LH et al (2010) Nature, occurrence and consequences of medication-related adverse events during hospitalization: a retrospective chart review in the Netherlands. *Drug Saf* 33(10):853–864
10. Klopotoska JE, Wierenga PC, Stuijt CCM, Arisz L, Dijkgraaf MGW, Kuks PFM, Asscheman H, de Rooij SE, Lie-A-Huen L, Smorenburg SM, on behalf of the WINGS Study Group (2013) Adverse drug events in older hospitalized patients: results and reliability of a comprehensive and structured identification strategy. *PLoS One* 8(8):e71045
11. Krahenbuhl-Melcher A et al (2007) Drug-related problems in hospitals: a review of the recent literature. *Drug Saf* 30(5):379–407
12. Saedder EA, Brock B, Nielsen LP, Bonnerup DK, Lisby M (2014) Identifying high-risk medication: a systematic literature review. *Eur J Clin Pharmacol* 70(6):637–645
13. Kim MK, Patel RA, Uchizono JA, Beck L (2012) Incorporation of Bloom's taxonomy into multiple-choice examination questions for a pharmacotherapeutics course. *Am J Pharm Educ* 76(6):114
14. Kramers C, Janssen BJ, Knol W, Hessel MHM, Mulder WM, Dumont G, Maassen van den Brink A, Tichelaar J (2017) A licence to prescribe. *Br J Clin Pharmacol* 83(8):1860–1861
15. Warle-van Herwaarden MF et al (2012) Targeting outpatient drug safety: recommendations of the Dutch HARM-wrestling task force. *Drug Saf* 35(3):245–259
16. Vries TPGM de HR, Hogerzeil HV, Fresle DA (1994) Guide to good prescribing – a practical manual. World Health Organisation, Geneva
17. Lawshe CH (1975) A quantitative approach to content validity. *Pers Psychol* 28(4):563–575
18. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3):297–334
19. Schmitt N (1996) Uses and abuses of coefficient alpha. *Psychol Assess* 8(4):350–353
20. Tavakol M, Dennick R (2011) Making sense of Cronbach's alpha. *Int J Med Educ* 2:53–55
21. Cureton EE (1966) Corrected item-test correlations. *Psychometrika* 31(1):93–96
22. Mitra, NK, et al (2009) The levels Of difficulty and discrimination indices in type a multiple choice questions Of pre-clinical semester 1 multidisciplinary summative tests. *IeJSME* 3:2–7
23. Ebel RL, Frisbie DA (1991) Essentials of education measurement. Prentice Hall, Englewood Cliffs
24. Kusurkar RA, ten Cate TJ, Vos CMP, Westers P, Croiset G (2013) How motivation affects academic performance: a structural equation modeling analysis. *Adv Health Sci Educ Theory Pract* 18(1):57–69
25. Keijsers CJ, Ross S (2015) A pharmacological approach to education. *Br J Clin Pharmacol* 80(3):329–330
26. Bos JM, Bemt PMLA, Smet PAGM, Kramers C (2017) The effect of prescriber education on medication-related patient harm in the hospital: a systematic review. *Br J Clin Pharmacol* 83(5):953–961
27. Maxwell, SRJ, et al., Prescribing safety assessment 2016: delivery of a national prescribing assessment to 7343 UK final-year medical students. *Br J Clin Pharmacol* 2017. 83(10): 2249–2258

#### STUDY HIGHLIGHTS

What is the current knowledge on the topic?

- Inappropriate prescribing leads to medication-related morbidity and mortality
- Education is a possible tool to reduce medication-related morbidity and mortality

What question did this study address?

- What is a feasible method to identify learning goals and content development of an education program concluded by a pharmacotherapy assessment in order to improve safe prescribing by future doctors in primary care and hospitals?

What does this study add to our knowledge?

- Joint national effort from a team of experts in the field of prescription and pharmacotherapy is an appropriate way to achieve a valid and reliable assessment.
- By passing this assessment, students demonstrate to have knowledge which is a prerequisite for safe prescribing (at a level comparable to the performance of experts in the field).

How might this change clinical pharmacology or translational science?

- We describe a method which could be internationally used as a feasible starting point in the development of a “license to prescribe.”

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