Development of quality indicators for appropriate antibiotic use in daily hospital practice

van den Bosch, C.M.A.

Citation for published version (APA):
van den Bosch, C. M. A. (2016). Development of quality indicators for appropriate antibiotic use in daily hospital practice
Chapter 5

Applicability of Generic Quality Indicators for Appropriate Antibiotic Use in Daily Hospital Practice

Caroline M.A. van den Bosch1, Marlies E.J.L. Hulscher2, Stephanie Natsch3, Jan Wille4, Jan M. Prins1 and Suzanne E. Geerlings1

1Department of Internal Medicine, Division of Infectious Diseases, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands,
2Department of Scientific Institute for Quality of Healthcare (IQ healthcare), Radboud University Medical Center, Nijmegen, the Netherlands,
3Department of Pharmacy, Radboud University Medical Centre, Nijmegen, the Netherlands,
4Department of Centre for Infectious Diseases Epidemiology and Surveillance, The National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Submitted
Abstract

Objective
The ability to monitor the appropriateness of hospital antibiotic use is a key element of an effective antibiotic stewardship program. A set of 11 generic quality indicators (QIs) was previously developed to assess the quality of antibiotic use in hospitalized adults treated for a bacterial infection. The primary aim of the current study was to assess the clinimetric properties of these QIs (nine process and two structure indicators) in daily clinical practice.

Methods
QIs were developed following a RAND-modified Delphi procedure. An observational multicenter study included 1890 in patients treated with antibiotics for a suspected bacterial infection. In this cohort we tested the measurability, applicability, reliability, room for improvement and case mix stability of these QIs.

Results
Low applicability (≤ 10%) was found for the QIs ‘therapeutic drug monitoring’, ‘adapting antibiotics to renal function’ and ‘discontinue empirical therapy in case of lack of clinical and/or microbiological evidence of infection’. For the latter, we also found a low inter-observer agreement (kappa < 0.4). One QI showed low improvement potential. The remaining seven QIs had sound clinimetric properties. Case-mix correction was necessary for most process QIs. For all QIs, we found ample room for improvement and large variation between hospitals.

Conclusion
Establishing the clinimetric properties was essential, as four of the eleven previously selected QIs showed unsatisfactory properties in this practice test. Since the quality of antibiotic use and the process of documenting data is changing over time and may vary per country, QIs should always be tested in practice first.

Introduction

Around 30-40% of patients do not receive care based on available scientific evidence. For antibiotic treatment this means unnecessary or inaccurate use of antibiotics, which can have negative consequences for patient outcome and the development of antimicrobial resistance. To curb the development of antibiotic resistance, better use of current agents and a decrease of inappropriate antibiotic use are necessary.

Studies show that antibiotic stewardship programs can improve antibiotic use and reduce healthcare costs without negatively influencing the quality of care provided. One of the key elements of an effective stewardship program is its ability to monitor the quality of hospital antibiotic use, with the aim to set priorities and focus improvement. Monitoring the appropriateness of hospital antibiotic use can be accomplished using quality indicators (QIs).

Quality indicators (QIs) function as measurable elements for which there is evidence or consensus that they can be used to assess the appropriateness of daily antibiotic care provided. QIs can be developed using recommendations and already available QIs that are systematically extracted from international guidelines and literature. After developing QIs, assessing their feasibility in daily practice is essential before using them to measure appropriateness of antibiotic use. This can be done by testing important clinimetric characteristics, like inter-observer reliability.

QIs measuring antibiotic use in hospitals have been described in literature before, but for none of these measures the clinimetric properties were tested in practice. We reported previously on the systematic development, using the RAND-modified Delphi procedure, of generic indicators for measuring the appropriateness of antibiotic use in hospitalized adults treated for a suspected bacterial infection, see table 1. The primary aim of this current study was to assess the clinimetric properties of these generic QIs in a large number of inpatients in daily clinical practice.
Table 1. List of generic quality indicators to monitor antibiotic use for all bacterial infections in hospitalized adult patients on non-ICU departments.

<table>
<thead>
<tr>
<th>Number</th>
<th>Quality indicator</th>
<th>Numerator description</th>
<th>Denominator description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Empirical systemic antibiotic therapy should be prescribed according to the national guideline</td>
<td>Number of patients who started with empirical systemic antibiotic therapy according to the national guideline</td>
<td>Total number of patients who started with empirical systemic antibiotic therapy</td>
<td>Blood cultures should be performed between one week before start of treatment and actual start of treatment</td>
</tr>
<tr>
<td>2</td>
<td>Before starting systemic antibiotic therapy at least two sets of blood cultures should be taken</td>
<td>Number of patients in whom at least 2 sets of blood cultures were taken before systemic antibiotic therapy was started</td>
<td>Total number of patients who started with systemic antibiotic therapy</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>When starting systemic antibiotic therapy specimens for culture from suspected sites of infection should be taken as soon as possible, preferably before antibiotics are started.</td>
<td>Number of patients in whom cultures from suspected sites of infection were taken within 24 hours after the systemic antibiotics were started</td>
<td>Total number of patients who started with systemic antibiotic therapy</td>
<td>Cultures should be taken between two weeks before start of treatment and one day after start of treatment</td>
</tr>
<tr>
<td>4</td>
<td>An antibiotic plan should be documented in the case notes at the start of systemic antibiotic therapy</td>
<td>Number of patients for whom an antibiotic plan was documented in the case notes</td>
<td>Total number of patients who started with systemic antibiotic therapy</td>
<td>Antibiotic plan includes indication, and name, dosage, route and interval of administration of the antibiotic</td>
</tr>
<tr>
<td>5</td>
<td>Systemic antibiotic therapy should be switched from intravenous to oral antibiotic therapy within 48-72 hours on the basis of the clinical condition and when oral treatment is adequate</td>
<td>Number of patients with intravenous antibiotics for 48-72h, in whom changing to oral antibiotic therapy on the basis of clinical conditions was done</td>
<td>Total number of patients with intravenous antibiotics for 48-72h, in whom changing to oral antibiotic therapy on the basis of the clinical condition was indicated</td>
<td>Definitions: see Box 1</td>
</tr>
<tr>
<td>6</td>
<td>Empirical antibiotic therapy should be changed to pathogen-directed therapy if culture results become available</td>
<td>Number of patients with empirical therapy whose culture became positive and changing to pathogen-directed therapy was done correctly</td>
<td>Total number of patients with empirical systemic antibiotics, whose culture became positive.</td>
<td>Therapy is pathogen-directed if it is in accordance with resistance pattern. If possible, antibiotics should be chosen from the group of narrow spectrum antibiotics (Box 2)</td>
</tr>
<tr>
<td>7</td>
<td>Dose and dosing interval of systemic antibiotic therapy should be adapted to renal function</td>
<td>Number of patients with a compromised renal function with a dosing regimen adjusted to renal function.</td>
<td>Total number of patients who started with systemic antibiotic therapy which should be dosed according to renal function, and who had an unknown or compromised renal function</td>
<td>Compromised renal function defined as an estimated Glomerular Filtration Rate &lt; 50mL/min/1.73m²</td>
</tr>
</tbody>
</table>

Definitions: see Box 1
### Methods

**Setting, study population and data collection**

This was an observational multicenter study where we measured the applicability of the developed QIs in four university and 18 non-university (medium-sized and small) hospitals. To include a representative patient group, we identified in each hospital on non-Intensive Care Unit (ICU) departments, all adult inpatients using antibiotics for >24 hours for a suspected hospital- or community-acquired bacterial infection on the day of a point prevalence measurement. This point prevalence measurement was performed biannually by PREZIES (prevention of nosocomial infections by surveillance), a department of the RIVM (National Institute for Public Health and the Environment, the Netherlands). Patients were included in October 2011, March 2012 and October 2012. Of all included patients clinical and laboratory data of the entire antibiotic pathway, from start of antibiotics till discharge, were extracted retrospectively from medical and nursing records, and medication charts. Patients were excluded if antibiotics were used as prophylaxis, or when antibiotic treatment was started on the ICU department or in another hospital, see figure 1. All data needed to compute the QIs or to assess potential determinants needed for case mix correction (see below) were extracted by trained research nurses and the study coordinator in a uniform way and entered in the database anonymously. The ethics committee assessed the study and concluded that it was deemed exempt from their approval.

### Quality indicators and definitions

The set of 11 QIs for appropriate antibiotic use of all bacterial infections in hospitalized adult patients admitted on a non-ICU department (table 1) were developed using the systematic RAND modified Delphi method.\(^9\) QIs one to nine are process indicators, measuring the process around appropriate antibiotic use at the patient level. QIs 10 and 11 are structure indicators, which measure requirements for appropriate antibiotic use at the hospital level. The QIs with their numerators and denominators are described in detail in Table 1. For some of the QIs, we had to develop more extensive working definitions:

**QI.1 Prescribe empirical antibiotic therapy according to the national guideline.**

We defined empirical therapy as the antibiotic (combination) prescribed on the day of diagnosis, before identification of a causative pathogen. If
initial therapy was directed at a previously cultured pathogen, this was not called ‘empirical therapy’ and these patients were excluded for this indicator. Prescribing in accordance with the guideline was evaluated in relation to the at that moment current Dutch national guideline, http://www.swab.nl/guidelines. Individual patient characteristics, in particular allergies, pregnancy, and previous extended-spectrum β-lactamase (ESBL) infection were taken into account when computing the algorithms for appropriate empirical therapy.

Only patients with one or two (possible) diagnoses covered by a guideline were included. Patients with more than two (possible) infections or with a (possible) infection not covered by a guideline were excluded. In case two diagnosed or possible infections affected a single organ system (e.g. upper and lower respiratory tract infection (LRTI)), it usually concerned a differential diagnosis, and empiric antibiotic therapy was considered correct if covered by one of the two applicable guidelines. If the two (possible) infections affected different organ systems (e.g. LRTI and urinary tract infection), antibiotic therapy was only considered correct if both diagnoses were covered according to their respective guidelines. Two infectious diseases specialists assessed the algorithms regarding guideline adherent antibiotic therapy for combinations of diagnoses.

**QI.5 Antibiotic therapy should be switched from intravenous to oral therapy within 48 – 72 hours.**

This QI was applicable to the subgroup of patients who started with intravenous (iv) antibiotic treatment and fulfilled the criteria for safe early switch, see Box 1. Clinical stability at 48 – 72 hours of antibiotic treatment was also considered to have been present when temperature or blood pressure was not mentioned in the medical records or white blood cells (WBC) had not been determined, because we assumed that if these parameters were abnormal, they would have been documented.

**QI.6 Empirical antibiotic therapy should be changed to pathogen-directed therapy if culture results become available.**

This QI applied to patients with a positive culture (denominator), and it measured the proportion of patients (numerator) that received pathogen-directed antibiotic treatment after culture results became positive. We estimated conservatively that in general at day 5 culture results are available. Therefore, the antibiotic (combination) given at day 5 was considered the final therapy. This therapy was considered pathogen-directed when it was in accordance with the resistance pattern of the cultured microorganism, regardless whether antibiotic therapy was changed or not. If possible, antibiotics should be chosen from a group of recommended (‘more’ narrow) antibiotics (Box 2).**

**QI.7 Dose and dosing interval of antibiotic therapy should be adapted to renal function.**

Creatinine clearance was measured using the Modification of Diet in Renal Disease (MDRD) formula. Antibiotics were grouped according to necessity to

---

**Figure 1. Patients Identified During the Point Prevalence Measurements**
Box 1.

**QI 5 Criteria for safe switch within 2-3 days**

1. Hemodynamic stable
2. No fever
3. Normal WBC count (4 – 12 x10⁹/L)
4. Patient must be able to take oral medication
5. Patient must have a functioning gastrointestinal tract, without signs of malabsorption
6. Safe switch **not** possible in case of: meningitis, intracranial abscess, endocarditis, mediastinitis, *Legionella* pneumonia and exacerbations of cystic fibrosis, inadequately drained abscess and empyema, severe soft tissue infection such as group A streptococcal infections, infection of foreign bodies, *Staphylococcus aureus* or *Pseudomonas aeruginosa* bacteremia, liver abscess and empyema, osteomyelitis and arthritis, chemotherapy related neutropenia.

Box 2.

**QI 6 Recommended (more narrow) antibiotics**

- Penicillin
- Ciprofloxacin
- Phenethicillin
- Norfloxacin
- Amoxicillin
- Ofloxacin
- Flucloxacillin
- Doxycycline
- Cephalexin
- Tetracycline
- Cephalothin
- Minocycline
- Cefazolin
- Nitrofurantoin
- Erythromycin
- Fosfomycin
- Co-trimoxazole
- Metronidazole
- Trimethoprim

Box 3.

**QI 9 Criteria for discontinuing empirical antibiotic therapy**

1. Antibiotic therapy was started empirical
2. Blood cultures and cultures from suspected site of infection have been performed
3. All cultures remained negative
4. Haemodynamic stable 2 – 3 days after start treatment
5. WBC count normal (4 – 12 x10⁹/L) 2 – 3 days after start treatment
6. No fever at day 6 or 7
7. No antibiotics in the 30 days before start of empirical treatment

Adapt the dosing regimen in case of a compromised renal function: no adaption to renal function necessary, or adaption necessary with estimated Glomerular Filtration Rate (eGFR) 30-50 mL/min/1.72m², 10-30 mL/min/1.72m², or <10 mL/min/1.72m². In patients using an antibiotic that should be dosed according to renal function, the plasma creatinine should therefore be checked sometime between 3 days before and 2 days after the start of antibiotics.

**QI.11 Local antibiotic guidelines should correspond to the national antibiotic guidelines, but should deviate based on local resistance patterns.**

We compared the local guidelines with the national guidelines in the Netherlands (http://www.swab.nl/guidelines). Since between various regions in the Netherlands minimal differences in local resistance rates exist, deviation based on local resistance patterns is in our setting not an issue. For the same reason we evaluated in QI.1 empirical antibiotic therapy according to the national guidelines, instead of the local guideline.

**Assessment of the clinimetric properties of the Quality indicators**

The denominator and numerator of the QIs were translated into computerized algorithms in SPSS syntax, in which data from clinical patient records were entered, to determine for every QI whether a patient qualified for the denominator and to estimate the numerator.

**Measurability** is defined as the availability of administrative data required to evaluate the indicator. An indicator was considered measurable if data necessary to score the indicator could be abstracted from the available data for >75% of the cases.

**Applicability.** A QI should be applicable to at least 10% of the reviewed patient records (with a minimum of 10 patients per hospital).

**Inter-observer reliability** (only estimated in process QIs). A second investigator rated 50 records from different departments in four medical centres (two university- and two non-university hospitals, two with and two without Electronic Patient Record System [EPRS]). A second investigator rated 50 records from different departments in four medical centres (two university- and two non-university hospitals, two with and two without Electronic Patient Record System [EPRS]). The percentage of agreement on the level of indicator outcome was calculated and expressed in Kappa ($\kappa$) coefficients. Scores of $\kappa >0.6$ were considered to be good, $\kappa$ values $<0.4$ unreliable.

**Potential room for improvement** measures the sensitivity of a potential indicator to detect variability in quality of care between and within hospitals. It is expressed as 100% minus the performance score, with performance...
expressing the percentage of adherence to an indicator. High QI scores with little variation between hospitals make indicators less sensitive and therefore less useful in daily practice, so the improvement potential was considered ‘low’ if the performance scores of all hospitals were >85%.

**Case mix stability** (only estimated in process QIs). For indicators that showed sound clinimetric properties, case-mix stability was analysed, to determine whether an indicator can monitor quality in a specific hospital over time, and to compare hospitals of different sizes and settings. The effects on the indicator scores of the following potential determinants were studied: sex, age, community- versus hospital-acquired infection, sepsis (SIRS criteria ≥2) versus no sepsis, severity of illness (MEWS score), co morbidity (Charlson Co morbidity Index), prior use of antibiotics and iv or oral start of antibiotics. The need for case-mix correction did not lead to elimination of the indicator.

**Correlation between performance scores of the Quality indicators**

We finally examined the relationships between the performance scores of the process QIs, with the ultimate aim of reducing the number of QIs, and therefore the required data collection necessary to assess ‘appropriate antibiotic use’. A correlation coefficient > 0.4 was considered to represent a moderate correlation between QIs.

**Analysis**

Collected data were entered in a database using the Statistical Package for the Social Sciences (SPSS 20·0 for Windows®, SPSS Inc., Chicago, IL, USA). Descriptive analyses were performed for each indicator. Performance was expressed as percentage of adherence to an indicator. Case-mix stability was analysed by a multilevel (mixed model) analysis. With these analyses we take into account the hierarchical structure: patients nested within hospitals. We performed a mixed model with a random intercept and all other variables fixed. For examining the relationship between the process QIs, we used the bivariate Spearman correlation coefficient.

**Role of the funding source**

The design and conduct of the study, the collection, analysis and interpretation of the data were funded by the Netherlands Organisation for Health Research and Development (ZonMW) [grant number 205100003].

The funders had no role in the preparation, review, or approval of the manuscript, and decision to submit the manuscript for publication.

**Results**

**Patients**

The QIs were tested in a sample of 1890 adult patients, admitted to non-ICU departments of 22 hospitals in the Netherlands and treated for a suspected bacterial infection. Of the 1890 patients, 51.5% were men, and 73.6% of the suspected bacterial infections were community-acquired. Further baseline characteristics are shown in table 2.

**Clinimetric properties of the QIs**

**Measurability:** Some indicators had more missing values than others, but for all QIs measurability was good (Table 3).

**Applicability:** QI.7, adapting antibiotics to renal function, applied to only 8.7% of patients, QI.8, on Therapeutic Drug Monitoring, applied to 5.8% of patients and QI.9, regarding discontinuing empirical antibiotic therapy, applied to 8.6%.

**Inter-observer reliability:** QI.9, discontinuing empirical antibiotic therapy, had a kappa of 0.24, indicating very low inter-observer reliability, partly because the impossibility to design a good algorithm for ‘lack of clinical evidence of infection’ left it subject to personal interpretation. All other indicators scored kappas >0.6, suggesting good reliability.

**Potential room for improvement:** Structure QI.10 states that a current local guide should be present in the hospital. Ninety-five percent of the hospitals had a local, recent guide, which leaves little room for improvement.

**Case-mix stability:** Multilevel analysis showed that in five of the six QIs one or more determinants significantly influenced indicator scores. The most important variables that influenced indicator scores were hospital- versus community-acquired infection, co morbidity, severity of illness and start of i.v. versus oral antibiotic treatment. The required corrections per QI, and all other clinimetric properties are shown in Table 3.
Correlation between QIs

The highest correlation coefficient, 0.17, was found between QI5 (oral switching) and QI8 (streamlining) (data not shown). So, no strong correlations between QIs were found and reducing the number of QIs is therefore not possible.

Table 2. Baseline characteristics of the patients (n = 1890)

<table>
<thead>
<tr>
<th>Sex</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>974 (51.5)</td>
<td>916 (48.5)</td>
</tr>
</tbody>
</table>

| Age, mean (SD) | 69·8 (17) |

<table>
<thead>
<tr>
<th>Infection</th>
<th>Community-acquired</th>
<th>Hospital-acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1391 (74)</td>
<td>499 (26)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Lower respiratory tract infection</th>
<th>Skin and soft tissue infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>177 (9)</td>
<td>17 (1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 diagnoses possible at start of antibiotic treatment</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1180 (62)</td>
<td>710 (38)</td>
</tr>
<tr>
<td>Ye</td>
<td>774 (41)</td>
<td>1116 (59)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SIRS criteria</th>
<th>≥ 2 (sepsis)</th>
<th>&lt; 2 (no sepsis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>774 (41)</td>
<td>1116 (59)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antibiotics</th>
<th>Started intravenously</th>
<th>Started orally</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1399 (74)†</td>
<td>489 (26)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Received antibiotics before start of (empirical) treatment</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1716 (91)†</td>
<td>174 (9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antibiotics started orally</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1186 (63)</td>
<td>606 (32)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antibiotics received before start of (empirical) treatment</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1180 (62)</td>
<td>710 (38)</td>
</tr>
</tbody>
</table>

Table 3. Clinimetric properties of the quality indicators

<table>
<thead>
<tr>
<th>INDICATOR</th>
<th>Measurability; missings % (n)</th>
<th>Applicability % (n)</th>
<th>Interobserver reliability (Kappa)</th>
<th>Improvement potential (%) (100-performance)</th>
<th>Case mix stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prescribe empirical antibiotic therapy according to guideline</td>
<td>1·1 (20)</td>
<td>73 (1361)</td>
<td>0·75</td>
<td>59</td>
<td>NoA,B,C</td>
</tr>
<tr>
<td>2. Before starting antibiotic therapy two sets of blood cultures should be taken</td>
<td>0</td>
<td>100 (1890)</td>
<td>0·74</td>
<td>64</td>
<td>NoA,D,E,F,G</td>
</tr>
<tr>
<td>3. When starting antibiotic therapy cultures should be taken from suspected sites of infection</td>
<td>0·9 (17)</td>
<td>65 (1217)</td>
<td>0·74</td>
<td>51</td>
<td>NoA,B,C,G,H</td>
</tr>
<tr>
<td>4. An antibiotic plan should be documented in the case notes at the start of antibiotic therapy</td>
<td>0</td>
<td>100 (1890)</td>
<td>0·77</td>
<td>39</td>
<td>NoA,B,C,G,H</td>
</tr>
<tr>
<td>5. Antibiotic therapy should be switched from intravenous to oral therapy within 48 – 72 hours</td>
<td>1 (19)</td>
<td>23 (422)</td>
<td>0·90</td>
<td>68</td>
<td>No</td>
</tr>
<tr>
<td>6. Empirical antibiotic therapy should be changed to pathogen-directed therapy if culture results become available</td>
<td>0·2 (3)</td>
<td>24 (435)</td>
<td>0·94</td>
<td>50</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Dose and dosing interval of antibiotic therapy should be adapted to renal function</td>
<td>0</td>
<td>9 (165)</td>
<td>0·87</td>
<td>62</td>
<td>-</td>
</tr>
<tr>
<td>8. Therapeutic drug monitoring should be performed when the therapy duration is longer than 3 days for aminoglycosides and 5 days for vancomycin</td>
<td>0·1 (2)</td>
<td>6 (112)</td>
<td>0·79</td>
<td>21</td>
<td>-</td>
</tr>
<tr>
<td>9. Empirical therapy should be discontinued in case of lack of clinical and/or microbiological evidence of infection. Maximum duration of therapy is 7 days</td>
<td>0·2 (3)</td>
<td>9 (162)</td>
<td>0·24</td>
<td>67</td>
<td>-</td>
</tr>
<tr>
<td>10. A current local antibiotic guideline should be present in the hospital, and an update should be done every 3 years</td>
<td>14 (3)</td>
<td>100 (19)</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>11. Local antibiotic guidelines should correspond to the national antibiotic guidelines, but should be adapted based on local resistance patterns</td>
<td>9 (2)</td>
<td>100 (20)</td>
<td>-</td>
<td>100</td>
<td>-</td>
</tr>
</tbody>
</table>

A Correction required for (A) age, (B) comorbidity (Charlson Index), (C) hospital- vs. community-acquired infection, (D) severity of illness (MEWS score), (E) sepsis versus no sepsis, (F) start antibiotic treatment IV versus orally.
Performance scores
The mean performance scores for each QI are shown in Table 4. The highest score (61%) was found for QI.4 (antibiotic plan should be documented). QI.11 (local guidelines should correspond to the national guidelines) had the lowest score: 0%. The mean performance score for all QIs together was 38%. For each QI we established a large variation between hospitals.

Table 4. Performance of the validated QIs in the 22 Dutch hospitals.

<table>
<thead>
<tr>
<th>INDICATOR</th>
<th>n</th>
<th>Performance % (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prescribe empirical antibiotic therapy according to guideline</td>
<td>563</td>
<td>41 (24 – 58)</td>
</tr>
<tr>
<td>2. Before starting antibiotic therapy two sets of blood cultures should be taken</td>
<td>674</td>
<td>36 (9 – 59)</td>
</tr>
<tr>
<td>3. When starting antibiotic therapy cultures should be taken from suspected sites of infection</td>
<td>595</td>
<td>49 (33 – 73)</td>
</tr>
<tr>
<td>4. An antibiotic plan should be documented in the case notes at the start of antibiotic therapy</td>
<td>1145</td>
<td>61 (23 – 98)</td>
</tr>
<tr>
<td>5. Antibiotic therapy should be switched from intravenous to oral therapy within 48 – 72 hours</td>
<td>134</td>
<td>32 (5 – 50)</td>
</tr>
<tr>
<td>6. Empirical antibiotic therapy should be changed to pathogen-directed therapy if culture results become available</td>
<td>228</td>
<td>50 (21 – 85)</td>
</tr>
<tr>
<td>11. Local antibiotic guidelines should correspond to the national antibiotic guidelines, but should deviate based on local resistance patterns</td>
<td>0</td>
<td>0 (63 – 94*)</td>
</tr>
</tbody>
</table>

*not one hospital had local guidelines that corresponded to the national guidelines completely, so performance was 0%, but the % of overlap between the local and the national guidelines ranged from 63 – 94%, with a mean of 80%.

Discussion
We demonstrated the importance of testing the clinimetric properties of previously developed QIs before using them, for example in antibiotic stewardship programs. In this practice test in 22 hospitals in the Netherlands, seven of the eleven generic QIs showed sound clinimetric properties. We also demonstrated that there was no relevant correlation between the process QIs, precluding a reduction in total number of QIs and data required. This study also showed low adherence to the QIs, with a wide performance range across hospitals. Four indicators scored below 50% adherence. This means a large room for improvement concerning the quality of antibiotic treatment in Dutch hospitals.

Although some QIs showed no sound clinimetric properties, this does not imply that these QIs are not important when assessing appropriate antibiotic use. However, they should not be used as generic QIs, but for instance in targeted audits or in specific settings (QI.7 renal function and QI.8 TDM) or after being redefined (QI.9 discontinue empiric therapy).

Our study has several strengths. To our knowledge, this is the first study to develop and test the clinimetric properties of a set of generic QIs describing appropriate antibiotic treatment for all bacterial infections in the hospital, along the entire antibiotic pathway from start to discontinuing antibiotics, on an individual patient level. Previously other studies have developed and used QIs for antibiotic use in the hospital, such as the ESAC group, but those QIs were not tested in clinical practice and e.g. inter-observer reliability was not measured. Testing the internal validity of data collected, and therefore knowing your data was uniformly collected is very important. The QIs of Thern et al. showed some overlap with our set, however most their indicators applied to specific patient groups, e.g. only to patients with community acquired infections. Our set of QIs is generic and therefore applies to all adult inpatients treated with antibiotics. Pulcini et al. developed QIs specifically for the reassessment of empirical antibiotic use on day 3, instead of the whole antibiotic pathway as measured in our study.

Secondly, the RAND modified Delphi procedure was used for the QI development and many international experts participated in this procedure. This systematic and rigorous consensus method for indicator development,
together with an applicability test in nearly 1900 patients, results in robust indicators.\textsuperscript{7-9,17,20,29}

Thirdly, all QIs showed high inter-observer reliability, except for QI.9. The overall high kappa was probably because the data were collected in a uniform and standardized way by the investigators. The interpretation was also uniform using the computerized algorithms in SPSS syntax to qualify a patient for the denominator and estimating the numerator.

This study also has potential limitations. One potential limitation could be that the QIs were developed and tested in a Dutch setting, so the results may not automatically translate to other countries. The clinimetric properties can be different for other countries. However, since the QIs were developed using international studies and with an international expert panel, we believe these indicators represent a theoretically sound set that can be used internationally.\textsuperscript{17} Nevertheless, feasibility should always be tested in practice first, because registration of data may vary between countries and sometimes even within clinical settings.

Another limitation with assessing the QI ‘prescribe according to national guideline’ was that some guideline recommendations may have been outdated. Therefore, although guidelines should be reassessed every three years,\textsuperscript{30} in some cases empirical treatment may have been scored as not correct, while in reality it was correct according to the most recent insights. We also recommend that QI ‘prescribe according to national guideline’ should be applicable to patients with one diagnosis only, because constructing algorithms regarding guideline-adherent therapy for all possible combinations of diagnoses was time-consuming and difficult.

A third limitation was the retrospective, observational design of the study, which makes it impossible to collect all relevant data. For example, not all data that influenced the physicians’ choices may have been documented properly. In addition, most Dutch hospitals still do not have a systematic and robust registration system, therefore collecting data is a time-consuming activity. The expectation is that the current introduction of EPRS may solve part of this problem over time.\textsuperscript{9} Since collecting data for process QIs is time-consuming, especially for QI.1, we explored what the consequences would be when certain patient characteristics (allergies, pregnancy, and previous ESBL infection) were not taken into account when computing the algorithm for QI.1 ‘prescribe empirical antibiotic therapy according to guideline’. Excluding these parameters only marginally decreased the mean performance rate: from 41.4\% to 40.3\%. Therefore, omitting these parameters when assessing QI.1 could be considered.

In summary, we evaluated a set of generic and robust QIs to measure and monitor the appropriateness of antibiotic use in the treatment of all bacterial infections in hospitalized adult patients. A practice test in the Netherlands demonstrated that seven out of the eleven QIs had sound clinimetric properties, and ample room for improvement with large variation between hospitals. Since the process of collecting data is changing over time and varies per country, theoretically sound QIs should always be tested in practice first. Future studies can use this article as a guide for testing the clinimetric properties of their QIs or use this reliable set as a tool to monitor the quality of hospital antibiotic use and to evaluate strategies to improve it in antibiotic stewardship programs.

Funding
This work was supported by the Netherlands Organisation for Health Research and Development (ZonMW) [grant number 205100003].

Transparency declarations
None declared by all authors
- all participating hospitals:
  Academic Medical Centre, Amsterdam
  Rijnstate Hospital, Arnhem
  Radboud University Medical Centre, Nijmegen
  University Medical Centre Utrecht, Utrecht
  Maxima Medical Centre, Veldhoven
  Gelderse Vallei, Ede
  Westfriesgasthuis, Hoorn
  Hospital Rivierenland, Tiel
  Diakonessenhuis, Utrecht
  Leiden University Medical Centre, Leiden
  Gemini Hospital, Den Helder
  Orbis Medical Centre, Sittard
  Hospital Bernhoven, Ude
  Medical Centre Haaglanden, Den Haag
  Jeroen Bosch Hospital, Den Bosch
  Vlietland Hospital, Schiedam
  Atrium Medical Centre Parkstad, Heerlen
  Franciscus Hospital, Roosendaal
  LangeLand Hospital, Zoetermeer
  Groene Hart Hospital, Gouda
  Lievensberg Hospital, Bergen op Zoom
  Medisch Spectrum Twente, Enschede

Reference List


