

## Design of a 'day 3 bundle' to improve the reassessment of inpatient empirical antibiotic prescriptions

Céline Pulcini<sup>1,2\*</sup>, Sylviane Defres<sup>3</sup>, Ila Aggarwal<sup>3</sup>, Dilip Nathwani<sup>3</sup> and Peter Davey<sup>2,3</sup>

<sup>1</sup>Service d'Infectiologie, Hôpital l'Archet 1, Centre Hospitalier Universitaire de Nice, Route St Antoine de Ginestière, BP 3079, 06202 Nice Cedex 3, France; <sup>2</sup>Informatics Section, Division of Community Health Sciences, MacKenzie Building, Kirsty Semple Way, Dundee DD2 4BF, Scotland, UK; <sup>3</sup>Infection Unit, East Block, Level 4, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland, UK

Received 8 November 2007; returned 10 January 2008; revised 19 February 2008; accepted 22 February 2008

**Objectives:** To develop and test a set of process measures of quality of care in the reassessment of inpatient empirical antibiotic prescriptions, to determine the inter-rater reliability of medical notes' review in assessment of these measures and to test these measures on one ward.

**Methods:** Measures of process of care were identified from a literature review. Forty sets of medical notes were reviewed by two independent doctors and the inter-rater reliability determined using observed percentage agreement and the kappa statistic. These measures were collected weekly and fed back to doctors in order to stimulate improvement.

**Results:** Four process measures were identified and were grouped together to create a 'day 3 bundle': antibiotic plan, review of the diagnosis, adaptation to microbiology and intravenous–oral switch. The inter-rater agreement was  $\geq 80\%$  for all measures. Data collection was feasible and was easily sustained over several weeks. The reassessment of antibiotic prescriptions around day 3 was better documented using real-time feedback of the measures to the medical team.

**Conclusions:** Our measures of care are suitable for the reassessment of empirical inpatient antibiotic prescriptions, with good inter-rater reliability. This quality intervention should be part of a more comprehensive and multifaceted plan to improve antibiotic use in hospitals.

**Keywords:** indicator, quality assurance, switch therapy, antibiotic therapy, medical decision-making, prescribing practice, hospital infections

### Introduction

Bacterial resistance is a major concern worldwide. At least one-third of antibiotics prescribed in hospitals are either unnecessary or inappropriate, and such misuse is a main driver for resistance.<sup>1</sup> Recommendations to improve antibiotic use in hospitals have been published in many countries and favour multifaceted interventions.<sup>1–3</sup> Reassessment of antibiotic prescriptions around day 3 is part of these recommendations and has been proved to trigger modification of the therapy more often and sooner.<sup>4</sup> Following the example of quality improvement initiatives for community-acquired pneumonia,<sup>5,6</sup> ventilator-acquired pneumonia,<sup>7</sup> catheter-related bacteremia,<sup>8</sup> sepsis<sup>9,10</sup> or surgical prophylaxis,<sup>11</sup> appropriate measures of

quality of care of patients receiving an antibiotic therapy are required in order to improve antibiotic use.<sup>12</sup>

The aim of our study was to develop and test process measures of quality of care that can then be used to assess and improve the reassessment of inpatient empirical antibiotic prescriptions around day 3. The intended purpose of our measures is for local improvement efforts rather than for benchmarking or use by an accreditation agency.<sup>13</sup> We followed published methods for developing and implementing a clinical performance measure.<sup>13</sup> Our study had three principal objectives: first to identify the key processes of the reassessment that need to be documented in the medical notes, second to assess the inter-rater reliability of these process measures and third to test these measures and the impact of feedback on the quality of care on one ward.

\*Correspondence address. Service d'Infectiologie, Hôpital l'Archet 1, CHU de Nice, Route de Antoine de Ginestière, BP 3079, 06202 Nice Cedex 3, France. Tel: +33-492035515; Fax: +33-493965454; E-mail pulcini.c@chu-nice.fr

## Patients and methods

### *Selection of key process measures for the reassessment of inpatient empirical antibiotic prescriptions*

We focused on the reassessment of curative empirical antibiotic therapies 24–96 h after the treatment was started, where day 1 was the first day the antibiotic was administered. Key recommendations were pre-selected by one reviewer from a literature search focusing on guidelines or recommendations on antimicrobial stewardship and original articles assessing reassessment of antibiotic therapies around day 3. Then a group of three infectious diseases (ID) physicians (C. P., D. N. and P. D.) met to select three to five measures thought to be valid and easy-to-collect.

### *Reliability of measures of processes of care from review of medical notes*

To determine the reliability of review of medical notes, we measured agreement between two independent reviewers (S. D. and I. A.). Two independent, single reviewers are thought to be sufficient for data abstraction and tests of inter-rater reliability.<sup>14</sup> Our data abstraction was carried out by internal clinical staff rather than external reviewers. Studies of data reliability have compared results between internal and external reviewers in the assessment of performance indicators used to compare standards between hospitals.<sup>15</sup> Minimal differences were demonstrated between internal and external reviewers and, where differences did exist, the internal reviewers did not always give more favourable results.<sup>15</sup> Using internal reviewers should be even less of a cause for concern for our purpose of assessing changes in standards over time in our own hospital rather than comparing results between hospitals.

The medical notes of 40 patients who received antibiotic treatment on the ID ward were reviewed. Data were extracted from the medical notes using a standard form. The review process was carried out independently by the two Registrars trained in ID, blind to each other's results. The two Registrars were also treating patients on the ID ward, and they could abstract data from their own medical notes.

The percentage agreement for each measure was calculated and the kappa statistic was used to indicate agreement adjusted for the level expected to occur by chance. The 95% confidence intervals for  $\kappa$  and the proportions of positive ( $p_{\text{pos}}$ ) and negative ( $p_{\text{neg}}$ ) agreements for each variable were calculated to assist in the interpretation of  $\kappa$ .  $p_{\text{pos}}$  and  $p_{\text{neg}}$  are of particular value in explaining the high agreement but low  $\kappa$  that occurs when the expected agreement is high.<sup>16,17</sup> Bias-adjusted  $\kappa$  and the expected agreement that could occur by chance were also calculated.<sup>16,17</sup> Statistical analyses were performed using the Diagnostic and Agreement Statistics spreadsheet.<sup>18</sup>

### *Test of feedback of process measures on improvement in the quality of care*

We then tested the impact of feedback of these process measures on one ward (the ID ward in Ninewells Hospital, Dundee, UK). There are 18 inpatient beds with ~1000 admissions annually. The medical team is made up of three ID consultants, two Registrars and three junior doctors. Five to 10 randomly selected medical notes were reviewed each week by one doctor (S. D. or I. A.) on the ward, and the results were fed back in real time to the whole medical team using run charts. Changes were implemented when needed in order to improve the reassessment of antibiotic prescriptions, using plan-do-study-act cycles.<sup>19</sup> A cycle is made up of four steps that can be carried out by one person in 1 day: (i) plan the change: list the

modifications to be done, try to predict what will happen and how you will adapt; (ii) do: carry out the intervention and identify the problems that may occur; (iii) study: collect data, summarize what was learned and which problems occurred, and discuss with the team; and (iv) act: determine what changes are to be made.

## Results

### *Selection of key process measures for the reassessment of inpatient empirical antibiotic prescriptions*

Based on our literature search,<sup>1,3,4,12,20,21</sup> four measures among seven were selected for documentation in the medical notes 24–96 h after the antibiotic course was started, thus around day 3 of treatment:

- (i) Was there an antibiotic plan (name, dose, route, interval of administration and planned duration)?
- (ii) Was there a review of the diagnosis?
- (iii) If positive microbiological results were available, was there any adaptation of the antibiotic treatment, for example streamlining or discontinuation?
- (iv) If the patient was initially started on intravenous (iv) antibiotic therapy, was the possibility of iv–oral switch documented?

A fifth measure was made up of the grouping of the preceding four measures to form a care bundle.<sup>12</sup> A care bundle is a grouping of best practices with respect to a disease process that individually improve care, but when applied together result in substantially greater improvement. Their application favours an 'all or none' approach as opposed to piecemeal measures. Bundles are dichotomous, so compliance is assessed in a simple yes/no measure.

- (v) The 'day 3 bundle' was deemed completed only if all preceding four measures were completed.

### *Reliability of measures of processes of care from medical notes' review*

Data are presented in Table 1. Levels of agreement for the measures were high (80.0% to 97.5%). Adjusted  $\kappa$  were  $>0.8$  for two measures (adaptation to microbiology and day 3 bundle), representing almost perfect agreement. The antibiotic plan measure reached a  $\kappa$  of 0.75, representing substantial agreement, and the iv–oral switch documentation measure had a  $\kappa$  of 0.56, representing moderate agreement. The review of the diagnosis measure obtained a  $\kappa$  of 0, contrasting with a 97.5% level of agreement. For all our measures,  $\kappa$  and bias adjusted  $\kappa$  were almost identical.

### *Test of feedback of these process measures for improvement of the quality of care*

The measures were easily collected in around 15 min per week, and the collection of measures was sustained over 15 consecutive weeks.

Baseline data (weeks 1 and 2) showed room for improvement (Table 2). Run charts were used to feedback weekly results in real time in order to motivate change (Figure 1). Run charts show the course of data over time, simply by 'plotting the dots', and are useful to present and analyse the data.<sup>22</sup> There was

**Table 1.** Inter-rater reliability for measures of quality of care

	Observed agreement (%)	Expected agreement (%)	$\kappa$ (95% confidence interval)	Positive agreement $p_{\text{pos}}$	Negative agreement $p_{\text{neg}}$	Bias adjusted $\kappa$
Antibiotic plan	36/40 (90.0)	60.1	0.75 (0.52–0.98)	0.93	0.82	0.75
Review of the diagnosis	39/40 (97.5)	97.5	0.00 (0.00–0.00)	0.99	0.00	–0.01
Adaptation to positive microbiology	38/40 (95.0)	56.0	0.89 (0.73–1.04)	0.92	0.96	0.89
iv–oral switch documentation	32/40 (80.0)	54.0	0.57 (0.30–0.83)	0.85	0.71	0.56
Day 3 bundle	37/40 (92.5)	55.3	0.83 (0.65–1.01)	0.94	0.89	0.83

**Table 2.** Evolution of the completion (%) of the process measures over time

Measure	Week														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Antibiotic plan	25	0	62.5	50	100	86	60	80	60	60	67	100	83	50	100
Review of diagnosis	75	50	100	100	80	100	100	100	100	100	100	100	100	100	100
Adaptation to positive microbiology	100	75	100	100	100	100	100	100	NA	NA	100	100	100	100	100
iv–oral switch documentation	60	50	75	100	50	80	100	100	50	75	100	100	100	100	100
Day 3 bundle	13	0	62.5	50	60	86	60	80	40	60	67	100	83	50	100

NA, not applicable.

an initial improvement from 0% to 13% compliance in the first 2 weeks to 63% in week 3. For the next 4 weeks, compliance varied between 50% and 80%.

In an attempt to improve the documentation of antibiotic reassessment, we introduced a sticker (Figure 2) of the elements of the bundle, in week 7, using a plan-do-study-act approach. The sticker seemed to moderately improve the completion of the bundle, but definitely improved the ease of evaluation of compliance. In week 12, we achieved 100% compliance with the whole bundle for the first time, which was repeated in week 15, after a drop at 50% compliance in week 14 attributable to work overload.

## Discussion

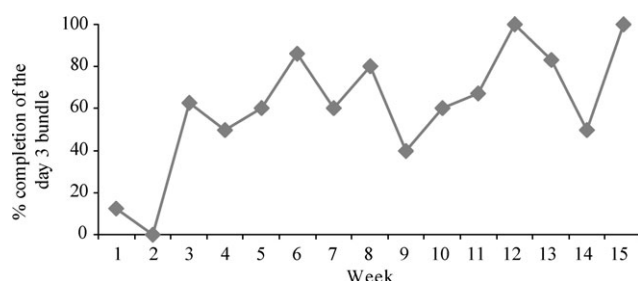
The first aim of this project was to identify and test key process measures in the reassessment of an inpatient empirical antibiotic prescription. We focused on the period 24–96 h after the antibiotic was started, thus around day 3 of treatment, since it is at that time that clinical evolution and the availability of culture results allow for the reassessment.<sup>1,3,4,20</sup> Two measures, adaptation to microbiology and iv–oral switch documentation, were widely cited in the literature.<sup>1,3,4,12,20,21</sup> Two others, antibiotic plan and review of the diagnosis, were considered essential by the three ID physicians, but were only cited explicitly in one French recommendation.<sup>3</sup> All four measures were grouped to form a care bundle, as suggested in the literature.<sup>12</sup> Our interpretation of the meaning of the term ‘care bundle’ was influenced by the Institute for Healthcare Improvement.<sup>23</sup> Their definition includes four criteria that distinguish care bundles from check lists or care pathways:

- The changes in the bundle are all necessary; so if there are four changes in the bundle and one is removed, the results would not be the same.
- The changes in the bundle are all based on evidence.
- The changes are clear-cut and involve all or nothing measurement such that each change can be recorded as a yes or no answer.
- Bundle changes occur in the same space and time, and can be delivered by the same clinical team.

In a question and answer session,<sup>23</sup> Carol Haraden says that each element of the bundle should be supported by evidence from randomized clinical trials. However, this is not true of the elements of the bundles that are listed by Carol Haraden, some of which are only supported by observational data. The Agency for Healthcare Research and Quality has a less exacting definition of evidence to support quality indicators: ‘A clinical practice guideline or other peer-reviewed synthesis of clinical evidence’ as well as ‘One or more research studies published in a National Library of Medicine indexed, peer reviewed journal’.<sup>24</sup> Our care bundle fulfils the criteria for defining several linked processes that need to be delivered to every patient. However, we accept that the separate elements in the bundle have varying levels of evidence to support them.

Therefore, we see our set of measures as a pragmatic first-step intervention to improve reassessment of empirical antibiotic therapies. Importantly, it should make the assessment of other quality indicators easier by decreasing the frequency of missing data, which is a common threat to the assessment of appropriateness of prescribing.<sup>25</sup> The need to record the information should also stimulate the team decisions around continuing antibiotic use.

## Measures to improve the reassessment of antibiotic prescriptions



**Figure 1.** Run chart showing the evolution of completion of the day 3 bundle measure.

Start date of antibiotic therapy: / / 2007		Date of review: / / 2007	
Review of initial <b>diagnosis</b>		Significant <b>bacteriological results</b>	
<b>Antibiotic plan</b>		Route of administration:	
Agent:			
Dosage:	Dosing interval:	Planned total duration: days	
If the patient receives iv antibiotic, is an <b>oral switch</b> possible?		Give the reason for that choice:	
<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Not applicable (not on iv)			

**Figure 2.** Sticker used to improve documentation of the measures in the medical notes.

The second aim of the study was to test the reliability of review of medical notes as a method of assessment of these measures of quality of care. A literature review identified studies of inter-rater reliability of prescribing appropriateness,<sup>25,26</sup> but only for long-term drugs, thus excluding antibiotic therapies. Other studies focused on specific infections, such as community-acquired pneumonia<sup>15,17</sup> or sepsis.<sup>10</sup> In epidemiological studies, kappa is the statistic most often used to measure agreement between two reviewers. Kappa compares the observed agreement with the agreement that could be expected to occur by chance.<sup>27</sup> The Cochrane Effective Practice and Organization of Care Group define a reliable primary outcome measure as at least 90% agreement between two or more raters or  $\kappa \geq 0.8$ .<sup>28</sup> However, there is considerable inconsistency in the level of  $\kappa$  that is used to indicate acceptable agreement, for example 0.7 for indicators of prescribing quality<sup>26</sup> and 0.4 for indicators of the quality of care of pneumonia.<sup>17</sup> Rather than selecting a single and somewhat arbitrary threshold for acceptability, it may be preferable to report agreement over the whole range of  $\kappa$  from 0 to 1. A  $\kappa$  of 0.00 is considered poor agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–1.00 almost perfect agreement.<sup>27,29</sup> Thus, two of our five measures reached  $\kappa$  values  $>0.8$  and one reached substantial agreement, demonstrating high inter-rater reliability. The iv–oral switch measure only reached moderate agreement ( $\kappa = 0.56$ ), probably because the reason for the switch was often implied in the medical notes, even if obvious for an ID specialist, leading to different interpretations for the medical notes' reviewers. The operational definition of the measure was then modified to be more precise: a clearly written reason for iv–oral switch is needed. The low  $\kappa$  score achieved for the review of the diagnosis measure was initially surprising because the level of

agreement was high (97.5%). This depression of the  $\kappa$  value is a paradoxical phenomenon, recognized to occur when data from pairs of raters are highly symmetrically imbalanced,<sup>16</sup> as in our case where nearly all diagnosis were rated as reviewed. The calculation of proportions of positive and negative agreement,  $p_{\text{pos}}$  and  $p_{\text{neg}}$ , respectively, allows interpretation of the practical effect of prevalence of positive and negative responses on  $\kappa$ .<sup>16,17</sup> The paradox of high agreement with a relatively low  $\kappa$  is explained by the wide discrepancy between  $p_{\text{pos}}$  and  $p_{\text{neg}}$ .<sup>16,17</sup>

The third aim was to test the collection of these measures on one ward and to try to improve practice. Collection of data proved to be feasible and sustainable. There was clear room for improvement, as shown by measures collected during the first 2 weeks. Then real-time audit and feedback was a powerful driver of improvement, as previously demonstrated.<sup>30</sup> Two measures (review of the diagnosis and adaptation to microbiology) rapidly achieved a 100% completion rate, which is not surprising on an ID ward. Others were more difficult to achieve, possibly because they seemed obvious to the ID physicians. Stickers were moderately useful to improve the documentation of these measures, probably because they were used by 5/8 members of the team, but definitely improved the ease of evaluation of compliance.

The main strengths of this study are that we have used a thorough approach for developing and implementing clinical performance measures,<sup>13</sup> and we have conducted a rigorous analysis of the reliability of key processes of care for reassessment of antibiotic prescriptions. Strengths in the analysis and reporting include the use of multiple reliability indicators and the reporting of raw data as well as per cent agreement and  $\kappa$ . Collecting, analysing and feeding back of the information are not hugely time-consuming and are very feasible. Involvement and engagement of the whole multidisciplinary healthcare team in this process are likely to further enhance the long-term sustainability of this process.<sup>19</sup>

This study has some limitations. Data were collected by the two Registrars who were also treating patients on the ward and were trained in infectious diseases. They could abstract data from their own medical notes. This quality improvement study is intended to be embedded into daily clinical practice, and involvement of the medical team is essential to ensure adherence to change. Given the high  $\kappa$  values we observed, it appears that potentially reviewing data from their own medical notes had a negligible influence on the data collected by the two ID Registrars. Our study was conducted on an ID ward and should be repeated in other settings, often where there is the greatest burden on suboptimal prescribing, in order to prove its external validity. However, given the high level of agreement found for all our measures, our operational definitions seem to be precise enough to be optimistic.

### Conclusions and further work

We have developed a reliable and valid set of measures to improve the documentation of inpatient empirical antibiotic prescription reassessment at around day 3. The data collection was achievable and sustainable on one ward. We plan to test these measures on other wards in our hospital and in other hospitals. There is ample room for improvement, and this 'day 3 bundle' could be useful as part of a multifaceted intervention<sup>21,31</sup> to improve antibiotic use in hospitals.



## Funding

Most data have been generated as part of the routine work of the hospital. C. P. was supported by grants from NHS Education Scotland, University of Medicine in Nice (France) and REDPIT association (REcherche et Développement en Pathologie Infectieuse et Tropicale, Nice, France).

## Transparency declarations

P. D. has been paid honoraria, consultancy fees or speaker fees by Johnson & Johnson, Optimer, Pfizer and Wyeth who are the manufacturers of various antibiotics and anti-infective products. D. N. declares that during the preparation of this document he was not in the employment of any pharmaceutical firm with interest in the content of the guidelines although he did accept appointments to the advisory boards of Pfizer, Wyeth, Johnson & Johnson and Novartis and has spoken at symposia supported by them. Other authors: none to declare.

C. P. designed the study, trained the doctors to collect the data, analysed the data and wrote the article. S. D. and I. A. have led the quality project on the ward, collected the data and reviewed the article. D. N. and P. D. have reviewed the study design, led the quality project on the ward and reviewed the article.

## References

- Dellit TH, Owens RC, McGowan JE Jr *et al.* Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. *Clin Infect Dis* 2007; **44**: 159–77.
- Nathwani D. Antimicrobial prescribing policy and practice in Scotland: recommendations for good antimicrobial practice in acute hospitals. *J Antimicrob Chemother* 2006; **57**: 1189–96.
- Société de Pathologie Infectieuse de Langue Française. Comment améliorer la qualité de l'antibiothérapie dans les établissements de soins? *Méd Mal Infect* 2002; **32**: 320–8.
- Senn L, Burnand B, Francioli P *et al.* Improving appropriateness of antibiotic therapy: randomized trial of an intervention to foster reassessment of prescription after 3 days. *J Antimicrob Chemother* 2004; **53**: 1062–7.
- Schouten JA, Hulscher ME, Trap-Liefers J *et al.* Tailored interventions to improve antibiotic use for lower respiratory tract infections in hospitals: a cluster-randomized, controlled trial. *Clin Infect Dis* 2007; **44**: 931–41.
- 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.
- Cocanour CS, Peninger M, Domonoske BD *et al.* Decreasing ventilator-associated pneumonia in a trauma ICU. *J Trauma* 2006; **61**: 122–9; discussion 129–30.
- Pronovost P, Needham D, Berenholtz S *et al.* An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006; **355**: 2725–32.
- Gao F, Melody T, Daniels DF *et al.* The impact of compliance with 6-hour and 24-hour sepsis bundles on hospital mortality in patients with severe sepsis: a prospective observational study. *Crit Care* 2005; **9**: R764–R770.
- Marwick C, Watts E, Evans J *et al.* Quality of care in sepsis management: development and testing of measures for improvement. *J Antimicrob Chemother* 2007; **60**: 694–7.
- van Kasteren ME, Mannien J, Kullberg BJ *et al.* Quality improvement of surgical prophylaxis in Dutch hospitals: evaluation of a multi-site intervention by time series analysis. *J Antimicrob Chemother* 2005; **56**: 1094–102.
- Cooke FJ, Holmes AH. The missing care bundle: antibiotic prescribing in hospitals. *Int J Antimicrob Agents* 2007; **30**: 25–9.
- 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.
- Hofer TP, Bernstein SJ, DeMonner S *et al.* Discussion between reviewers does not improve reliability of peer review of hospital quality. *Med Care* 2000; **38**: 152–61.
- Williams SC, Watt A, Schmaltz SP *et al.* Assessing the reliability of standardized performance indicators. *Int J Qual Health Care* 2006; **18**: 246–55.
- Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; **43**: 543–9.
- Scinto JD, Galusha DH, Krumholz HM *et al.* The case for comprehensive quality indicator reliability assessment. *J Clin Epidemiol* 2001; **54**: 1103–11.
- DAG-stats 2007. *Diagnostic and Agreement Statistics*. [http://www.mhri.edu.au/biostats/DAG\\_Stat/](http://www.mhri.edu.au/biostats/DAG_Stat/) (19 February 2008, date last accessed).
- Pulcini C, Crofts S, Campbell D *et al.* Design, measurement, and evaluation of an education strategy in the hospital setting to combat antimicrobial resistance. *Dis Manage Health Outcomes* 2007; **15**: 151–63.
- Stephenson J. CDC campaign targets antimicrobial resistance in hospitals. *JAMA* 2002; **287**: 2351–2.
- Vogtlander NP, Van Kasteren ME, Natsch S *et al.* Improving the process of antibiotic therapy in daily practice: interventions to optimize timing, dosage adjustment to renal function, and switch therapy. *Arch Intern Med* 2004; **164**: 1206–12.
- Lloyd R. *Quality Health Care: A Guide to Developing and Using Indicators*. Sudbury, MA, USA: Jones and Bartlett, 2004.
- Haraden C. *What Is a Bundle?* <http://www.ihl.org/IHI/Topics/CriticalCare/IntensiveCare/ImprovementStories/WhatIsaBundle.htm> (19 February 2008, date last accessed).
- National Quality Measures Clearinghouse. *Measure Validity*. [http://www.qualitymeasures.ahrq.gov/resources/measure\\_domains.aspx](http://www.qualitymeasures.ahrq.gov/resources/measure_domains.aspx) (19 February 2008, date last accessed).
- Tully MP, Cantrill JA. The validity of explicit indicators of prescribing appropriateness. *Int J Qual Health Care* 2006; **18**: 87–94.
- Tully MP, Javed N, Cantrill JA. Development and face validity of explicit indicators of appropriateness of long term prescribing. *Pharm World Sci* 2005; **27**: 407–13.
- Kunac DL, Reith DM, Kennedy J *et al.* Inter- and intra-rater reliability for classification of medication related events in paediatric inpatients. *Qual Saf Health Care* 2006; **15**: 196–201.
- Cochrane Effective Practice Organisation of Care Group 2002. *The Data Collection Checklist*. <http://www.epoc.uottawa.ca/checklist2002.doc> (19 February 2008, date last accessed).
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74.
- Jamtvedt G, Young JM, Kristoffersen DT *et al.* Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2006; CD000259.
- Gyssens IC, Blok WL, van den Broek PJ *et al.* Implementation of an educational program and an antibiotic order form to optimize quality of antimicrobial drug use in a department of internal medicine. *Eur J Clin Microbiol Infect Dis* 1997; **16**: 904–12.