MAJOR ARTICLE

Quality of Antibiotic Use for Lower Respiratory Tract Infections at Hospitals: (How) Can We Measure It?

J. A. Schouten,[1,2,3] M. E. J. L. Hulscher,[1] H. Wollersheim,[1,3] J. Braspennning,[1] B. J. Kullberg,[2,3] J. W. M. van der Meer,[2,3] and R. P. T. M. Grol[1]

[1]Centre for Quality of Care Research, [2]Nijmegen University Centre for Infectious Diseases, and [3]Department of General Internal Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Background. To assess and improve the quality of antibiotic use in patients with community-acquired pneumonia (CAP) and acute exacerbation of chronic bronchitis or chronic obstructive pulmonary disease (AECB), a valid set of quality indicators is required. This set should also be applicable in practice.

Methods. Guidelines and literature were reviewed to derive potential indicators for quality of antibiotic use in treating hospitalized patients with lower respiratory tract infection (LRTI). To assess the evidence base of each indicator, a literature review was performed. Grade A recommendations were considered valid. For grade B–D recommendations, an expert panel performed a consensus procedure on the indicator's relevance to patient health, reduction of antimicrobial resistance, and cost containment. To test applicability in practice, feasibility, opportunity for improvement, reliability, and case-mix stability were determined for a data set of 899 hospitalized patients with LRTI.

Results. None of the potential indicators from guidelines and literature were supported by grade A evidence. Nineteen indicators were selected by consensus procedure (12 indicators for CAP and 7 indicators for AECB). Lack of feasibility and of opportunity for improvement led to the exclusion of 4 indicators. A final set of 15 indicators was defined (9 indicators for CAP and 6 indicators for AECB).

Conclusions. A valid set of quality indicators for antibiotic use in hospitalized patients with LRTI was developed by combining evidence and expert opinion in a carefully planned procedure. Subjecting indicators to an applicability test is essential before using them in quality-improvement projects. In our demonstration setting, 4 of the 19 indicators were inapplicable in practice.

Community-acquired lower respiratory tract infection (LRTI) is a common cause of acute illness in adults. The spectrum of disease ranges from mild mucosal colonization or infection, to acute bronchitis or acute exacerbation of chronic bronchitis or chronic obstructive pulmonary disease (AECB), to overwhelming parenchymal infection in patients with community-acquired pneumonia (CAP). Antibiotic treatment is rarely indicated for acute bronchitis and is sometimes indicated for the more severe cases of AECB, but it is always indicated for CAP. It may be difficult to differentiate between viral and bacterial LRTI or between bronchitis, AECB, and CAP. This may be one of the reasons why antibiotics are prescribed to more than two-thirds of patients with LRTI in Europe and the United States. In view of the worldwide development of antibiotic resistance, this is not a desirable situation [1].

Recommendations for the rational use of antibiotics in hospitalized patients with LRTI have been formulated in national and international guidelines [2–7]. Guidelines describe, in essence, "the right thing to do." They assist in making practitioner and patient decisions prospectively for specific clinical circumstances. To make a valid and reliable assessment of current practice in patients with LRTI, key recommendations from these guidelines can be translated into measurable elements—so-called "indicators" [8]. Indicators serve as measurable elements of practice performance for which there is evidence or consensus that they can be used to assess the quality (and, therefore, change in the quality) of care provided [9].

Several articles have suggested quality indicators for

Received 10 January 2005; accepted 6 April 2005; electronically published 8 July 2005.

Reprints or correspondence: Dr. J. A. Schouten, Centre for Quality of Care Research (KWAZO 229), Radboud University Nijmegen Medical Centre, P.O. Box 9101, Nijmegen, The Netherlands (J.Schouten@aig.umcn.nl).

Clinical Infectious Diseases 2005; 41:450–60

1058-4838/2005/4104-0005$15.00

the management of CAP, but the results have often varied in terms of their relevance, scientific soundness, and interpretability [10]. To our knowledge, no quality indicators have been suggested for antibiotic use in hospitalized patients with AECB. Where possible, indicators should be based directly on scientific evidence. However, like in many fields of medicine, there is only a limited scientific basis for recommendations regarding antibiotic use in cases of LRTI [10]. To develop valid quality indicators, it is necessary to use a systematic procedure that combines available evidence and expert opinion to assess additional aspects of care for which evidence alone is insufficient, absent, or methodologically weak [11, 12].

Development of valid quality indicators is, however, not enough. Validity in itself does not guarantee applicability in a specific setting. To assess applicability, again, a rigorous approach is required: indicators should be tested on important clinimetric characteristics, such as feasibility, reliability, opportunity for improvement, and case-mix stability [13].

In this article, we describe how we systematically developed a valid set of quality indicators for antibiotic use in hospitalized patients with LRTI. In addition, we describe a method to assess their applicability in daily practice. We used our quality-improvement project for antibiotic use in LRTI in 8 Dutch hospitals as a test case.

METHODS

A set of indicators was developed in 4 steps. Applicability was tested in 5 steps. Figure 1 shows a flowchart detailing the steps involved in developing the indicators and testing their applicability.

Development of a Valid Set of Indicators

Preselection of potential indicators. Four independent investigators (J.S., M.H., and 2 guideline experts) preselected key recommendations from national guidelines [6, 7]. Quality indicators that were published in international guidelines for CAP [2–5] or in the literature [10, 14–33] were added to the list of potential indicators (table 1 and table A1 in the Appendix). For the latter purpose, a literature search was performed (table 2).

Evidence-based assessment of the indicators. Every potential indicator was investigated to determine the degree of scientific evidence that linked indicator performance to outcome (i.e., mortality, morbidity, length of hospitalization, and cost-effectiveness). We started by reviewing whether the source (guideline or literature) of the potential indicator specified any references. A search of the PubMed database was then performed using search terms specific to the quality indicator topic. On the basis of the available literature, all of the potential indicators were given 1 of 4 grades (A–D) of supporting evidence (table A2 in the Appendix). Potential indicators with contradictory evidence were excluded, and grade A recommendations were immediately accepted as valid (i.e., evidence-based) indicators. The remaining indicators (grades B, C, or D) were tested further in an expert consensus procedure.

Rating and adding procedure by an expert panel. A panel of 11 opinion leaders in medical microbiology, infectious diseases, respiratory medicine, and quality-of-care medicine were asked to conduct a consensus procedure for the preselected set of indicators [11, 34]. In the 2-round consensus procedure, the panel judged the potential indicators on the basis of 3 criteria: (1) clinical relevance to the patient health benefit, (2) relevance to reducing antimicrobial resistance, and (3) relevance to cost-effectiveness. A 5-category Likert scale was used that varied from "completely disagree" (category 1) to "completely agree" (category 5). An extra answer category could be marked if the expert could not decide about a particular question. A definition of these constructs was provided in the covering letter. In the second round, the expert panel had the opportunity to comment on the proposed indicators and to add or modify potential indicators for evaluation.

Only indicators with >70% agreement between the experts on 1 criterion were selected in the first round. Indicators with >70% disagreement on all 3 criteria were rejected [34]. All of the other indicators, including those added or modified by the experts, were reevaluated in the second round.

Final set of indicators. The final step in devising the set of indicators consisted of operationalizing them by defining numerators and denominators. An algorithm for every indicator revealed how it had been deduced from the available data.

Assessment of Applicability of Quality Indicators in a Specific Patient Sample

Setting and study population. To test applicability in a specific setting, feasibility of data collection, reliability, opportunity for improvement, and case-mix stability were determined in a demonstration data set (Dutch LRTI quality-improvement project). A prospective observational audit was performed at 8 medium-sized hospitals, including both teaching and nonteaching facilities, in the southeastern part of The Netherlands. Patients with CAP and AECB were selected on the basis of formal inclusion criteria.

Data collection. During a 6-month period, trained research assistants made twice-weekly reviews of the charts of all of the patients admitted to internal medicine and respiratory medicine hospital wards. All of the relevant patients were followed-up during their period of hospitalization and until 30 days after discharge from the hospital. Data were collected from admission sheets, medical and nursing records, medication charts, and microbiological and radiological testing reports. After recording data on the preprinted standardized data forms, 2 assistants entered the results into a database.

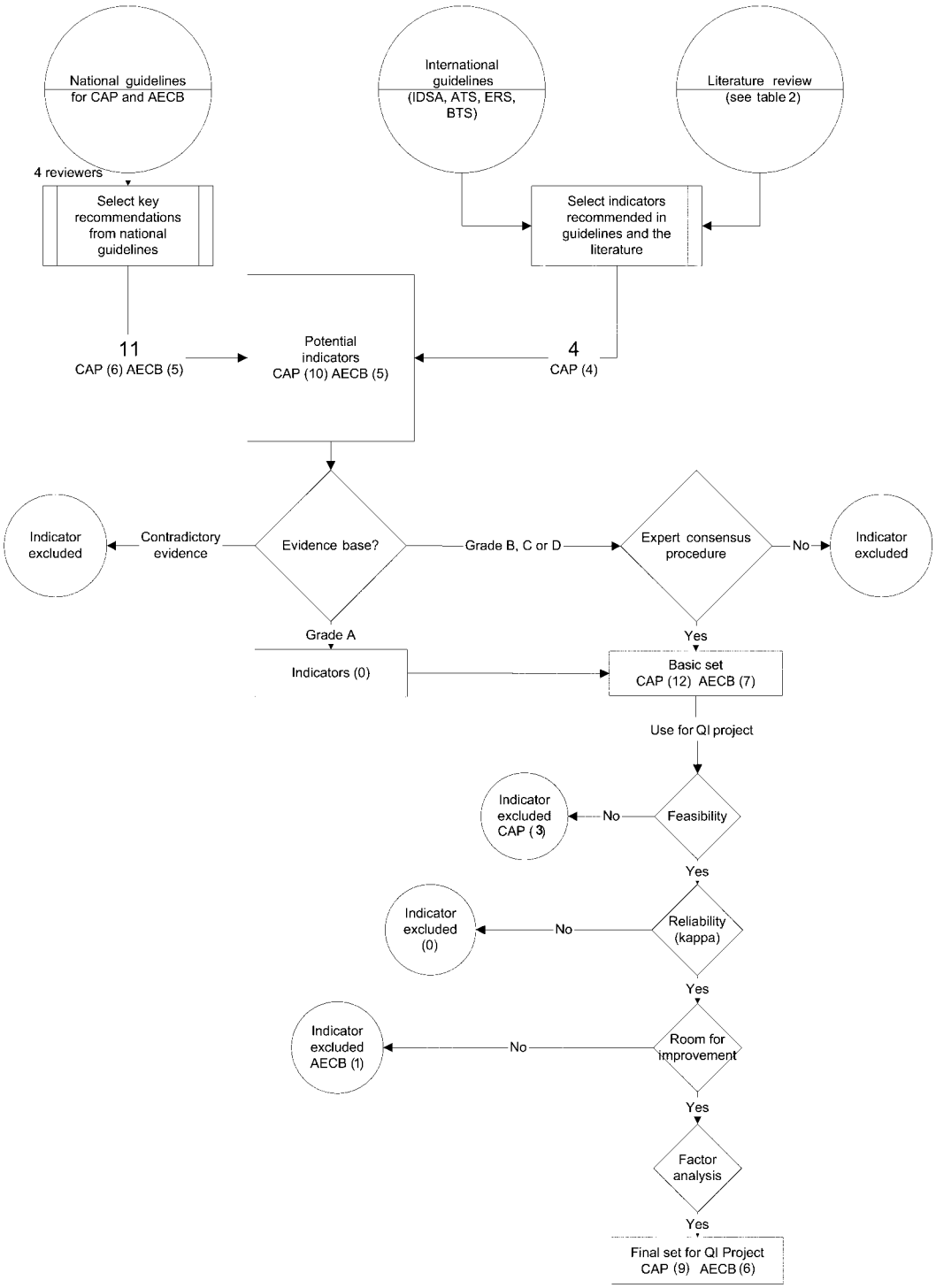

Figure 1. Procedural flowchart showing the steps involved in the development of indicators for assessing and improving the quality of antibiotic use in patients with community-acquired pneumonia (CAP) and acute exacerbation of chronic bronchitis or chronic obstructive pulmonary disease (AECB), as well as the steps involved in testing the applicability of those indicators. ATS, American Thoracic Society; BTS, British Thoracic Society; ERS, European Respiratory Society; IDSA, Infectious Diseases Society of America; QI, quality improvement.

Table 1. Rating and adding procedure for the development of quality indicators for antibiotic use in lower respiratory tract infection.

Disease, recommendation	Supporting evidence[a]	Selection round	
		First round	Second round
CAP			
No. of recommendations selected		4	8
1. Initiate antibiotic therapy <4 h after presentation	B [22, 38]	Selected	
2. Include coverage of *Legionella* species in empirical antibiotic therapy for severe CAP	B [44–49]	Added	Rejected
3. Prescribe empirical antibiotic therapy in adherence with national guidelines	B [27, 30, 39]	Selected	
4. Adapt dose and dose interval of antibiotics to renal function	D	Added	Selected
5. Switch from intravenous to oral antibiotic therapy according to existing criteria and clinical stability	B [40, 41]	No decision	Selected
6. Change broad-spectrum empirical therapy to pathogen-directed therapy as soon as culture results become available	C [3, 6]	Selected	
7. Stop antibiotic therapy if no fever for 3 days	D	Added	Selected
8. Change antibiotic therapy if no clinical improvement within 72 h of initiation	D	Added	Selected
9. Perform Gram stain and culture of a sputum sample	D [3, 6]	Selected	
10. Perform culture of 2 blood samples	B [50, 51]	No decision	Selected
11. Perform cultures <24 h after presentation	B [22]	Changed to rec. 12	
12. Perform blood cultures <24 h after presentation	B [22]	Modified from rec. 11	Rejected
13. Perform cultures before empirical therapy	B [22]	Changed to recs. 14 and 15	
14. Perform 2 blood cultures before empirical therapy	B [22]	Modified from rec. 13	Selected
15. Perform Gram stain and culture of sputum sample before empirical therapy	D [6]	Modified from rec. 13	Selected
16. Perform serological tests for atypical microorganisms on clinical suspicion	D [3, 6]	No decision	Rejected
17. Perform urine antigen testing against *Legionella* species on clinical suspicion	B [52]	Added	Selected
Exacerbation of CB or COPD			
No. of recommendations selected		4	3
1. Prescribe antibiotic therapy for exacerbation of CB or COPD only when indicated	B [53, 54]	Selected	
2. Do not use macrolide as first choice of antibiotic	B [55, 56]	Selected	
3. Adapt dose and dose interval of antibiotics to renal function	D	Added	Selected
4. Switch from intravenous to oral antibiotic therapy according to existing criteria and clinical stability	D [7]	Selected	
5. Change broad-spectrum empirical therapy to pathogen-directed therapy as soon as culture results become available	D [7]	No decision	Selected
6. Perform Gram stain and culture of a sputum sample	D [7]	Selected	
7. Optimal duration of antibiotic therapy should be 5–7 days	D	Added	Selected

NOTE. CAP, community-acquired pneumonia; CB, chronic bronchitis; COPD, chronic obstructive pulmonary disease.

[a] Supporting evidence is rated on a scale of grade A–grade D (see table A2 in the Appendix).

Applicability steps and analysis. Clinimetric characteristics of quality indicators—including feasibility, reliability, opportunity for improvement, and case-mix stability—were defined and determined in the demonstration data set. In addition, factor analysis was performed for data (i.e., indicator) reduction purposes.

Feasibility. Feasibility of data abstraction was defined as the percentage of missing values per indicator (i.e., the percentage of indicator values that could not be calculated because $\geqslant$1 element of the algorithm could not be retrieved from the available records). Feasibility was considered poor if this percentage exceeded 25%.

Reliability. To assess the reliability of our data collection, the percentage of agreement between 2 data reviewers on the level of indicator outcome, corrected for chance, was expressed in κ coefficients. A sample consisting of 10% of the records of 2 hospitals was collected by 2 independent data reviewers. Scores of $0.41 \leqslant \kappa \leqslant 0.6$ were considered to be moderate, $0.61 \leqslant \kappa \leqslant 0.8$ were considered to be good, and $\kappa > 0.8$ were considered to be very good [35]. Values of <0.4 were considered to be poor and led to elimination of the indicator.

Potential opportunity for quality improvement. Quality measures must be capable of detecting changes in the quality of care to discriminate between and within subjects. If indicator performance is invariably high, with little variation, this renders an indicator less sensitive and thus less successful as an indicator. From the viewpoint of internal quality improvement, indicators with a performance score >85% were defined as having limited room for improvement. Indicators with a performance score >85% in all participating hospitals were not selected [8, 36].

Case-mix stability. Case-mix stability is an important in-

Table 2. Summary of a systematic literature search for quality indicators for antibiotic use in the treatment of community-acquired pneumonia (CAP) and acute exacerbation of chronic bronchitis or chronic obstructive pulmonary disease (COPD).

Methods
Medical subject heading terms used were "quality of care," "quality," "performance," "recommended processes of care," "indicators," "antibiotic use," "antibiotics" combined with "community-acquired pneumonia," "pneumonia," and "CAP" or combined with "exacerbation," "exacerbation COPD," and "exacerbation of chronic bronchitis."
Original peer-reviewed articles were selected between January 1990 and July 2004 from the PubMed database.
Proposed indicators could cover the entire process of care from admission to discharge, but had to relate in some way to antibiotic use. Diagnostic procedures were included as long as they were likely to influence the use of antibiotics (i.e., choice of drug, dosage, or duration of therapy).
Excluded articles included all articles with children as study subjects, articles that did not specifically address CAP or exacerbation of COPD, articles that did not address antibiotic use, articles that were letters or case reports, articles that addressed patients with HIV/AIDS, and articles that only addressed patient outcome indicators (e.g., mortality and length of hospitalization).
Results
The initial search yielded 177 articles involving CAP; 156 of these studies were excluded on the basis of titles and abstracts. Reasons for exclusion were as follows: article did not specifically address CAP (77 articles); article did not address indicators or only addressed outcome indicators for CAP (59); article addressed patients with HIV/AIDS (7); article had children as study subjects (6); and article was letter or case report (7). The remaining 21 articles [10, 14–33] were reviewed for indicators. From these studies, 4 additional potential indicators (i.e., indicators that had not already been derived from national guidelines) were added to the list. These included initiation of antibiotic use <4 h after presentation, prescription of antibiotic therapy in adherence with national guidelines, performance of blood and sputum cultures <24 h after presentation, and performance of blood and sputum cultures before empirical therapy.
No relevant articles were found for acute exacerbation of chronic bronchitis or COPD that provided indicators for antibiotic use.

dicator asset, enabling application of an indicator to monitor quality in a specific hospital over time and to compare hospitals of different sizes and settings [13]. The relationship between certain patient characteristics and the indicator result was analyzed to decide whether correction for case mix was necessary. In the CAP indicators, we studied the distribution of outcome according to age (either $\geq$70 years or <70 years), sex, and Pneumonia Severity Index [37] (either $\leq$III or >III). For AECB, no validated severity-of-illness score was available, so we used the most recent forced expiratory volume in 1 second (FEV_1) value (expressed as a percentage of the predicted value) as a substitute. The need for case-mix correction did not lead to exclusion from our final set.

Factor analysis. Factor analysis was performed to detect relationships between indicators, thus potentially leading to a reduction in the number of indicators [36]. To perform data reduction through factor analysis, a minimum of correlation between the items is required. We used Bartlett's sphericity test (in which P should be <.05), Kaiser-Maier-Olkin measure of sampling adequacy (MSA) (in which MSA should be >0.5), and R^2 (in which R^2 should be >0.20). This procedure was performed separately for the sets of CAP and AECB indicators.

RESULTS

Development of a valid set of indicators. In the first step, 10 potential indicators for CAP and 5 for AECB were preselected from national and international guidelines and the literature (figure 1 and table 2). No "good supporting evidence" (grade A) could be found that linked process to outcome in any of these indicators. None of the indicators had to be excluded because of contradictory evidence (table 1).

All 15 potential indicators were entered into the iterated consensus procedure. In the first round, 4 recommendations were immediately selected for both CAP and AECB. No indicators were eliminated. Six new items (4 for CAP and 2 for AECB) were added by the expert panel. One potential indicator was revised, and another was split into 2 separate indicators. In the second round, 8 potential indicators (5 for CAP and 3 for AECB) out of 14 (11 for CAP and 3 for AECB) remaining recommendations were selected. The ultimate set was considered to be valid. It consisted of 12 indicators for CAP and 7 indicators for AECB.

Assessment of applicability of quality indicators in a specific patient sample. All 19 validated indicators were tested using a sample of 443 hospitalized patients with CAP and 456 hospitalized patients with AECB (tables 3 and 4). A review was made of the distribution of performance of the indicators over the 8 hospitals. Although there was wide variability in outcome (table 3), this was not because of the patient mix (data not shown).

Feasibility. Feasibility of 3 indicators was poor in our demonstration data set of Dutch patients with LRTI. Performing cultures of blood and sputum samples before the first antibiotic dose was administered showed poor feasibility: 55% of the subjects had missing values for timely culturing of blood samples, and 75% had missing values for timely culturing of sputum samples. These 2 indicators were rejected. It proved to be impossible to construct an algorithm for 1 of the 12 indicators for CAP ("change antibiotic therapy if no clinical improvement within 72 h of initiation"). We were unable to operationalize "no clinical improvement." This indicator was considered to be nonfeasible (100% of the subjects had missing values) and was also rejected.

Reliability. One indicator (timeliness of antibiotic admin-

Table 3. Applicability of quality indicators for antibiotic use in 443 patients hospitalized with community-acquired pneumonia in 8 Dutch hospitals.

Indicator	Feasibility, % of patients with missing values	Opportunity for improvement, % (range)[a]	Interobserver reliability, κ	Case-mix correction
Prescribe empirical therapy adherent to national guidelines	0.7	45 (5–59)	0.7	No
Initiate antibiotic therapy within 4 h after presentation	21[b]	68 (35–87)	0.5	No
Adapt dose and dosing interval of antibiotics to renal function	17	77 (40–100)	1	Yes[c]
Switch from intravenous to oral antibiotic therapy according to existing criteria and clinical stability	5	81 (35–93)	0.7	No
Change broad-spectrum empirical therapy to pathogen-directed therapy as soon as culture results become available	8	80 (50–100)	0.7	No
Change antibiotic therapy if no clinical improvement occurs within 72 h after initiation	100	NA[d]	NA[d]	NA[d]
Stop antibiotic therapy 3 days after defervecence	18	11 (0–41)	1	No
Perform Gram stain and culture of a sputum sample	0	54 (20–63)	1	Yes[e]
Perform culture of 2 blood samples	0	57 (48–67)	0.6	Yes[c]
Perform Gram stain and culture of a sputum sample before empirical therapy	75	24 (0–100)	0.8	NA
Perform culture of 2 blood samples before empirical therapy	55	85 (70–100)	0.8	NA
Perform a urine antigen test against *Legionella* species on clinica suspicionl	0	84 (67–100)	ND	No

NOTE. NA, not applicable; ND, not enough data available.

[a] Opportunity for improvement is given as the median percentage of patients for whom the indicator was performed per hospital (range).
[b] Data excludes 1 outlier hospital (see Results).
[c] Correction for age required.
[d] No algorithm could be created for this indicator.
[e] Correction for sex required.

istration) received a score of $\kappa = 0.5$, indicating moderate interobserver reliability. All other indicators showed κ scores of >0.6 (i.e., good or very good). No indicator was rejected.

Opportunity for improvement. The AECB recommendation to not prescribe macrolides as a first-choice antibiotic showed a high outcome (performed in >90% of cases) in each participating hospital, and thus it showed little room for improvement. This indicator was rejected for use as a quality indicator in our quality-improvement project. For the AECB indicator "adapt dose and dose interval to renal function," there was a high median performance rate (96%), but an outlier hospital with a performance rate of 73% of cases was detected, and the indicator was, therefore, not rejected.

Need for case-mix correction. Two CAP indicators needed correction for age: "adapting dose and dose interval of antibiotics to renal function" ($P = .0001$) and "obtaining samples for blood cultures" ($P = .003$). Regarding sex, sputum samples were obtained significantly more often from men than from women ($P = .001$). All of the other CAP indicators showed stable patterns of distribution over the 3 patient characteristics (age, sex, and severity of illness). In the population of patients with AECB, sputum cultures were performed more consistently in patients with an FEV_1 value of ⩽60% ($P = .033$). The need for case-mix correction did not lead to exclusion from our set but should be taken into account for interpretation of performance scores in our group of hospitals.

Factor analysis. For CAP and AECB, Bartlett's sphericity test ($P = .091$ and $P = .161$), Kaiser-Maier-Olkin MSA (0.463 and 0.489), and R^2 (0.147 and 0.137) indicated that no relevant correlation was detected between the indicators. Subsequently, further factor analysis, performed in an attempt to reduce the number of indicators, was not considered useful. Our set comprises intrinsically strong indicators.

DISCUSSION

On the basis of a carefully planned procedure that combined evidence and expert opinion, we developed a set of valid quality indicators for antibiotic use in hospitalized patients with LRTIs. In addition, we showed the importance of subjecting these indicators to a practice test before using them to measure and improve the quality of care in a specific setting. In our example, only a part of the valid set (15 of 19 indicators) turned out to be applicable in daily practice.

None of our potential indicators could rely on a firm body of evidence that linked process to outcome of care. "Timely administration of antibiotics" and "prescription of an empirical antibiotic regimen according to current guidelines" were consistently associated with improved survival in patients with

CAP, but this was only in observational, retrospective studies [38, 39]. Several prospective interventional trials have demonstrated that early-switch strategies are cost-effective and safe, but no randomized, controlled trials have yet confirmed these results [40, 41]. No firm associations were found between outcome and most of the other suggested indicators. Results from our expert consensus procedure demonstrated that these non–evidence-based recommendations may still be regarded as valuable by professionals. Changing from broad-spectrum to narrow-spectrum therapy once culture results become available will probably not directly affect short-term outcome for the individual patient. However, it has a theoretical effect on reducing the development of resistance, and it may thus turn out to be crucial for the outcome of future patients [42]. Unfortunately, studies that link process indicators with resistance patterns are confronted with large methodological difficulties, so it will be difficult to prove any definite relationship. Using a technique that systematically combined evidence and consensus enabled us to assess (and thus improve) a broader range of aspects than would have been possible if quality indicators had been restricted to evidence only.

Even if our set of indicators is considered to be valid, its applicability in daily practice has several other important prerequisites. In our demonstration data set of hospitalized Dutch patients with LRTI, most of the indicators showed reasonable applicability (i.e., they were found to be feasible and reliable and showed room for improvement). Unfortunately, the feasibility of data collection turned out to be poor for some indicators. These findings support our belief that Dutch hospitals do not have systematic and robust registration systems (e.g., for registering the timing of hospital procedures). This currently constitutes a major barrier against the application of these kinds of quality indicators in The Netherlands, not only for research purposes, but also for monitoring the quality of daily practice. Once Dutch hospitals are required to collect these data—for example, as part of their normal review process—documentation will probably pick up. Timing of procedures caused major feasibility problems in our example, but in other countries, the feasibility of these indicators may be very different. US hospitals, for example, readily collect data for antibiotic timing as part of their normal review and accreditation process. In the United States, however, other data collection problems may arise, jeopardizing feasibility. This underlines the importance of performing an applicability test before using indicators to measure and improve the quality of care in a specific setting.

In our applicability test, we used a performance rate of 85% of cases (for each participating hospital) as a cut-off value to exclude indicators. From the viewpoint of internal quality improvement, indicators that score >85% in all hospitals have little room for improvement. Quality measures must be capable of detecting changes to discriminate between and within subjects. If indicator performance is invariably high with little interhospital variation, this renders an indicator less sensitive and thus less successful as an indicator [36]. The main goal of subjecting our set of indicators to this criterion was to prioritize the indicators most in need of improvement in a quality-improvement project (i.e., those indicators with low performance rates and/or large interhospital variation). Using "opportunity for quality improvement" as a selection criterion is particularly important for internal quality-improvement efforts. If, on the other hand, the indicator set is to be used for accreditation purposes, for example, room for improvement might not be desirable as a selection criterion; the trend in regulating and accrediting organizations is to provide indicator sets that highlight excellent performance, as well as merely meet minimal standards [13].

Table 4. Applicability of quality indicators for antibiotic use in 456 patients hospitalized with acute exacerbation of chronic bronchitis or chronic obstructive pulmonary disease.

Indicator	Feasibility, % of patients with missing values	Opportunity for improvement, % (range)[a]	Interobserver reliability, κ	Case-mix correction
Do not use macrolide therapy as first-choice treatment for bronchitis in exacerbations of chronic obstructive pulmonary disease.	0	97 (91–100)	0.85	NA
Adapt dose and dose interval of antibiotics to renal function	7	97 (73–100)	1	NA
Administer antibiotics only on strict indication	1	50 (18–78)	1	No
Switch from intravenous to oral antibiotic therapy according to existing criteria and clinical stability	2	79 (15–100)	0.7	No
Change broad-spectrum empirical therapy to pathogen-directed therapy as soon as culture results become available	0	80 (0–100)	0.7	No
Perform Gram stain and culture of sputum sample	2	51 (11–68)	1	Yes[b]
Optimal duration of antibiotic therapy should be 5–7 days	3	26 (18–79)	1	No

[a] Opportunity for improvement is given as the median percentage of patients for whom the indicator was performed per hospital (range).
[b] Correction for forced expiratory volume in 1 second required.

Earlier sets of indicators, developed using somewhat different methodology, show many similarities to our set [10, 24, 43]. In some of these studies, clinimetric criteria, such as feasibility, reliability, and opportunity for improvement, were appraised by the clinical judgement of experts [10] and not on the basis of empirical data from real practice. However, in our experience, the feasibility of data collection is often overrated by professionals. All members of our expert panel believed that it was feasible to measure the time lag between performance of blood cultures and first antibiotic administration, but in reality, this could be done for only 25% of patients.

In summary, we developed a robust set of intrinsically strong indicators using rigorous methodology that combined the available evidence and expert opinion. Performance assessment in a practical test showed that some indicators were flawed by poor feasibility of data collection. Our experience demonstrates that, before implementation of a theoretically sound set of indicators, a practice test should be performed to assess its applicability in daily practice.

Acknowledgments

We thank B. Frijling and W. Wijnands, for independently selecting key recommendations from national guidelines; the panel of 11 experts who performed the consensus procedure (A. Cox, H. Wollersheim, E. Stobberingh, I. Gyssens, T. Casparie, W. Wijnands, Y. Hekster, B. J. Kullberg, W. Boersma, A. Schreurs, and R. Aleva); and Janine Trap, for administrative and statistical support.

Financial support. Grant support from Zon/Mw and the Dutch Department of Health.

Potential conflicts of interest. All authors: no conflicts.

APPENDIX

Table A1. International comparison of quality indicators for community-acquired pneumonia (CAP), by group or organization.

Indicator topic	Present report	Joint Commision on the Accreditation of Healthcare Organizations	Centers for Medicare and Medicaid Services	Agency for Healthcare Research and Quality (National Quality Measures Clearinghouse)
Antibiotic choice	Empirical antibiotic regimen according to current national guidelines	Initial antibiotic selection consistent with current recommendations	Initial antibiotic therapy consistent with current recommendations	Percentage of immunocompetent patients with CAP who receive an initial antibiotic regimen during the first 24 h that is consistent with current guidelines
Antibiotic timing	Timely initiation of antibiotic therapy (within 4 h after presentation)	Time from initial hospital arrival to first dose of antibiotic	Antibiotic timing, percentage of patients with pneumonia who received first dose of antibiotics within 4 h after hospital arrival	Percentage of patients who received their initial dose of antibiotics within 4 h of hospital arrival; median time from hospital arrival to administration of first antibiotic dose
Blood cultures	Obtaining 2 sets of blood samples for culture; culture of 2 blood samples before empirical therapy	Blood cultures performed prior to first antibiotic administration	Blood cultures performed prior to first antibiotic administration	Percentage of patients whose initial blood culture specimen was collected prior to the first hospital dose of antibiotics; percentage of patients who had blood cultures performed within 24 h prior to or after hospital arrival
Sputum culture	Obtaining sputum samples for Gram stain and culture; obtaining sputum samples for Gram stain and culture before administration of empirical therapy	…	…	…
Urine antigen test	Urine antigen testing against *Legionella* species on clinical suspicion	…	…	…
Oxygenation assessment	…	Oxygenation assessment within 24 h of hospital arrival	Percentage of patients who received an oxygen assessment within 24 h prior to or after hospital arrival	Percentage of patients who received oxygenation assessment within 24 h prior to or after arrival at the hospital
Smoking	…	Adult CAP smoking cessation counseling	Smoking cessation advice/ counseling	Percentage of adult patients with a history of smoking cigarettes who are given smoking cessation advice/counseling during hospital stay
Vaccination	…	Inpatients screened for and/or given pneumococcal vaccination	Pneumococcal screening and vaccination; influenza screening/ vaccination	Percentage of Medicare members aged ⩾65 years who ever received a pneumococcal vaccination; percentage of applicable patients admitted to the hospital for pneumonia who received pneumococcal immunization prior to admission; percentage of patients who were screened for pneumococcal vaccine status and were vaccinated prior to discharge, if indicated; Percentage of patients admitted to the hospital for CAP who received influenza immunization in the preceding influenza period; percentage of patients who received influenza vaccination
Changing therapy	Switching from intravenous to oral therapy, according to existing criteria and when clinically stable; changing broad-spectrum empirical therapy into pathogen-directed therapy (streamlining therapy); adapting dose and dose interval of antibiotics to renal function; stopping antibiotic therapy after 3 consecutive days of defervescence	Appropriate timing of intravenous to oral antibiotic switch[a]; excessive antibiotic use[a]	…	…
Patient outcome	…	Risk-adjusted pneumonia mortality rate[a]	…	Pneumonia mortality rate; bacterial pneumonia hospital admission rate

[a] Recommended for future core measure completion by the Joint Commision on the Accreditation of Healthcare Organizations.

Table A2. Level of supporting evidence linking indicator performance to outcome.

Evidence level	Definition	Example of a study providing the specified level of evidence
A1	A good systematic review of studies designed to answer the question of interest	Systematic review of randomized, controlled trials
A2	One or more rigorous studies designed to answer the question but not formally combined	Randomized, controlled trial
B	One or more prospective clinical studies that illuminate but do not rigorously answer the question	Prospective cohort study; underpowered or poor quality randomized, controlled trial; nonrandomized, controlled trial
C	One or more retrospective clinical studies that illuminate but do not rigorously answer the question	Audit or retrospective case-control study
D	Formal combination of expert views or other information	Delphi study; expert opinion; informal consensus

NOTE. From [57].

References

1. Ortqvist A. Treatment of community-acquired lower respiratory tract infections in adults. Eur Respir J Suppl **2002**; 36:40s–53s.
2. ERS Task Force Report. Guidelines for management of adult community-acquired lower respiratory tract infections. European Respiratory Society. Eur Respir J **1998**; 11:986–91.
3. British Thoracic Society guidelines for the management of community-acquired pneumonia in adults. Thorax **2001**; 56(Suppl 4):IV1–64.
4. Bartlett JG, Dowell SF, Mandell LA, et al. Practice guidelines for the management of community-acquired pneumonia in adults. Infectious Diseases Society of America. Clin Infect Dis **2000**; 31:347–82.
5. Niederman MS, Mandell LA, Anzueto A, et al. Guidelines for the management of adults with community-acquired pneumonia: diagnosis, assessment of severity, antimicrobial therapy, and prevention. Am J Respir Crit Care Med **2001**; 163:1730–54.
6. van Kasteren ME, Wijnands WJ, Stobberingh EE, et al. Optimization of the antibiotics policy in the Netherlands. II. SWAB guidelines for the antimicrobial therapy of community-acquired pneumonia [in Dutch]. Dutch Foundation of the Working Party on Antibiotic Policy. Ned Tijdschr Geneeskd **1998**; 142:952–6.
7. van Kasteren ME, Wijnands WJ, Stobberingh EE, et al. Optimizing of antibiotics policy in the Netherlands. III. SWAB guidelines for antimicrobial therapy in adults hospitalized with bronchitis [in Dutch]. Dutch Foundation of the Working Party on Antibiotic Policy. Ned Tijdschr Geneeskd. **1998**; 142:2512–5.
8. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. London: BMJ Publishing Group, **2004**.
9. Lawrence M, Olesen F. Indicators of quality in health care. Eur J Gen Pract **1997**; 3:103–8.
10. Rhew DC, Goetz MB, Shekelle PG. Evaluating quality indicators for patients with community-acquired pneumonia. Jt Comm J Qual Improv **2001**; 27:575–90.
11. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ **2003**; 326:816–9.
12. Campbell SM, Hann M, Hacker J, et al. Quality assessment for three common conditions in primary care: validity and reliability of review criteria developed by expert panels for angina, asthma and type 2 diabetes. Qual Saf Health Care **2002**; 11:125–30.
13. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. Int J Qual Health Care **2001**; 13:489–96.
14. Quality of care improvements for patients with pneumonia. Florida Medical Quality Assurance. Eval Health Prof **1998**; 21:514–24.
15. Battleman DS, Callahan M, Thaler HT. Rapid antibiotic delivery and appropriate antibiotic selection reduce length of hospital stay of patients with community-acquired pneumonia: link between quality of care and resource utilization. Arch Intern Med **2002**; 162:682–8.
16. Bratzler DW, Murray CK, Bumpus LJ, Moore LL. Community-acquired pneumonia in Oklahoma: characteristics and management of hospitalized Medicare beneficiaries. J Okla State Med Assoc **1996**; 89:87–92.
17. Chu LA, Bratzler DW, Lewis RJ, et al. Improving the quality of care for patients with pneumonia in very small hospitals. Arch Intern Med **2003**; 163:326–32.
18. Dedier J, Singer DE, Chang Y, et al. Processes of care, illness severity, and outcomes in the management of community-acquired pneumonia at academic hospitals. Arch Intern Med **2001**; 161:2099–104.
19. Fine JM, Fine MJ, Galusha D, et al. Patient and hospital characteristics associated with recommended processes of care for elderly patients hospitalized with pneumonia: results from the medicare quality indicator system pneumonia module. Arch Intern Med **2002**; 162:827–33.
20. Fortune G, Elder S, Jaco D, et al. Opportunities for improving the care of patients with community-acquired pneumonia. Clin Perform Qual Health Care **1996**; 4:41–3.
21. Jencks SF, Cuerdon T, Burwen DR, et al. Quality of medical care delivered to Medicare beneficiaries: a profile at state and national levels. JAMA **2000**; 284:1670–6.
22. Meehan TP, Fine MJ, Krumholz HM, et al. Quality of care, process, and outcomes in elderly patients with pneumonia. JAMA **1997**; 278: 2080–4.
23. Metersky ML, Galusha DH, Meehan TP. Improving the care of patients with community-acquired pneumonia: a multihospital collaborative QI project. Jt Comm J Qual Improv **1999**; 25:182–90.
24. Nathwani D, Williams F, Winter J, et al. Use of indicators to evaluate the quality of community-acquired pneumonia management. Clin Infect Dis **2002**; 34:318–23.
25. Rhew DC. Quality indicators for the management of pneumonia in vulnerable elders. Ann Intern Med **2001**; 135:736–43.
26. Schade CP, Cochran BF, Stephens MK. Using statewide audit and feedback to improve hospital care in West Virginia. Jt Comm J Qual Saf **2004**; 30:143–51.
27. Barlow GD, Lamping DL, Davey PG, Nathwani D. Evaluation of outcomes in community-acquired pneumonia: a guide for patients, physicians, and policy-makers. Lancet Infect Dis **2003**; 3:476–88.

28. Halm EA, Horowitz C, Silver A, et al. Limited impact of a multicenter intervention to improve the quality and efficiency of pneumonia care. Chest **2004**; 126:100–7.
29. Lawrence SJ, Shadel BN, Leet TL, et al. An intervention to improve antibiotic delivery and sputum procurement in patients hospitalized with community-acquired pneumonia. Chest **2002**; 122:913–9.
30. Malone DC, Shaban HM. Adherence to ATS guidelines for hospitalized patients with community-acquired pneumonia. Ann Pharmacother **2001**; 35:1180–5.
31. Meehan TP, Weingarten SR, Holmboe ES, et al. A statewide initiative to improve the care of hospitalized pneumonia patients: the Connecticut Pneumonia Pathway Project. Am J Med **2001**; 111:203–10.
32. Yealy DM, Auble TE, Stone RA, et al. The emergency department community-acquired pneumonia trial: methodology of a quality improvement intervention. Ann Emerg Med **2004**; 43:770–82.
33. Milo LA, Smucker W, Logue E, et al. Shoot, ready, aim: pneumonia care quality and costs in a community hospital. Am J Med Qual **2003**; 18:214–9.
34. Cantrill JA, Sibbald B, Buetow S. Indicators of the appropriateness of long-term prescribing in general practice in the United Kingdom: consensus development, face and content validity, feasibility, and reliability. Qual Health Care **1998**; 7:130–5.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics **1977**; 33:159–74.
36. Streiner D, Norman G. Health measurement scales. 2nd ed. Oxford: Oxford University Press, **1994**.
37. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med **1997**; 336:243–50.
38. Houck PM, Bratzler DW, Nsa W, et al. Timing of antibiotic administration and outcomes for medicare patients hospitalized with community-acquired pneumonia. Arch Intern Med **2004**; 164:637–44.
39. Menendez R, Ferrando D, Valles JM, Vallterra J. Influence of deviation from guidelines on the outcome of community-acquired pneumonia. Chest **2002**; 122:612–7.
40. Ramirez JA, Vargas S, Ritter GW, et al. Early switch from intravenous to oral antibiotics and early hospital discharge: a prospective observational study of 200 consecutive patients with community-acquired pneumonia. Arch Intern Med **1999**; 159:2449–54.
41. Rhew DC, Tu GS, Ofman J, et al. Early switch and early discharge strategies in patients with community-acquired pneumonia: a meta-analysis. Arch Intern Med **2001**; 161:722–27.
42. Leibovici L, Shraga I, Andreassen S. How do you choose antibiotic treatment? BMJ **1999**; 318:1614–6.
43. Nathwani D, Rubinstein E, Barlow G, Davey P. Do guidelines for community-acquired pneumonia improve the cost-effectiveness of hospital care? Clin Infect Dis **2001**; 32:728–41.
44. Gleason PP, Meehan TP, Fine JM, et al. Associations between initial antimicrobial therapy and medical outcomes for hospitalized elderly patients with pneumonia. Arch Intern Med **1999**; 159:2562–72.
45. Rello J, Catalan M, Diaz E, et al. Associations between empirical antimicrobial therapy at the hospital and mortality in patients with severe community-acquired pneumonia. Intensive Care Med **2002**; 28:1030–5.
46. Baddour LM, Yu VL, Klugman KP, et al. Combination antibiotic therapy lowers mortality among severely ill patients with pneumococcal bacteremia. Am J Respir Crit Care Med **2004**; 170:440–4.
47. Martinez FJ. Monotherapy versus dual therapy for community-acquired pneumonia in hospitalized patients. Clin Infect Dis **2004**; 38(Suppl 4):S328–40.
48. Stahl JE, Barza M, DesJardin J, et al. Effect of macrolides as part of initial empiric therapy on length of stay in patients hospitalized with community-acquired pneumonia. Arch Intern Med **1999**; 159:2576–80.
49. Waterer GW, Somes GW, Wunderink RG. Monotherapy may be suboptimal for severe bacteremic pneumococcal pneumonia. Arch Intern Med **2001**; 161:1837–42.
50. Campbell SG, Marrie TJ, Anstey R, et al. The contribution of blood cultures to the clinical management of adult patients admitted to the hospital with community-acquired pneumonia: a prospective observational study. Chest **2003**; 123:1142–50.
51. Waterer GW, Jennings SG, Wunderink RG. The impact of blood cultures on antibiotic therapy in pneumococcal pneumonia. Chest **1999**; 116:1278–81.
52. Lettinga KD, Verbon A, Weverling GJ, et al. Legionnaires' disease at a Dutch flower show: prognostic factors and impact of therapy. Emerg Infect Dis **2002**; 8:1448–54.
53. Saint S, Bent S, Vittinghoff E, Grady D. Antibiotics in chronic obstructive pulmonary disease exacerbations: a meta-analysis. JAMA **1995**; 273:957–60.
54. Anthonisen NR, Manfreda J, Warren CP, et al. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. Ann Intern Med **1987**; 106:196–204.
55. Pihlajamaki M, Kotilainen P, Kaurila T, et al. Macrolide-resistant *Streptococcus pneumoniae* and use of antimicrobial agents. Clin Infect Dis **2001**; 33:483–8.
56. SWAB (Dutch Foundation of the Working Party on Antibiotic Policy). NethMap 2004: consumption of antimicrobial agents and antimicrobial resistance among medically important bacteria in the Netherlands. Amsterdam: SWAB c/o Department of Infectious Diseases, Tropical Medicine, and AIDS, Academic Medical Centre, **2004**.
57. Guyatt GH, Sackett DL, Sinclair JC, et al. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. JAMA. **1995**; 274:1800–4.