Evaluation of Quality Indicators for Dutch Community Pharmacies Using a Comprehensive Assessment Framework

Tim W.A. Schoenmakers, PharmD; Martina Teichert, PharmD, PhD; Jozé Braspenning, PhD;
Lydia Vunderink, MSc; Peter A.G.M. De Smet, PharmD, PhD; and Michel Wensing, PhD

ABSTRACT

BACKGROUND: In 2008, the Dutch Health Care Transparency Programme (*Zichtbare Zorg*) was set up to develop and apply quality indicators (QIs) for health care. These QIs serve a range of purposes and can be categorized into those for internal use—for meeting quality standards and to continuously measure improvement (formative)—and external use—to enable patients and health insurance companies to distinguish between health care providers (summative). In order to assess the validity of QIs, a comprehensive Indicator Assessment Framework (IAF) was developed. This framework specifies the following criteria for validation: content validity, absence of selection bias, absence of measurement bias, and statistical reliability. Because of the intended summative use, the IAF was used for structural assessment of the QIs set for Dutch community pharmacists.

OBJECTIVE: To assess the validity of the current set of 52 QIs for community pharmacies using the IAF.

METHODS: An expert panel applied the IAF criteria to the set of QIs collected in 1,807 Dutch community pharmacies on their performance in 2011. The QIs were judged as meeting, partly meeting, or not meeting the requirements regarding these criteria. The judgments were evaluated for QI type (structure, process, or outcome) and for predefined domains.

RESULTS: Thirteen QIs (25%) were judged as meeting the requirements for all criteria. Among them were 12 structure indicators and 1 process indicator. For process indicators, the criterion for measurement bias poorly met the requirements, and content validity was unsatisfactory for outcome indicators. The 13 overall valid QIs covered 6 out of 10 predefined domains: continuity of care, clinical risk management, compounding, dispensing of medication, management, and quality management.

CONCLUSIONS: When subjecting the QI set for community pharmacies to the requirements of the IAF, only a quarter of the QIs met all requirements. To increase the number of valid process and outcome indicators, meaningful aspects for the outcome of pharmaceutical care have to be defined, and uniform measurement of relevant processes has to be implemented.

What is already known about this subject

- Quality indicators (QIs) play an increasing role in health care and serve a range of purposes, which can be broadly categorized as formative for internal audit and feedback and summative for external evaluation, benchmarking, pay for performance, and public reporting.
- QIs should be validated according to their purpose in order to provide useful and reliable information.
- Those QIs used for summative purposes require more rigorous validation than those used for formative purposes.

What this study adds

- When subjecting an existing QI set for community pharmacies to a structured assessment, only 25% of the QIs met all requirements.
- Our findings illustrate that development of a QI set should be a continuous structural process in order to obtain valid QIs for all relevant domains.
- Uniform pharmaceutical care process registrations are needed to increase the number of valid QIs.

In managed care, quality indicators (QIs) play an important role in monitoring and continuously improving health care.[1,2] In addition to formative use for internal audit and feedback,[3,4] QIs are increasingly used for summative purposes such as external evaluation,[5] benchmarking,[6] pay for performance,[7] and public reporting.[8] To warrant reliable information, QIs should be systematically developed and tested for validity, acceptability, feasibility, reliability, and sensitivity to change.[9] A practice test or a more rigorous evaluation (when used for pay for performance or public reporting) should be part of the development process.[8,10,11]

The development of QIs vary per country. In some countries, sets of indicators have been developed by specific agencies, such as the National Institute for Health and Clinical Excellence in the United Kingdom[10,12] and the Institute for Applied Improvement and Research in Health Care in Germany.[6,13] In the United States, QIs have been developed by medical societies or governmental agencies, such as the Agency for Healthcare Research and Quality, and have been endorsed by the National Quality Forum.[8,14] With knowledge of the key features of QI development in other Western countries, the Dutch Health Care Transparency Programme (*Zichtbare Zorg*) in 2008 started the development of quality indicators (QIs) and enhancement of their uptake in routine monitoring of the quality and outcomes of health care.[15]

To meet this ambitious aim for each health care profession (e.g., general practitioners, physiotherapists, dentists, midwives, and community pharmacists), task groups were formed consisting of the following representatives of stakeholders: (a) care providers who want to monitor and continuously improve the quality of their care; (b) organizations of patients and consumers, which require information for their members to

make reasonable choices for the most appropriate care; (c) the Healthcare Inspectorate, which needs information for early detection of patient safety risks and supervision of health care processes; and (d) health insurance companies using QIs for contracting the best care for their clients at reasonable costs. To evaluate the validity of the QIs, a comprehensive Indicator Assessment Framework (IAF) was developed.[16] Based on epidemiologic principles, the IAF provides 4 criteria for the assessment of QIs: content validity, absence of selection bias, absence of measurement bias, and statistical reliability.

Within this setting, community pharmacists collected data in 2012 for a set of 52 QIs concerning their performance for the year 2011. Two-thirds of this QI set originated from a former set, which was developed for formative purposes.[5] These QIs had been selected by relevance for quality of care improvement or risk of harm to patients. The additional QIs in the present set were selected by relevance to patients, particularly with regard to aspects of counseling and safety, and health insurance companies, who would be interested in distinguishing community pharmacies in patient-oriented pharmaceutical care. Assessment of the QI set according to the criteria of the IAF should lead to publication of valid QI results traceable for each individual community pharmacy to provide all stakeholders with meaningful information. This summative purpose, different from the former set, needed extra requirements for the QIs in order to allow valid comparisons between community pharmacies,[8,17] based on a structured approach that included an empirical or practice test.[4,11] The purpose of this study was to assess the validity of the QI set for community pharmacies using the IAF.

Methods
Study Design
Results of QI assessment based on IAF criteria were evaluated for IAF criteria (content validity, absence of selection bias, absence of measurement bias, and statistical reliability); indicator type (structure, process, and outcome); and domain (e.g., compounding, counseling).

Setting
The pharmacy task group of the Health Care Transparency Programme requested the independent research institute Significant to facilitate and guide the assessment of the QI set during June and August 2012. Significant had participated in the development of the general IAF and thus could adapt this general tool to the characteristics of the QI set for community pharmacies. The current version of the IAF became available in February 2012[16] and is based on epidemiological principles and consensus techniques following national and international guidelines.[9,18] An expert panel was formed, which consisted of 6 pharmacists who had participated in the data collection—5 practicing community pharmacists and a pharmacist/epide-

miologist, who served as project leader in the development and data collection process. The community pharmacists were carefully chosen as experienced representatives from urban as well as rural areas and from different settings, such as independent pharmacies, pharmacies in pharmacy chains, or pharmacies in health centers.

QI Set and Data in 2011
The set of 52 QIs covered the following 10 domains (Table 1): continuity of pharmaceutical care (4 QIs); patient counseling (3), clinical risk management (12), compounding (3), dispensing (6), monitoring of medication use (10), self-care support and over-the-counter medications (2), logistics (5), quality management (6), and professional development (1). The QI set addressed structures (21), processes (19), and outcomes (12) for community pharmacies. The majority of the structure indicators were measured by dichotomous questions, whereas process indicators and outcome indicators were mostly measured by numerical measures.

Between April and May 2012, all 1,987 Dutch community pharmacies were requested to complete an online questionnaire about the QIs. This questionnaire contained 1 or more questions for each QI with options for dichotomous or categorical answers or fields to provide numerators and denominators for numerical QIs. Per QI, community pharmacists could add comments and information in a free text field in the online questionnaires. During data collection, a help desk was available for questions. Results from all community pharmacy questionnaires were used to generate benchmark reports. The benchmark results, response rates, comments, and questions were accessible to the expert panel and thus could serve as information from a practice test. For dichotomous QIs, the benchmark reports provided insight into discriminatory power and ceiling effects.

Measures
The IAF supplies 4 criteria for assessing each QI: content validity, selection bias, measurement bias, and statistical reliability (used only for QIs measured on a numerical scale). Furthermore, the IAF judges the content validity for the whole set. The expert panel rated content validity and absence of selection and measurement bias as meeting the requirements, not meeting the requirements, or partly meeting the requirements. Statistical reliability was assessed by statistical measures according to IAF requirements, and results were expressed in the same way. To validate the pharmacy QI set, the 4 IAF criteria were applied as described in the following sections.

Content Validity. Content validity is considered the most important criterion within the IAF. It is defined as the degree to which the QI directly reflects the performance of the community pharmacist or pharmacy team. Strongest evidence for a

| TABLE 1 | Expert Panel Assessment Scores of QIs Grouped by Domain, 2011 |

No.[c]	Domain	Type[d]	Data Collection with Questionnaire[a]		Final Rating of Expert Assessment[b]			
			N (Response %)[e]	Comments (%)	Content Validity[f]	Absence of Selection Bias	Absence of Measurement Bias	Statistical Reliability[g]
1	Continuity of Care[h]							
1.1	Checking and registration of current medication use	S	1,807 (98.8)	254 (14.1)	1	1	0	n/a
1.2	Management of interactions between oral anticoagulants and co-trimoxazole	S	1,807 (98.7)	344 (19.0)	1[33]	1	1	n/a
1.3	Participation of pharmacotherapy audit meetings with GPs and community pharmacists	S	1,807 (99.0)	151 (8.4)	0.5[34]	0.5	0.5	n/a
1.4	Percentage of patients with actual and verified documentation of current medication use prior to and after dispensing of medication	P	1,807 (72.1)	721 (39.9)	1	1	0	0.5
	Mean score of QIs within this domain				0.88	0.88	0.38	0.50
2	Patient Counseling							
2.1	Percentage of users of inhalation medication who received instructions for the use of inhaler devices first dispensed	P	1,807 (85.5)	567 (31.4)	0.5	0.5	0	1
2.2	Percentage of benzodiazepine users who received verbal information about dependency with a follow-up prescription for benzodiazepines	P	1,807 (91.3)	440 (24.3)	0	0.5	0	n/a
2.3	Percentage of patient counseling pharmacy staff with demonstrable competency in patient counseling	S	1,807 (91.4)	321 (17.8)	0	1	0.5	0
	Mean score of QIs within this domain				0.17	0.67	0.17	0.50
3	Clinical Risk Management							
3.1	Availability of protocols concerned with asking patients about potential contraindications	S	1,807 (98.8)	131 (7.2)	1	1	1	n/a
3.2	Asking the prescriber for confirmation of suspected contraindications	P	1,807 (98.6)	311 (17.2)	0.5	0	0.5	n/a
3.3	Percentage of patients chronically using loop diuretics and RAS inhibitors who are dispensed NSAIDs	O	1,807 (98.0)	166 (9.2)	0[31]	0.5	0.5	0
3.4	Percentage of patients with documented contraindication of heart failure who are dispensed NSAIDs	O	1,807 (74.5)	532 (29.4)	0.5[31]	0.5	0	0.5
3.5	Availability of protocols concerned with asking patients about possible drug intolerabilities	S	1,807 (99.5)	108 (6.0)	1	1	1	n/a
3.6	Documentation of the management of drug-drug interactions	S	1,807 (99.3)	209 (11.6)	1	1	0	n/a
3.7	Number of patients who concurrently use oral anticoagulants and co-trimoxazole	O	1,807 (99.6)	264 (14.6)	0[35]	0	0	n/a
3.8	Number of patients who concurrently use oral anticoagulants and miconazole	O	1,807 (99.6)	89 (4.9)	1	0.5	0	n/a
3.9	Percentage of dispenses with quality assurance steps applied in the dispensing process	P	1,807 (83.5)	704 (39.0)	0.5[36]	0	0	1
3.10	Availability of protocols concerning dose control of compounded medicine for systemic use	S	1,807 (98.1)	n/a	1	1	0	n/a
3.11	Percentage of compounded medicine for systemic use with dose control	P	1,807 (99.0)	311 (17.2)	1	1	0	n/a
3.12	Patients using metformin in a daily dosage of 1,000 mg and above, with a creatinine clearance below 50 mL/min with medication adjustment according to creatinine clearance.	P	1,807 (71.2)	767 (42.4)	1	0.5	0	0
	Mean score of QIs within this domain				0.71	0.58	0.25	0.38
4	Compounding							
4.1	Arrangements made for checks by compounding pharmacy in case of delegated compounding	S	1,807 (94.4)	334 (18.5)	1	1	1	n/a
4.2	Releasing technician compounded medicines by a pharmacist before dispensing to patients	S	1,807 (83.0)	n/a	1	1	1	n/a

TABLE 1 Expert Panel Assessment Scores of QIs Grouped by Domain, 2011 *(continued)*

No.[c]	Domain	Type[d]	N (Response %)[e]	Comments (%)	Content Validity[f]	Absence of Selection Bias	Absence of Measurement Bias	Statistical Reliability[g]
			Data Collection with Questionnaire[a]		Final Rating of Expert Assessment[b]			
4.3	Stock compounding according to written protocols and nonstandardized compoundings according to guidelines	P	666[i] (72.7)	229 (34.4)	1	1	0	0
	Mean score of QIs within this domain				1.00	1.00	0.67	0
5	Dispensing of Medication							
5.1	Percentage of community-dwelling patients aged >65 years with medication dispensed in individualized dosage systems	P	1,807 (95.0)	231 (12.8)	0	0	0.5	1
5.2	Registration of errors during dispensing of medication, dispensed to patient	S	322[j] (99.0)	n/a	1	1	0	1
5.3	Percentage of registered errors during dispensing of medication, dispensed to patient	P	1,807 (94.8)	n/a	0.5	0.5	0	n/a
5.4	Availability of protocols on generic substitution for information to patients, prescribers, and persons who administer medicines	S	1,807 (97.1)	203 (11.2)	0.5	0.5	0	n/a
5.5	Compliance of individualized automated dispensing packages with national standards	S	1,807 (98.0)	159 (8.8)	1	1	1	n/a
5.6	Availability of control procedures for medication in weekly dosage forms (not automated)	S	1,807 (99.6)	115 (6.4)	1	1	1	n/a
	Mean score of QIs within this domain				0.67	0.67	0.42	1.00
6	Monitoring of Medication Use							
6.1	Percentage of benzodiazepine users (aged >65 years) who chronically use benzodiazepines	O	1,807 (98.2)	109 (6.0)	0	0	1	0.5
6.2	Number of medication reviews performed for patients with polypharmacy (>5 different medication groups)	P	1,807 (85.9)	442 (24.5)	0.5[37-39]	0.5	0	n/a
6.3	Percentage of classical NSAID users (aged >70 years) with gastro protection	O	1,807 (98.3)	304 (16.8)	1[31]	0.5	0.5	0.5
6.4	Percentage of patients who concurrently use nitrates and antithrombotics	O	1,807 (97.7)	242 (13.4)	1[31]	0.5	0.5	0.5
6.5	Percentage of users of strong opiates who take laxatives (lactulose, magnesium oxide, bisacodyl, macrogol, or sennosides)	O	1,807 (98.0)	536 (29.7)	1[31]	0.5	0.5	0.5
6.6	Percentage of dispensed third-generation chinolones	O	1,807 (98.9)	124 (6.9)	0.5	0.5	1	n/a
6.7	Monitoring and counseling of adherence to chronic medication used	P	1,807 (89.4)	378 (20.9)	0.5	0	0	n/a
6.8	Percentage of new and existing users of simvastatin who use simvastatin in a lower dose than according to national standards	O	1,807 (92.3)	309 (17.1)	0	0.5	0	0.5
6.9	Percentage of patients who use an antidepressant (SSRI) longer than 24 months continuously	O	1,807 (97.5)	383 (21.2)	0	0	0	0.5
6.10	Percentage of patients who continuously use proton pump inhibitors longer than 12 months	O	1,807 (98.3)	147 (8.1)	0.5	0.5	1	1
	Mean score of QIs within this domain				0.50	0.35	0.45	0.57
7	Self-Care Support							
7.1	Following protocols for dispensing high-risk OTC medication	S	1,807 (83.8)	n/a	1	0.5	0	n/a
7.2	Percentage of high-risk OTC medication dispensings according to protocol	P	1,807 (66.9)	n/a	0.5	0.5	0	n/a
	Mean score of QIs within this domain				0.75	0.50	0.00	
8	Logistics							
8.1	Structures available for assessment of suppliers	S	1,807 (99.6)	n/a	1	1	1	n/a
8.2	Assessment of reliability of suppliers	P	322[9] (100.0)	n/a	1	1	1	n/a
8.3	Availability of a valid expiration system	S	1,807 (98.4)	n/a	1	1	1	n/a
8.4	Number of expired medicines dispensed to patients	P	1,807 (83.8)	n/a	1	0.5	0	n/a

TABLE 1	Expert Panel Assessment Scores of QIs Grouped by Domain, 2011 *(continued)*							
			Data Collection with Questionnaire[a]		**Final Rating of Expert Assessment[b]**			
No.[c]	Domain	Type[d]	N (Response %)[e]	Comments (%)	Content Validity[f]	Absence of Selection Bias	Absence of Measurement Bias	Statistical Reliability[g]
8.5	Number of inaccurately performed recalls of medicine	P	1,807 (88.8)	346 (19.1)	1	1	0.5	n/a
	Mean score of QIs within this domain				1.00	0.90	0.70	
9	Quality Management							
9.1	Registration of errors made during dispensing of medication, noticed within the pharmacy	S	1,807 (86.7)	454 (25.1)	1	1	0	1
9.2	Year of most recent evaluation of patients' experiences	S	1,807 (95.0)	207 (11.5)	1	1	1	n/a
9.3	Number of registered complaints made by patients	P	1,807 (47.3)	285 (15.8)	0.5	0.5	0	1
9.4	Number of adverse drug reactions reported to the Netherlands Pharmacovigilance Centre	P	1,807 (66.8)	205 (11.3)	0.5	0.5	0.5	1
9.5	Availability of certified quality management system	S	1,807 (99.8)	201 (11.1)	1	1	1	n/a
9.6	Parameters for clinical risk management in the pharmacy information system are set up according to the prevailing standards	S	1,807 (89.5)	123 (6.8)	1	1	1	n/a
	Mean score of QIs within this domain				0.83	0.83	0.58	1.00
10	Professional Education							
10.1	Professional development of pharmacy staff	P	1,807 (87.1)	514 (28.4)	0.5	1	0	0

[a]Results from national data collection.

[b]Scores based on consensus meeting; for all criteria: 1 = meeting the requirements; 0.5 = partly meeting the requirements; 0 = not meeting the requirements.

[c]QIs assessed as valid on all criteria in bold.

[d]Indicator type: S = structure; P = process; O = outcome.

[e]N = number of pharmacies elegible; % = response percentage.

[f]Only references that were available to the expert panel are listed in the table.

[g]Only assessed on numerical indicators, expressing the power of the indicator. Where statistic validity was not applicable, this was counted as 1 (and presented between parentheses).

[h]Indicator domain.

[i]Only to be answered by compounding pharmacies.

[j]Only to be answered by pharmacies without certified quality management system (QI 9.5) because assessing this process is part of the audit.

GP = general practitioner; mg = milligram; mL/min = milliliter per minute; n/a = not available; NSAID = nonsteroidal anti-inflammatory drug; OTC = over the counter; QI = quality indicator; RAS = renin-angiotensin system; SSRI = selective serotonin re-uptake inhibitor.

causal relationship between pharmaceutical care and the result of a QI is delivered by randomized controlled trials, followed by cohort, case control studies, and case reports from the literature. If this evidence is lacking, consensus of the expert panel was accepted. When the score of the QI was likely to directly reflect the quality of the activities of the pharmacist or the pharmacy team, the criterion was rated as meeting the requirements. A QI was scored as not meeting the requirements when activities of the pharmacist and pharmacy team were not likely to affect the score of that QI at all, or when individual responses to the questionnaire or to the help desk showed that the question had not been understood correctly, and it was assumed that the answers given were not likely to be a response to the question as intended. When the pharmacist or the pharmacy team could only partly or indirectly influence the score of the QI, or when the QI did not strongly reflect their activities, content validity was considered as partly meeting the requirements.

Absence of Selection Bias. This criterion reflects the degree to which differences between populations of pharmacies with regard to age, drug use, morbidity, or social economic status could have influenced the results of a QI. This criterion was rated as meeting the requirements when differences in populations were not likely to influence the results of a QI. As data were collected per pharmacy, no information on the patient level was available for case-mix adjustments. A QI was judged as not meeting the requirements for this criterion when the results were more likely to represent differences in populations between pharmacies rather than differences in the pharmaceutical care provided. A QI was scored as partly meeting the requirements when population bias could not fully be excluded, but the QI results were likely to depend at least equally on pharmaceutical care as on differences in populations. Information on this aspect was also derived from comments or questions during data collection.

Absence of Measurement Bias. This criterion deals with differences in data collection by community pharmacies that were likely to bias comparisons between QI results. This criterion

TABLE 2	Expert Panel Assessment Scores of 2011 QIs, Means Per Indicator Type			
Type of Indicator	Content Validity	Absence of Selection Bias	Absence of Measurement Bias	Statistical Reliability
Overall	0.70	0.67	0.40	0.57
Structure indicators (S)	0.90	0.93	0.62	0.66
Process indicators (P)	0.63	0.55	0.16	0.61
Outcome indicators (O)	0.46	0.38	0.42	0.50
QI = quality indicator.				

was scored as meeting the requirements when all pharmacies were considered to have collected information on the aspect in a comparable way. A QI was rated as not meeting the requirements when information for a numerical QI could not be derived from registrations within the primary process or when free text or the logbook showed that the questions of a QI were probably misunderstood. A low response rate on specific numerical QIs ($< 50\%$) was examined as a potential indication of measurement bias. A QI was rated as partly meeting the requirements when measurement bias could not be excluded, but the QI results were likely to depend at least equally on pharmaceutical care as on differences in data collection.

Statistical Reliability. To detect differences between community pharmacies with sufficient statistical confidence, results from numerical QIs were subjected to a statistical test. Numerical QIs were rated for detectable variation by Cohen's effect size. As such, we measured whether the sample size was high enough to truly detect a difference for a power of at least 80% at a confidence level of 5%. A QI met the requirements for detecting small differences with a Cohen's $d \leq 0.2$; partly met the requirements for detecting medium differences with a Cohen's d between 0.2 and 0.5; and did not met the requirements with a Cohen's $d > 0.5$.[16]

Judgment in Consensus Meeting

To score all 52 QIs with respect to 3 of the 4 criteria, a table was developed according to the IAF methodology to facilitate structured assessment of all QIs by the expert panel. The table recited the relevant points per criterion as addressed by the IAF and thus guided the experts to their final judgment per QI. In preparation for a consensus meeting, each panel member individually rated the QI set for content validity, absence of selection bias, and measurement bias. During the consensus meeting, the individual scores were discussed, and a unanimous assessment score was allocated per QI for content validity, absence of selection bias, and measurement bias. For numerical QIs, the estimation of statistical reliability was added from statistical tests. This criterion was regarded as nonapplicable for non-numerical QIs.

Analysis

For this study, we constructed overall validation scores per QI for the IAF criteria, QI types, and domains. To calculate means, we assigned 1 point to meeting the requirements; 0.5 points to partly meeting the requirements; and 0 points to not meeting the requirements. Analysis was performed with Microsoft Excel 2007 (Microsoft Corp., Redmond, WA).

■ Results

For the validation process, the information from the online questionnaire was available, provided by pharmacists from 1,807 (91%) community pharmacies (Table 1). Most of the QIs (45) had a response rate between 83% and 100%. Six of the remaining QIs had a response rate between 67% and 74%. The lowest response rate (47%) was due to technical problems with the online questionnaire, according to the comments made. In total, 17,565 comments were made on the 52 QI indicators; 2,419 of these comments comprised general statements about availability of a certified quality management system and were consequently not considered relevant for the validation process. The remaining 15,146 comments were related to difficulties in measuring, understanding the questions, and aspects affecting the validity and reliability of the answers given.

In order to assess the validity of the QI set for community pharmacies, the expert panel judged the 52 QIs for content validity and absence of selection and measurement bias, along with content validity for the complete set, and thus allocated 157 validity scores (Table 1). The scores for the different criteria reached on average 0.70 for content validity, 0.67 for the absence of selection bias, 0.40 for the absence of measurement bias, and 0.57 for statistical reliability (Table 2). Content validity for the complete set was judged as partly meeting the requirements. Regarding the QI types, structure indicators scored the highest on all criteria. Process indicators scored lowest on measurement bias, whereas outcome indicators scored low on all criteria, including content validity.

Within the separate criteria, 30 (58%) QIs met the requirements for content validity. Eight (15%) of the 52 QIs did not meet those requirements: 5 were outcome indicators; 2 were process indicators; and 1 was a structure indicator. An example

of an indicator that did not meet the requirements for content validity was QI 6.9 ("Percentage of patients who use an antidepressant longer than 24 months continuously without a retraceable action of the pharmacist to alert the prescriber"). Content validity for this outcome indicator was assessed as not meeting the requirements because an effect of pharmaceutical care on this result is not supported by literature or primary care guidelines and was doubted by the expert panel.[19] Twenty-four QIs (46%) met the requirements for absence of selection bias. For QI 6.9, selection bias seemed probable because the indication for antidepressants could vary between pharmacy populations and could thus influence the duration of antidepressant use. For absence of measurement bias, 16 QIs (31%) met the requirements. Measurement bias was most prominent for process indicators, resulting in a low average score of 0.16. Pharmacists' comments in the questionnaire's free text field also suggested that pharmaceutical care processes could not be measured in an unambiguous and uniform way across all participating community pharmacies. For QI 6.9, measurement bias was plausible because contacts with prescribers were not unambiguously registered in community pharmacies, as was indicated in most of the 383 comments in the questionnaire. To some extent, this issue was applicable for other outcome indicators, such as QI 6.4 ("Percentage of patients who concurrently use nitrates and antithrombotics"). Statistical reliability was tested on the 21 numerical QIs. Eight (38%) of these met the requirements. For some QIs, measurement bias was due to ambiguous definitions in the process, as was elucidated by the comments on QI 9.3 ("Number of registered complaints made by patients").

Thirteen QIs (25%) met all 4 criteria—12 were structure indicators, and 1 was a process indicator. Structure indicator 9.5 ("Does the community pharmacy have a certified quality management system?") is an example of a QI that is valid for all criteria. This measure concerns whether pharmacies conduct a structured assessment and improvement of quality of care, externally assessed and certified, according to national quality and safety standards.[20] For this QI, content validity was allocated by expert agreement on the assumption that external assessment warrants a high standard of pharmaceutical care in community pharmacies. All other criteria were met: the presence of a quality management system did not depend on the characteristics of the patient population; no measurements were needed; and statistical reliability was not assessed for this dichotomous QI, while 80% answering with "yes" still left enough range for improvement. The QIs meeting all criteria were present in 6 of the 10 domains of the set: continuity of care (1 out of 4 QIs), clinical risk management (2/12), compounding (2/3), dispensing of medication (2/6), management

(3/5), and quality management (3/6). In the domains of patient counseling, monitoring of medication use, self-care support, and professional development, no fully valid QIs were present.

◼ Discussion

To assess validity, the 52 QIs for community pharmacies were judged by an expert panel using the IAF. Thirteen QIs (25%) were assessed as meeting all requirements for all criteria of the IAF and thus were judged as valid. This was lower than expected, since the majority of QIs from the set was assessed during their development and have been used for several years already.[5] This conservative judgment might be explained by the fact that the initial set was mainly developed for formative use, whereas the actual set was also validated for summative use, including publication of QI results for each community pharmacy. Earlier studies have indicated that summative use requires more rigorous evaluation to warrant valid and fair comparisons between health care providers.[8,17] Additionally, the relationship between process indicators and patient outcome can be weak when used for pay for performance or public reporting.[21,22]

From the 4 criteria validated, measurement bias showed as the main problem for process indicators because of lack of uniform registrations of pharmaceutical care with a reasonable effort as part of the routine care process. This requires national consensus on the way care registration is handled in community pharmacies and implementation in the pharmacy software systems. Content validity was particularly an obstacle to meeting the requirements for outcome indicators. It has to be noted that the outcomes measured in this QI set dealt with the outcome of pharmacy processes and did not consider patient outcomes. Information from patient-reported outcome measures (PROMs), indicating the result of care for the individual patient, have to be derived from patients.[23-25] At the time of this study, PROMs were rarely available for the assessment of pharmaceutical care; consequently, they should be developed in the future.

The IAF criteria were appropriate to assess the set of QIs for community pharmacies. The assessment results of the QI set show that the expert panel applied rigorous requirements for QI validity, possibly because of the potential use of QIs for summative purposes. The development of a QI set is a continuously structured process in order to provide a balanced set that covers the relevant pharmacy domains and meets the needs of all stakeholders—pharmacists, the Healthcare Inspectorate, patients, and health insurance companies.[26,27] Future development of the set includes improving the QIs that partly met the requirements, as well as removing the QIs that did not meet the requirements and could not be improved. In particular,

the QIs that lack content validity (such as QI 6.9—"Percentage of patients who use an antidepressant longer than 24 months continuously") should be removed or replaced by superior measures. QIs that have lost their relevance should be suspended, as well. New QIs that reflect new needs or recent guidelines should be introduced. The IAF can play a role in this continuous maintenance of the QI set by validating new QIs.

The nationwide development of an integrated QI set might not directly be applicable for other countries; however, our findings may be valuable. To be more specific, QIs and their assessment scores could be useful in developing or evaluating quality measures; for instance, as a performance measure within one of the outcome domains in the catalog of Pharmacy Quality Indicators by the Academy of Managed Care Pharmacy.[2] In particular, QIs that resemble measures in other sets, such as QI 6.4 ("Percentage of patients who concurrently use nitrates and antithrombotics"), deserve special attention.[28]

Limitations

Within the validation process, a number of limitations have to be addressed. First, the expert panel consisted only of pharmacists; no other stakeholders were represented. This setting was chosen to warrant a qualified assessment based on sufficient experience in the daily practice of community pharmacies. An earlier study on QIs assessment for cardiovascular risk management compared assessment between professionals and other stakeholders.[29] The study concluded that professionals were more qualified in assessing QIs than other stakeholders, which may justify our approach. Second, the members of the expert panel were directly affected as payees by potential use of the QI set for pay-for-performance programs. Therefore, they were likely to apply the IAF criteria in a strategically defensive way, leading to conservative judgments.[30] Third, evidence for content validity was mainly based on expert assessment, the lowest level of evidence as defined in the IAF.[16] Some publications on effects of pharmaceutical care were available, although these were not randomized controlled trials.[31,32] Lack of a strong evidence base, however, is common for QIs in many areas of primary care.[9] Since the assessment depended on the consensus of the expert panel, the panel members were carefully chosen from different regions and care settings. Furthermore, an independent research institute guided the structured testing by the IAF and the validation process.

■ Conclusions

The assessment of a set of QIs using the IAF elucidated the QIs on which community pharmacies were willing to be judged for summative purposes. Furthermore, the assessment revealed the critical issues for further improvement of the QI set.

Authors

TIM W.A. SCHOENMAKERS, PharmD, is PhD Candidate, Radboud University Medical Centre, Radboud Institute for Health Sciences, Department of IQ Healthcare, Nijmegen, The Netherlands, and Pharmacist, Zorgapotheek Nederland BV, Utrecht, The Netherlands; MARTINA TEICHERT, PharmD, PhD, is Researcher, Radboud University Medical Centre, Radboud Institute for Health Sciences, Department of IQ Healthcare, Nijmegen, and Scientific Advisor, Royal Dutch Pharmacists Association (KNMP), The Hague, The Netherlands; and Peter A.G.M. DE SMET, PharmD, PhD, is Professor, Radboud University Medical Centre, Radboud Institute for Health Sciences, Department of Clinical Pharmacy and IQ Healthcare, Nijmegen, and Scientific Advisor, Royal Dutch Pharmacists Association (KNMP), The Hague. JOZÉ BRASPENNING, PhD, is Assistant Professor, and MICHEL WENSING, PhD, is Professor of Implementation Science, Radboud University Medical Centre, Radboud Institute for Health Sciences, Department of IQ Healthcare, Nijmegen. LYDIA VUNDERINK, MSc, is Senior Consultant, Significant, Barneveld, The Netherlands.

AUTHOR CORRESPONDENCE: Tim W.A. Schoenmakers, PharmD, Radboud Institute for Health Sciences, Department of IQ Healthcare, Radboud University Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands. Tel.: +31(0)24.3615305; Fax: +31(0)24.3540166; E-mail: tim.schoenmakers@radboudumc.nl.

REFERENCES

1. Curtiss FR, Fry RN, Avey SG. Framework for pharmacy services quality improvement—a bridge to cross the quality chasm. Part I. The opportunity and the tool. *J Manag Care Pharm.* 2004;10(1):60-78. Available at: http://www.amcp.org/WorkArea/DownloadAsset.aspx?id=6908.

2. Academy of Managed Care Pharmacy. Catalog of pharmacy quality indicators: select those that are important to your organization v.2.0. 2007. Available at: http://www.amcp.org/WorkArea/DownloadAsset.aspx?id=9273.

3. Nouwens E, Van Lieshout J, Adang E, Bouma M, Braspenning J, Wensing M. Effectiveness and efficiency of a practice accreditation program on cardiovascular risk management in primary care: study protocol of a clustered randomized trial. *Implement Sci.* 2012;7:94.

4. Giesen P, Willekens M, Mokkink H, Braspenning J, Van Den Bosch W, Grol R. Out-of-hours primary care: development of indicators for prescribing and referring. *Int J Qual Health Care.* 2007;19(5):289-95.

5. De Bie J, Kijlstra NB, Daemen BJ, Bouvy ML. The development of quality indicators for community pharmacy care. *BMJ Qual Saf.* 2011;20(8):666-71.

6. Szecsenyi J, Broge B, Eckhardt J, Heller G, Kaufmann-Kolle P, Wensing M. Tearing down walls: opening the border between hospital and ambulatory care for quality improvement in Germany. *Int J Qual Health Care.* 2012; 24(2):101-04.

7. Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. *New Engl J Med.* 2009;361(4):368-78. Available at: http://www.nejm.org/doi/full/10.1056/NEJMsa0807651. Accessed January 5, 2015.

8. Romano PS, Mull HJ, Rivard PE, et al. Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. *Health Serv Res.* 2009;44(1):182-204.

9. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care.* 2002;11(4):358-64. Available at: http://qualitysafety.bmj.com/content/11/4/358.long. Accessed January 5, 2015.

10. Campbell SM, Kontopantelis E, Hannon K, Burke M, Barber A, Lester HE. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. *BMC Fam Pract.* 2011;12:85. Available at: http://www.biomedcentral.com/1471-2296/12/85. Accessed January 5, 2015.

11. Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators—a systematic review. *Implement Sci.* 2012;7:21.

12. National Institute for Health and Clinical Excellence. Developing clinical and health improvement indicators for the Quality and Outcomes Framework (QOF): interim process guide. May 2009. Available at: http://www.nice.nhs.uk/media/742/32/QOFProcessGuide.pdf. Accessed January 5, 2015.

13. AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH. Allgemeine Methoden im Rahmen der sektorenübergreifenden Qualitätssicherung im Gesundheitswesen nach § 137a SGB V. Version 2.0. [General methods for cross-sectoral quality assurance in health care in accordance with §137a Volume V of the Social Insurance Code]. June 2010. Available at: http://www.aqua-institut.de/aqua/upload/CONTENT/Projekte/137a/Methodenpapier/AQUA_AllgemeineMethoden_Version_2-0.pdf. Accessed January 5, 2015.

14. National Quality Forum. Measure evaluation criteria 2015. Available at: http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx. Accessed January 5, 2015.

15. Visser J, Zuidgeest M, Appelman M, Koolman X. Routekaart. Te maken keuzes bij de ontwikkeling van betrouwbare kwaliteitsindicatoren in de zorg. [Roadmap for developing reliable quality indicators in health care]. December 2012. Available at: http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/rubrieken/kwaliteit/1212-routekaart/1212-routekaart/Routekaart.pdf. Accessed January 5, 2015.

16. Koolman X, Zuidgeest M, Visser J, Appelman M. Indicatorstandaard: methodologische criteria voor de ontwikkeling van betrouwbare kwaliteitsindicatoren in de zorg. [Indicator standard: methodological criteria for development of reliable quality indicators in health care]. December 2012. Available at: http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/rubrieken/kwaliteit/1212-indicatorstandaard/1212-indicatorstandaard/Indicatorstandaard.pdf. Accessed January 8, 2015.

17. Kaafarani HM, Borzecki AM, Itani KM, et al. Validity of selected Patient Safety Indicators: opportunities and concerns. *J Am Coll Surg.* 2011(6);212:924-34.

18. De Koning JS, Smulders A, Klazinga NS. *Appraisal of Indicators through Research and Evaluation (AIRE).* 2.0 ed. Utrecht: Orde van Medisch Specialisten, 2007.

19. van Weel-Baumgarten EM, van Gelderen MG, Grundmeijer HGLM, et al. NHG-Standaard Depressie (tweede herziening). *Huisarts en Wetenschap.* 2012;55(6):252-59.

20. Zorginstituut Nederland. HKZ-keurmerken. [HKZ (Foundation for Harmonization of Quality Review in Health Care and Welfare) ISO compatible standards for quality and patient safety]. Available at: http://www.kiesbeter.nl/Kwaliteitvanzorg/Keurmerkenindezorg/Instellingsgerelateerdekeurmerken/HKZkeurmerken.aspx. Accessed January 16, 2015.

21. Nicholas L, Osborne N, Birkmeyer J, Dimich J. Hospital process compliance and surgical outcomes in Medicare beneficiaries. *Arch Surg.* 2010;145(10):999-1004.

22. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA.* 2006;296(22):2694-702.

23. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. *BMJ.* 2010;340:c186.

24. Kerr EA, Hayward RA. Patient-centered performance management: enhancing value for patients and health care systems. *JAMA.* 2013;310(2):137-38.

25. Basch E, Torda P, Adams K. Standards for patient-reported outcome-based performance measures. *JAMA.* 2013;310(2):139-40.

26. Mattke S. When should measures be updated? Development of a conceptual framework for maintenance of quality-of-care measures. *Qual Saf Health Care.* 2008;17(3):182-86.

27. Meyer GS, Nelson EC, Pryor DB, et al. More quality measures versus measuring what matters: a call for balance and parsimony. *BMJ Qual Saf.* 2012;21(11):964-68.

28. National Committee for Quality Assurance. List of NQF-endorsed measures. April 2008. Available at: http://www.ncqa.org/LinkClick.aspx?fileticket=p6Oa24wNQ6Y%253d&tabid=59&mid=1604&forcedownload=true%20. Accessed January 8, 2015.

29. van Lieshout J, Nouwens E, Bouma M, Spreeuwenberg C, Wensing M. Consistency of performance indicators for cardiovascular risk management across procedures and panels. *Qual Saf Health Care.* 2010;19(5):e31.

30. Tillema S. Public sector benchmarking and performance improvement: what is the link and can it be improved? *Public Money Manag.* 2010;30(1):69-75.

31. Warlé-van Herwaarden MF, Kramers C, Sturkenboom MC, Bemt PM, De Smet PA. Targeting outpatient drug safety: recommendations of the Dutch HARM-Wrestling Task Force. *Drug Saf.* 2012;35(3):245-59.

32. Morris CJ, Cantrill JA, Hepler CD, Noyce PR. Preventing drug-related morbidity—determining valid indicators. *Int J Qual Health Care.* 2002;14(3):183-98. Available at: http://intqhc.oxfordjournals.org/content/14/3/183.full.pdf. Accessed January 8, 2015.

33. Schalekamp T, Smit C, van Geest-Daalderop JH, de Vries-Goldschmeding H, de Boer A. Discrepancies between medication records of anticoagulation clinics and pharmacy records. *Pharmacoepidemiol Drug Saf.* 2006;15(11):823-28.

34. Teichert M, van der Aalst A, de Wit H, Stroo M, De Smet PAGM. How useful are prescribing indicators based on the DU90% method to distinguish the quality of prescribing between pharmacotherapy audit meetings with different levels of functioning? *Eur J Clin Pharmacol.* 2007;63(12):1171-77. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2071964/. Accessed January 8, 2015.

35. Schalekamp T, van Geest-Daalderop JH, Kramer MH, van Holten-Verzantvoort AT, de Boer A. Coumarin anticoagulants and co-trimoxazole: avoid the combination rather than manage the interaction. *Eur J Clin Pharmacol.* 2007;63(4):335-43.

36. Buurma H, de Smet PAGM, van den Hoff OP, Egberts ACG. Nature, frequency and determinants of prescription modifications in Dutch community pharmacies. *Br J Clin Pharmacol.* 2001;52(1):85-91. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2014512/. Accessed January 8, 2015.

37. Schnipper JL, Kirwin JL, Cotugno MC, et al. Role of pharmacist counseling in preventing adverse drug events after hospitalization. *Arch Intern Med.* 2006;166(5):565-71. Available at: http://archinte.jamanetwork.com/article.aspx?articleid=409914. Accessed January 8, 2015.

38. Finkers F, Maring J, Boersma F, Taxis K. A study of medication reviews to identify drug-related problems of polypharmacy patients in the Dutch nursing home setting. *J Clin Pharm Ther.* 2007;32(5):469-76.

39. Krska J, Cromarty JA, Arris F, et al. Pharmacist-led medication review in patients over 65: a randomized, controlled trial in primary care. *Age Ageing.* 2001;30(3):205-11.